On Carrier Accumulation, and the Properties of Certain Semiconductor Junctions†

By J. B. Gunn

Royal Radar Establishment, Great Malvern‡

[Received August 21, 1957]

§ 1. Introduction

In a recent paper, Low (1955) has shown that, in the region of a semiconductor near an electrode boundary, there are four types of solution to the equations governing hole and electron flow in an electric field. Which of the types is to be chosen in any given case depends on the composition of the current at the boundary, and on the direction of the electrical forces on the minority carriers. One of these types of solution relates to a situation, termed by Low *carrier accumulation*, in which minority carriers are drawn by the electric field towards the boundary, across which, however, they are able to pass only with difficulty. The carrier density at the boundary then becomes larger than the equilibrium value.

Unlike the other phenomena considered in the paper quoted, carrier accumulation had not then been observed experimentally, as no suitable electrode structure was known. It is the purpose of the present paper to show that a junction between relatively pure and relatively impure regions of a semiconductor of uniform conductivity type possesses a degree of impermeability to minority carriers which should permit carrier accumulation to be observed in the lightly doped section. In addition, it will be shown that carrier accumulation can lead to other effects which render the problem of more than academic interest.

In order that the underlying physical principles shall not be obscured by the mathematical detail, a brief description will be given in this introduction of the approach, methods, and results of the paper. The analysis starts from five basic equations, of which two are relations between the hole and electron currents and the corresponding density and potential gradients, two are continuity equations for the currents, and the fifth is Poisson's equation written for the space-charge of holes, electrons, and ionized impurity atoms. The equations are written in dimensionless form, since this eliminates a number of constants which are not of primary importance, and permits a greater degree of conciseness in the resulting analysis. It might be thought that a space-charge equation was necessary only in dealing with a region of the semiconductor in which the electric

† Communicated by Dr. A. F. Gibson.

‡ Now at The University of British Columbia, Vancouver, Canada.

field was large. That this is not so may be seen by observing, firstly, that such an equation provides the only coupling between the equations for hole and for electron flow; and, secondly, that space-charge must arise in any current-carrying conductor in which the conductivity σ is dependent on position. For, if ψ is the potential, J the current, and κ the dielectric constant, the space-charge ρ is given by

$$4\pi\rho/\kappa = -\nabla^2\psi = \nabla \cdot (J/\sigma) = -(1/\sigma^2)J \cdot \nabla\sigma \neq 0.$$

It follows from the necessity of taking account of space-charge that a general solution can be obtained by treating the five basic equations as a system of simultaneous differential equations, in which the potential is unknown; this is to be contrasted with the more usual restricted solutions in which the potential field is assumed to be known. It is not possible to obtain an exact analytical solution along these lines, however, and it must be decided for a given problem whether exact numerical or approximate analytical results will be the more useful. The latter may be expected to give the clearer insight into the nature of the solution, and we shall therefore follow this method where possible. An attempt will be made, none the less, to delay the introduction of approximations as long as possible so that the results quoted are as general as can be obtained without recourse to numerical computation or the use of expressions so cumbersome as to obscure the general nature of the results.

The analysis given in §2 begins by quoting the five basic equations and showing that they can be combined into a single equation of a relatively tractable type under three different assumptions, the assumptions of *neutrality*, *quasi-neutrality*, and *space-charge*. This equation is expressed in terms of the total current rather than the potential gradient; because it is solenoidal, the current vector field is usually among the data of a particular problem. In the first two assumptions, which are capable of dealing with non-equilibrium situations, it is supposed that the net space-charge density is always a negligible fraction of the density of charge of one sign. Space-charge effects then appear only in the term $\nabla^2\psi$, which can be eliminated from the set of equations. The resulting differential equation for the carrier density is then of the second order and non-linear, and can be treated analytically only by imposing further restrictions; limiting the variation of carrier density leads to a second order linear equation, while a first order non-linear equation is obtained by neglecting the effect of recombination of holes and electrons. Both of the resulting equations have solutions in terms of elementary functions. The assumption of small density variations is equivalent to that of uniform conductivity and therefore of zero space-charge density, or neutrality. It should be noted that space-charge effects still produce a strong coupling between holes and electrons even in this case. The relaxation of the restriction on density variations permits the development of some space-charge, but its density usually remains small and the assumption is that of quasi-neutrality.

A simple solution under conditions of gross space-charge can only be derived if both the hole and electron currents are assumed to be zero, which is nearly equivalent to the assumption of thermal equilibrium. Even so, the resulting second-order equation can, in the general case, only be integrated once in terms of known functions. The results do, however, provide a useful guide to non-equilibrium conditions.

The next step is to apply the differential equation just discussed to a physical system consisting of a single crystal of semiconductor in which a pure or lightly doped region (l-region) is separated from an impure or heavily doped region (h-region) of the same conductivity type by an abrupt change of impurity concentration or l–h *junction*. For simplicity, it is assumed that the concentration change is discontinuous; this is a convenient mathematical fiction which is a good approximation for most practical types of junction. It turns out that when accumulation takes place in such a structure, the neutral assumption is the best for the h-region and the quasi-neutral for the l-region. Near the junction itself there occurs a very thin region in which the space-charge assumption is appropriate. Little generality is lost by restricting the analysis to one-dimensional systems, as most of the cases occurring in practice will be of this type.

The analysis of the h-region given in §3 is straight-forward and is concerned mainly with well-known neutral solutions for carrier flow. The solutions are obtained in convenient general forms and a new approximate solution is given for the problem of carrier flow in a system with a radial current flow field. This section also introduces the concept of the *transport* velocity, which is defined as the ratio of minority carrier current density at a point to the excess of the minority carrier density there over its thermal equilibrium value. The transport velocity plays a similar part in the theory of carrier flow to that of the concept of impedance in electromagnetic theory. Both quantities serve to characterize, by a single number, whole families of solutions of the respective differential equations, and both provide a powerful method of treating boundary conditions. The section concludes by calculating for a number of structures the transport velocity v_{0h} measured at a point in the h-region very near the l–h junction. This velocity can easily be quite small.

Section 4 discusses the potential distribution within the space-charge region and the continuity relations between the carrier densities on the two sides of the region, derived on the assumption that the hole and electron currents are separately zero. An interesting property of the potential distribution, which is given in a general from applicable to both l–h and p–n junctions, is that the potential rise in the former is confined almost entirely to the l-region. The potential in the h-region never exceeds kT/q; this is in contrast to a p–n junction, in which no such limiting value is found. It is also shown that the space-charge region of an l–h junction is very thin, and it is this fact which enables the results given to be extended, as a good approximation, to cover the

finite currents likely to be encountered in practice. Very large electric fields may occur within the space-charge region.

The thinness of the space-charge region leads to continuity relations across the junction which are, firstly, that the hole and electron currents are separately conserved, and secondly, that the product of the hole and electron densities on one side must be equal to that on the other. Together with the normal neutral or quasi-neutral equation for the difference of the densities, these lead to a relation between v_{0l}, the transport velocity in the l-region near the junction, and v_{0h}. The ratio v_{0l}/v_{0h} may be as small as the ratio of the impurity densities in the two regions, and since v_{0h} may itself be a small quantity, it is clear that v_{0l} may, under suitable conditions, be reduced to a very small value indeed. From this it follows that an l–h junction may be made very nearly impermeable to minority carriers entering or leaving the l-region, while remaining completely permeable to majority carriers. Because it is a measure of the departure from semi-permeable behaviour, we shall refer to v_{0l} as the *leakage velocity* of the junction.

This semi-permeability is the characteristic property of an l–h junction, and leads to the possibility of observing a number of new effects†. Among these are the accumulation which should arise when a minority carrier current is forced across such a junction, and a resulting increase in current gain which should result when accumulation takes place in a class of devices, *conductance transistors*, allied to the filamentary transistor of Shockley *et al.* (1949). The way in which this current gain comes about may be seen qualitatively by observing that in a region where carriers are accumulating the diffusion resulting from the density gradient common to both kinds of carrier assists the electric field in producing a large current of majority carriers but opposes its action on the minority carriers, the current of which is therefore small. The current gain, which is the derivative of the total current with respect to the minority carrier current, at constant potential difference between two specified points in the system, may therefore exceed considerably the normal value for a filamentary transistor. The accumulation region ordinarily extends only a short distance from the junction, however, and in order to realize a large current gain it is necessary that the resistance of this region should be the dominant term in determining the output resistance of the device. This condition is most easily met in a system having spherical symmetry, and such systems are therefore considered in some detail.

In the analysis of these effects given in §5 it is found necessary to use the quasi-neutral approximation, not only because the carrier density may become very large near an accumulating junction, but also because the alternative assumption of neutrality implies that the conductivity of the semiconductor is not appreciably altered by the presence of additional carriers, and is therefore fundamentally incapable of dealing with conductivity modulation. Recombination effects are thus excluded from exact analytical treatment, but this is not a serious matter as adequate approximations can be derived from the recombination-free solutions. These solutions include two very general equations whose validity is restricted only by the neglect of recombination. The first, relating the local potential and carrier density, may be regarded as an extension of Boltzmann's equation to non-equilibrium conditions in an ambipolar semiconductor. The second describes the spatial distribution of potential in a generalized one-dimensional system, and includes as parameters the total and minority-carrier currents and one boundary condition. It thus forms the basis for a general theory of conductance transistors in which the dependence of the current gain on boundary conditions is brought out.

Corrections for recombination are discussed in §6. It is shown that, subject to a certain restriction, the previously derived exact solutions for the potential and carrier density remain approximately true even when recombination effects are included, and the minority carrier current lost by recombination can therefore be calculated to a good approximation. The restriction is that the recombination current must be much less than the total current in the system; the analysis, which is given in a generalized form restricted only to one-dimensional systems, therefore includes most cases of interest. The relative importance of recombination compared with junction leakage in determining the rate of carrier loss depends on a parameter which, for a planar l-region of thickness x_e and lifetime τ_l, is equal to $2v_{0l}\tau_l/x_e$. If this is much greater than unity, the important mechanism is leakage; if less, it is recombination.

Section 7 contains a number of applications of the theory given in the preceding sections. These include a proposed method for the measurement of the most important parameter of an l–h junction, its leakage velocity, and expressions for the current gain and collector and emitter impedances of a simply constructed type of conductance transistor with a large current gain.

The paper concludes, in §8, with a discussion of the foregoing sections, and some topics not previously considered.

§ 2. The Basic Equations

The five simultaneous differential equations which form the basis of the theory of carrier flow can be combined into a single equation in one variable only if certain approximations are made. This equation has three different forms, according to the approximation employed, and it is with the derivation of these forms that the present section is concerned. In order to simplify the basic equations, it will be assumed at the outset that the temperature T is constant, that the semiconductor is not degenerate, and that the electric field and its time derivative are not so large that the charge carriers are out of equilibrium with the lattice (Shockley 1951).

† Compare the inverse type of semi-permeability characteristic of a p–n junction which may be made permeable to minority, and nearly impermeable to majority, carriers.

In writing down the basic equations we shall use a system of reduced units (Gunn 1954) in which the unit of carrier density is n_i, the intrinsic density, and the unit of potential is the thermal potential kT/q. The unit of length, $(kT\kappa/4\pi q^2 n_i)^{1/2}$, is related to the Debye length in the medium, while the unit of time, $\kappa/(4\pi\mu q n_i)$, is based on the intrinsic dielectric relaxation time (κ=dielectric constant, μ=hole mobility, q=absolute electronic charge). In this notation, the equations have the dimensionless forms:

$$j_n = bn\nabla\psi - b\nabla n \quad (2.1)$$

$$j_p = -p\nabla\psi - \nabla p \quad (2.2)$$

$$\nabla \cdot j_n = -\partial n/\partial t + G(p, n) \quad (2.3)$$

$$\nabla \cdot j_p = -\partial p/\partial t + G(p, n) \quad (2.4)$$

$$-\nabla^2\psi = p - n - \Omega. \quad (2.5)$$

Here n=electron density; p=hole density; j_n=electron flux (density × velocity); j_p=hole flux; ψ=electrostatic potential; b=electron mobility/hole mobility; G=net rate of hole–electron pair generation (or recombination) and Ω=net impurity density (acceptors minus donors).

It should be noted that eqns. (2.1) to (2.4) can be solved independently *only when ψ is a known function of position*, which is not generally the case. The introduction of ψ as an unknown variable converts the system into one of simultaneous equations, which can only be solved for p and n by the elimination of $\nabla\psi$. The alternative procedure of assuming that space-charge effects are negligible is incorrect, since it excludes from consideration the vital coupling between holes and electrons, and may lead to results which are physically absurd.

Before using eqns. (2.1) to (2.5) to obtain more useful relations it is necessary to define the quantities in terms of which these relations will be obtained. These are: the total current density $J \equiv j_p - j_n - \partial/\partial t(\nabla\psi)$ the effective carrier density $N \equiv p + bn$, and the density difference $Q \equiv p - n$. The total current density is a particularly useful quantity since, unlike $\nabla\psi$, it is always a solenoidal vector ($\nabla \cdot J = 0$). Consequently, in certain simple flow patterns (all those in which the flow lines of J are straight, for instance), it is a known function of position.

By subtracting eqn. (2.1) from eqn. (2.2), the following expression for $\nabla\psi$ is obtained in terms of J, Q and N:

$$-\nabla\psi = \frac{1}{N}\left(J + \frac{\partial}{\partial t}\nabla\psi\right) - \frac{b-1}{b+1}\frac{1}{N}\nabla N + \frac{2b}{b+1}\frac{1}{N}\nabla Q. \quad (2.6)$$

Alternatively, $\nabla\psi$ may be eliminated between (2.1) and (2.2), yielding

$$(b+1)j_p = (1 + bQ/N)(J + \partial/\partial t\nabla\psi) - \left\{\frac{2b}{b+1} + \frac{b(b-1)}{b+1}\frac{Q}{N}\right\}\nabla N - \left\{\frac{b(b-1)}{b+1} - \frac{2b^2}{b+1}\frac{Q}{N}\right\}\nabla Q. \quad (2.7)$$

Finally, on taking the divergence of equ. (2.7) and putting $\partial\Omega/\partial t = 0$, we obtain

$$\begin{aligned}(b+1)\nabla \cdot j_p = &-\frac{\partial N}{\partial t} - b\frac{\partial Q}{\partial t} + (b+1)G \\ = &-b\frac{Q}{N^2}\left(J + \frac{\partial}{\partial t}\nabla\psi\right)\cdot\nabla N \\ &-\left\{\frac{2b}{b+1} + \frac{b(b-1)}{b+1}\frac{Q}{N}\right\}\nabla^2 N \\ &+\frac{b(b-1)}{b+1}\frac{Q}{N^2}(\nabla N)^2 \\ &-\frac{1}{N}\left\{\frac{b(b-1)}{b+1} + \frac{2b^2}{b+1}\frac{Q}{N}\right\}\nabla Q \cdot \nabla N \\ &+\frac{b}{N}\left(J + \frac{\partial}{\partial t}\nabla\psi\right)\cdot\nabla Q \\ &-\left\{\frac{b(b-1)}{b+1} - \frac{2b^2}{b+1}\frac{Q}{N}\right\}\nabla^2 Q \\ &+\frac{2b^2}{b+1}\frac{1}{N}(\nabla Q)^2 - \left(1 + b\frac{Q}{N}\right)\frac{\partial Q}{\partial t}. \quad (2.8)\end{aligned}$$

These equations are, of course, rigorously true for both p- and n-type material †.

Under conditions where Ω is constant, and the space-charge density small, but not negligible, each term in ∂Q may be neglected by comparison with the corresponding larger term in ∂N. Thus Ω/N may be written for Q/N, and the equations reduced to approximate forms in N alone:

$$-\nabla\psi = \frac{1}{N}(J + \partial/\partial t\,\nabla\psi) - \frac{b-1}{b+1}\frac{1}{N}\nabla N \quad (2.9)$$

$$(b+1)j_p = \left(1 + \frac{b\Omega}{N}\right)\left(J + \frac{\partial}{\partial t}\nabla\psi\right) - \left\{\frac{2b}{b+1} + \frac{b(b-1)}{(b+1)}\cdot\frac{\Omega}{N}\right\}\nabla N \quad (2.10)$$

$$\begin{aligned}(b+1)\nabla \cdot j_p = &-\frac{\partial N}{\partial t} + (b+1)G \\ = &-\frac{b\Omega}{N^2}\left(J + \frac{\partial}{\partial t}\nabla\psi\right)\cdot\nabla N \\ &-\left\{\frac{2b}{b+1} + \frac{b(b-1)}{(b+1)}\frac{\Omega}{N}\right\}\nabla^2 N \\ &+\frac{b(b-1)}{(b+1)}\frac{\Omega}{N^2}(\nabla N)^2. \quad (2.11)\end{aligned}$$

While the necessary conditions will be found to prevail in many cases where the space-charge arises only from a non-equilibrium distribution of carriers, they do not, of course, occur in a region of gross space-charge such as an exhaustion layer. In this case, however, the electric field

† A somewhat similar discussion has been given by Van Roosbroeck (1950, 1953)

resulting from the space-charge will be so large that any field resulting from the flow of carriers may be neglected by comparison. Equations (2.1) and (2.2) may then be integrated directly, with j_n and j_p neglected, to give

$$n=\exp \psi \qquad (2.12)$$

$$p=\exp(-\psi) \qquad (2.13)$$

where the zero of potential has been chosen to make $p=n=1$ when $\psi=0$.

On substituting these relations in eqn. (2.5), we obtain the approximate equation

$$\nabla^2\psi=2\sinh\psi+\Omega \qquad (2.14)$$

which is, of course, rigorously true when

$$j_p=j_n=0.$$

Equations (2.9)–(2.11) may be termed the *quasi-neutral* approximations, whereas eqn. (2.14) is the *space-charge* approximation. In neither case is the important term $(Q-\Omega)$ neglected, as may sometimes be justified when the ratio of majority to minority carriers is large. Since consideration of space-charge is in this case omitted altogether, it may be called the *neutral* approximation; it is the only case in which recombination phenomena can be analysed exactly.

§ 3. Carrier Flow in the Heavily Doped Region

In this section we shall calculate the minority carrier flow in the heavily doped region (h of fig. 1) of the junction. In general, recombination will be important in this region, both in the volume of the semiconductor and at the surface S. Consequently, the solution must start from eqn. (2.11), since only this equation contains G explicitly. The circumstance of heavy doping permits a considerable simplification to be made in this equation, which is fortunate since the exact form has no analytical solutions.

If $N_{\infty h}$ is the equilibrium value of N corresponding to the impurity density Ω_h in this region, then, provided $|N-N_{\infty h}|$ is always much less than $|\Omega_h|$, eqns. (2.10) and (2.11) can be written, in the static state,

$$(b+1)j_p=\left(1+b\frac{\Omega_h}{N_{\infty h}}+\frac{N-N_{\infty h}}{N_{\infty h}}\right)J - \left\{\frac{2b}{b+1}+\frac{b(b-1)}{b+1}\frac{\Omega_h}{N_{\infty h}}\right\}\nabla N \qquad (3.1)$$

$$(b+1)G=-b\frac{\Omega_h}{N_{\infty h}^2}J\,.\,\nabla N - \left\{\frac{2b}{b+1}+\frac{b(b-1)}{b+1}\frac{\Omega_h}{N_{\infty h}}\right\}\nabla^2 N \qquad (3.2)$$

The validity of the substitution of $N_{\infty h}$ for N is obvious, subject to the given restriction; that of the omission of $(\nabla N)^2$ will be proved later. Subject also to this restriction, the process of carrier recombination can be exactly represented by a single lifetime τ_h, so that G is given by $G=N-N_{\infty h}/(b+1)\tau_h$.

To fix our ideas, we shall now restrict the discussion to n-type material. Then, provided $|\Omega_h|$ is much greater than 1, $N_{\infty h}$ is approximately equal to $-b\Omega_h$, and eqn. (3.1) becomes

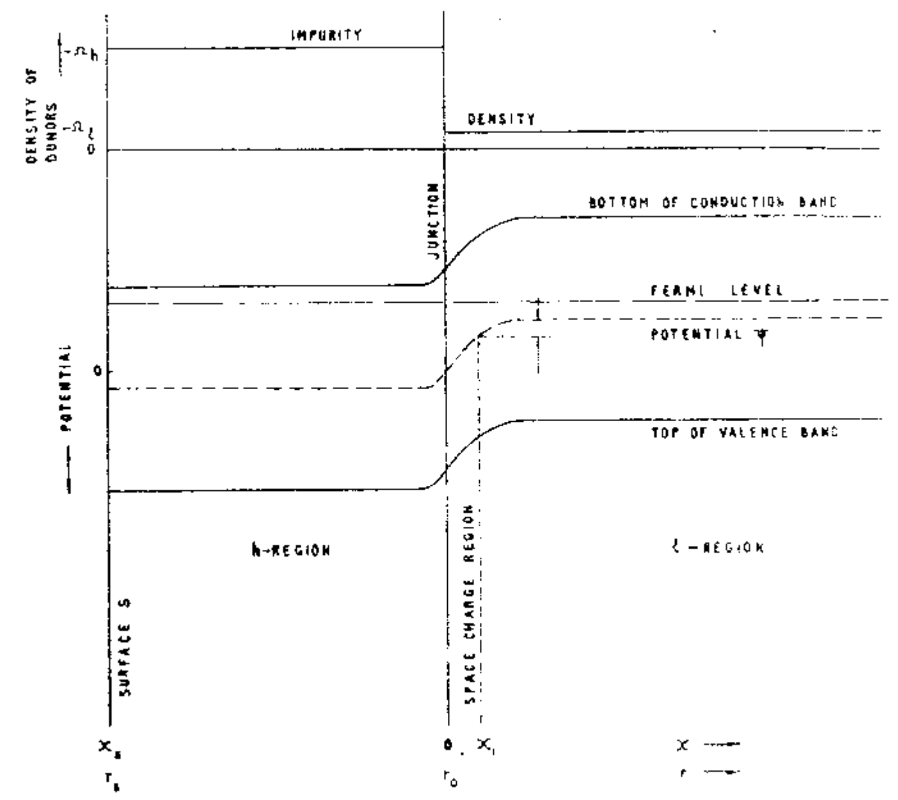
$$(b+1)j_p=-\frac{N-N_{\infty h}}{b\Omega_h}J-\nabla N \qquad (3.3)$$

while eqn. (3.2) becomes

$$\frac{N-N_{\infty h}}{\tau_h}-\frac{1}{b\Omega_h}J\,.\,\nabla N-\nabla^2 N=0. \qquad (3.4)$$

(The corresponding equations for p-type material are obtained by putting $N_{\infty h}\sim\Omega_h$.)

Fig 1



3.1. *Planar Flow*

Near an infinite plane junction, the solutions of eqn. (3.4) will have planar symmetry, and can consequently be expressed in terms of a single space coordinate x. (We shall write J for Jx, and so on, as no confusion will arise.) The equation then becomes

$$\frac{N-N_{\infty h}}{\tau_h}-\left(\frac{J}{b\,\Omega_h}\right)\frac{dN}{dx}-\frac{d^2N}{dx^2}=0. \qquad (3.1.1)$$

This has the solution

$$N-N_{\infty h}=C_1 \exp\frac{x}{L_1}+C_2 \exp\frac{x}{L_2} \quad . \quad . \quad . \quad . \quad (3.1.2)$$

where

$$L_1^{-1}=-\frac{J}{2b\Omega_h}-\left(\frac{J^2}{4b^2\Omega_h^2}+\frac{1}{\tau_h}\right)^{1/2}$$

$$L_2^{-1}=-\frac{J}{2b\Omega_h}+\left(\frac{J^2}{4b^2\Omega_h^2}+\frac{1}{\tau_h}\right)^{1/2}.$$

If J is taken to be negative, then, with the previously assumed negative Ω_h, L_1 is the *upstream* and L_2 the *downstream* diffusion length (Low 1955). In order to obtain an expression more suited to the fitting of boundary conditions, we now introduce a new variable, the transport velocity v, which is defined in general by $v=(b+1)j_p/(N-N_\infty)$. In the present case, by combining the appropriate form of eqn. (3.3) with eqn. (3.1.2), we find that v is given by

$$v+J/b\Omega_h=-\frac{1/L_1+1/L_2(C_2/C_1)\exp\{x(L_2^{-1}-L_1^{-1})\}}{1+C_2/C_1\exp\{x(L_2^{-1}-L_1^{-1})\}} \quad (3.1.3)$$

The surface S, placed at $x=x_s$, is assumed to be characterized by the recombination velocity v_s $(v_s<0)$, so that the boundary condition is $v=v_s$ at $x=x_s$. By combining this with eqn. (3.1.3.), and writing $v_s'=v_s+(J/b)\Omega_h$ for brevity, we find that v_{0h}, the transport velocity on the h-side of the junction at $x=0$, is given by

$$v_{0h}+\frac{J}{b\Omega_h}=-\frac{1}{L_2}\,\frac{L_2/L_1\exp\{x_s(L_2^{-1}-L_1^{-1})\}-(v_s'+L_1^{-1})/(v_s'+L_2^{-1})}{\exp\{x_s(L_2^{-1}-L_1^{-1})\}-(v_s'+L_1^{-1})/(v_s'+L_2^{-1})}. \quad (3.1.4)$$

In subsequent sections, we shall require the following special forms of eqn. (3.1.4):

$$J=0, \text{ therefore } -L_1=L_2=\tau_h^{1/2}$$

$$v_{0h}=\frac{1}{\tau_h^{1/2}}\,\frac{\exp(2x_s\tau_h^{-1/2})+(v_s-\tau_h^{-1/2})/(v_s+\tau_h^{-1/2})}{\exp(2x_s\tau_h^{-1/2})-(v_s-\tau_h^{-1/2})/(v_s+\tau_h^{-1/2})} \quad . \quad (3.1.5)$$

$$x_s(L_2^{-1}-L_1^{-1})\ll -1$$

$$v_{0h}=-\frac{J}{2b\Omega_h}-\frac{1}{\tau_h^{1/2}}\left(1+\frac{J^2\tau_h}{4b^2\Omega_h^2}\right)^{1/2} \quad . \quad . \quad . \quad (3.1.6)$$

$$-v_s\gg L_2^{-1}$$

$$v_{0h}=-\frac{J}{b\Omega_h}-\frac{1}{L_2}\,\frac{L_2/L_1\exp\{x_s(L_2^{-1}-L_1^{-1})\}-1}{\exp\{x_s(L_2^{-1}-L_1^{-1})\}-1}. \quad . \quad (3.1.7)$$

The neglect of $(\nabla N)^2$ in eqn. (3.2) can now be justified, since we see that

$$\frac{\Omega_h}{N_{\infty h}^2}\left(\frac{d}{dx}C\exp\frac{x}{L}\right)^2\bigg/\frac{d^2}{dx^2}C\exp\frac{x}{L}=\frac{\Omega_h}{N_{\infty h}}\cdot\frac{C\exp x/L}{N_{\infty h}}=\frac{\Omega_h}{N_{\infty h}}\cdot\frac{N-N_{\infty h}}{N_{\infty h}}$$

and this is small when $(N-N_{\infty h})/N_{\infty h}$ is small.

3.2. *Radial Flow*

If the junction covers a very small area of a planar surface, the flow near the junction will have approximately spherical symmetry, and little inaccuracy will be introduced by analysing the system as one of radial flow, occupying a solid angle of 2π. The single space variable is then the radius r, and the appropriate form of eqn. (3.4) is thus

$$\frac{N-N_{\infty h}}{\tau_h}-\frac{J_r}{b\Omega_h}\frac{dN}{dr}-\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dN}{dr}\right)=0.$$

If $I=2\pi r^2 J_r$, I is a constant, and we have

$$\frac{N-N_{\infty h}}{\tau_h}-\frac{I}{2\pi b\Omega_h}\frac{1}{r^2}\frac{dN}{dr}-\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dN}{dr}\right)=0. \quad . \quad . \quad (3.2.1)$$

Before attempting to solve this, we make the substitutions

$$W=(N-N_{\infty h})r\exp\left(\frac{-I}{4\pi b\Omega_h}\cdot\frac{1}{r}\right)$$

$$Z=r/\tau_h^{1/2}$$

$$F=\frac{I}{4\pi b\Omega_h\tau_h^{1/2}}.$$

Equation (3.2.1) now becomes

$$\frac{d^2W}{dZ^2}-\left(\frac{F^2}{Z^4}+1\right)W=0. \quad . \quad . \quad . \quad (3.2.2)$$

This equation has, unfortunately, no analytical solutions, but approximate solutions are readily obtained when either (*a*) $F^2/Z^4\ll 1$, or (*b*) $F^2/Z^4\gg 1$. In case (*a*), the solution is

$$W\approx C_A\exp Z+C_B\exp(-Z) \quad . \quad . \quad . \quad . \quad (3.2.3)$$

while in case (*b*) (Riccatis equation), which is not of immediate interest, it is

$$W\approx C_D Z\exp(F/Z)+C_E Z\exp(-F/Z).$$

The resulting expression for $N-N_{\infty h}$ in case (*a*) is thus

$$N-N_{\infty h}\simeq C_1 Z^{-1}\exp(F/Z-Z)+C_2 Z^{-1}\exp(F/Z+Z). \quad (3.2.4)$$

The two terms in this equation again represent upstream and downstream diffusion.

In calculating v, we shall restrict consideration to the case where v_s is very large at the surface S, which we suppose to be a hemisphere of radius r_s. This corresponds to the boundary condition $N-N_{\infty h}=0$ at $r=r_s$, and well represents the conditions at the surface of the metal rich phase of an alloyed l-h junction. The calculation for arbitrary v_s, following the lines of the corresponding calculation in § 3, presents no difficulties, but renders the resulting expressions needlessly cumbersome for the use which we shall make of them.

On making use of this boundary condition, eqn. (3.2.4), and the appropriate form of eqn. (3.3), we find that v is given by

$$v=-\frac{I}{4\pi b\Omega_h}\frac{1}{r^2}-\frac{1}{\tau_h^{1/2}}\left\{\coth\left(\frac{r-r_s}{\tau_h^{1/2}}\right)-\frac{\tau_h^{1/2}}{r}\right\}.$$

If we take the l–h junction to be a hemisphere of radius r_0, we see that

$$v_{0h}=-\frac{I}{4\pi b\Omega_h}\cdot\frac{1}{r_0^2}-\frac{1}{\tau_h^{1/2}}\left\{\coth\left(\frac{r_0-r_s}{\tau_h^{1/2}}\right)-\frac{\tau_h^{1/2}}{r_0}\right\}. \qquad (3.2.5)$$

In the special case where (r_0-r_s) is much less than both r_0 and $\tau_h^{1/2}$, this equation may be approximated by:

$$v_{0h}\simeq-\frac{I}{4\pi b\Omega_h}\cdot\frac{1}{r_0^2}-\frac{1}{r_0-r_s}. \qquad (3.2.6)$$

We may remark in passing that the method of approximation which we have used above retains first order terms in $I/(4\pi b\Omega_h r^2)$, neglecting only second- and higher-order terms. It is therefore more accurate than the usual treatment which neglects the first-order term in eqn. (3.2.1). Finally, similar considerations to those given in the last section may be involved to justify the neglect of $(dN/dr)^2$ in this case also.

§ 4. Potential Distribution in the Space-Charge Region, and Continuity between the l- and h-Regions

Even in thermal equilibrium the electrostatic potential will have different values deep in the l- and h-regions, since the Fermi level must be constant everywhere. The electric field necessary to produce this potential difference results from gross space-charge occurring near $x=0$, and we now analyse the conditions in this, the space-charge region. Since an exact analytical solution is impossible for non-equilibrium conditions, we shall base the analysis on the approximate space-charge eqn. (2.14), calculating first the situation in thermal equilibrium. Only a simple modification is then necessary to take account of non-equilibrium conditions.

4.1. *The Space-Charge Region in Equilibrium*

Because a significant density of space-charge extends over only a short distance near the junction, we shall be justified in using a planar geometry, even though the system as a whole may be of radial type. Near $x=0$ the potential distribution has the form shown in fig. 1, with space-charge in both l- and h-regions.

In one dimension, eqn. (2.14) has the form

$$\frac{d^2\psi}{dx^2}=2\sinh\psi-\Omega$$

which may be integrated once by the substitution $u=d\psi/dx$, giving

$$\tfrac{1}{2}u^2=2\cosh\psi-\Omega\psi+C. \qquad (4.1.1)$$

The values of Ω and C will, of course, be different in the l- and h-regions. If $\psi_{\infty h}$ and $\psi_{\infty l}$ are the values of ψ deep in the h- and l- regions respectively, they are given by

$$2\sinh\psi_{\infty h}=-\Omega_h \text{ and } 2\sinh\psi_{\infty l}=-\Omega_l.$$

These lead to the boundary conditions

$$u=0 \text{ when } \psi=\psi_{\infty h} \text{ and when } \psi=\psi_{\infty l}.$$

Hence C has the values

$$\left.\begin{aligned}C_h&=2\psi_{\infty h}\sinh\psi_{\infty h}-2\cosh\psi_{\infty h}\\ C_l&=2\psi_{\infty l}\sinh\psi_{\infty l}-2\cosh\psi_{\infty l}\end{aligned}\right\} \qquad (4.1.2)$$

in the h- and l-regions respectively.

The exact shape of the potential rise can only be obtained by integrating eqn. (4.1.1) a second time, and this can only be done numerically (see below). Additional information about the distribution can, however, be obtained by matching solutions at the origin. Since an infinite density of space-charge is not possible, both ψ and u must be continuous on passing from h to l material. If u_h and u_l are the solutions in h- and l-regions, and ψ_0 is the potential at $x=0$, then we have $u_h(\psi_0)=u_l(\psi_0)$, and hence, from eqn. (4.1.1),

$$\begin{aligned}(\psi_0-\psi_{\infty h})\sinh\psi_{\infty h}&+\cosh\psi_{\infty h}\\ &=(\psi_0-\psi_{\infty l})\sinh\psi_{\infty l}+\cosh\psi_{\infty l}\end{aligned}$$

or

$$\frac{\psi_0-\psi_{\infty h}+\coth\psi_{\infty h}}{\psi_0-\psi_{\infty l}+\coth\psi_{\infty l}}=\frac{\sinh\psi_{\infty l}}{\sinh\psi_{\infty h}}=\frac{\Omega_l}{\Omega_h}. \qquad (4.1.3)$$

This is a general result, and it applies both to p–n and l–h junctions. It shows that $(\psi_{\infty h}-\psi_0)$, the potential rise in the h region, can exceed the value $\coth\psi_{\infty h}$ if Ω_h/Ω_l is negative, i.e. in a p–n junction. For an l–h junction, this potential rise is always less than $\coth\psi_{\infty h}$ and tends to this value as Ω_h/Ω_l tends to $+\infty$. Since, in a practical junction, $\coth\psi_{\infty h}$ will be negligibly different from unity and Ω_h/Ω_l will be large, it follows that in such a case the majority of the potential rise occurs in the l-region.

It is also of interest to calculate u_{max}, the maximum electric field in the junction. This occurs at $x=0$, and is found by combining eqns. (4.1.1) and (4.1.2) with $\psi=\psi_0$. For simplicity, we assume that Ω_h/Ω_l is large, so that $\psi_{\infty h}-\psi_0$ is approximately unity, and $\cosh(\psi_{\infty h}/\psi_0)$ is approximately $\exp(\psi_{\infty h}/\psi_0)$. Then u_{max} is given by

$$\begin{aligned}\tfrac{1}{2}u_{max}^2&=2(\cosh\psi_0-\cosh\psi_{\infty h})+\Omega_h(\psi_0-\psi_{\infty h})\\ &\sim(e^{-1}-1)\exp\psi_{\infty h}-\Omega_h\end{aligned}$$

or, since $-\Omega_h\sim\exp\psi_{\infty h}$, by

$$u_{max}=(-2\Omega_h/e)^{1/2}. \qquad (4.1.4)$$

As an extreme case, assume that Ω_h is about -10^5 (even larger values may be expected in alloy structures). Then u_{max} is about 270, corresponding to a field of about 7×10^4 V cm^{-1}. At fields of this magnitude the

electron 'temperature' will be considerably increased, and the assumption stated in §2 will no longer be valid. It is not entirely clear what difference this should make to the foregoing results, but it is obvious that eqns. (2.12) and (2.13) will no longer be valid. It is the total energy of an electron which enters into the Boltzmann equation, and in principle it should be sufficient to add to ψ a second term in $\nabla\psi$ taking account of the kinetic energy of the electron, which is now no longer negligible. We shall not pursue this point further, however, as it is shown below that it is unlikely that the precise nature of the potential distribution will have much effect on the properties of the l-h junction as a whole.

In order to calculate the width of the space-charge region, we need take account of only the l region, as shown above. In this region, we have, recalling eqns. (4.1.1) and (4.1.2)

$$x=\pm\tfrac{1}{2}\int_{\psi_0}^{\psi}\frac{d\psi'}{\{\cosh\psi'-(1+\psi'-\psi_{\infty l})\sinh\psi_{\infty l}\}^{1/2}}$$

$$=X(\psi,\psi_0,\psi_{\infty l}),\text{ say.} \qquad (4.1.5)$$

(X is to be positive, so it is the negative sign which must be taken.) It is not necessary to calculate X explicitly in order to obtain an estimate of the width. It may be supposed that space-charge effects are negligible beyond the point x_1 at which $\psi-\psi_{\infty l}$ is less than unity. Evidently, x_1 is given $x_1=X(\psi_{\infty l}+1,\psi_0,\psi_{\infty l})$. This quantity is plotted in fig. 2.

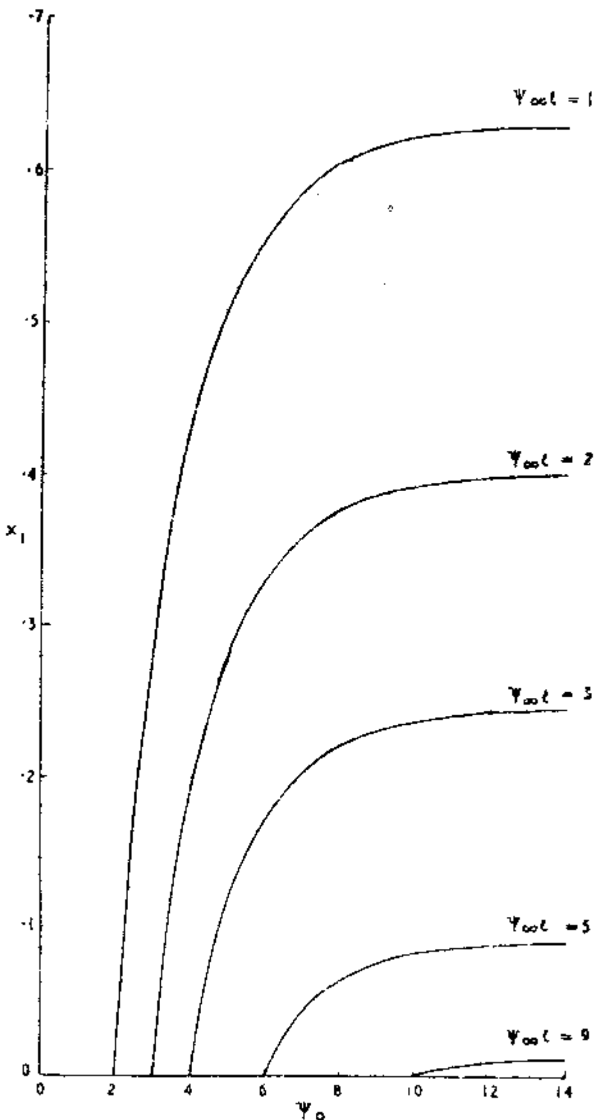
4.2. *Continuity Relations for Non-Equilibrium Conditions*

The exact calculation of ψ in the space-charge region when the junction is not in thermal equilibrium presents formidable difficulties. The differential equation to be solved is of the third order and non-linear, even in the simple case where one of the carrier currents is zero. Fortunately, it appears possible to treat the situation, with reasonable accuracy, as a quasi-equilibrium one. In this approach, it is assumed that only insignificant changes in the quasi Fermi levels, ϕ_p and ϕ_n, for holes and electrons respectively, occur on passing from one side of the space-charge region to the other (Shockley 1949). This will be true provided two conditions are fulfilled. The first of these is that the region must be so thin compared with the diffusion length for holes that recombination within it is negligible. The second is that the electrostatic potential difference across the region resulting from the flow of carriers must be negligible compared with that arising from difference of concentration.

In the first of these conditions, it is $\tau_l^{1/2}$, the diffusion length in the l-region, which is of primary importance, since the space-charge penetrates a negligible distance into the h-region. In transistor material $\tau_l^{1/2}$ may perhaps be as small as 10^2, in the worst case, and comparison with the values of x_1 in fig. 4 shows that the effective width of the space-charge region is likely to be much less than this. The second condition will be satisfied if the value E_{0l} of the electric field in the l-region, at points near to the space-charge region, is small compared with the value of

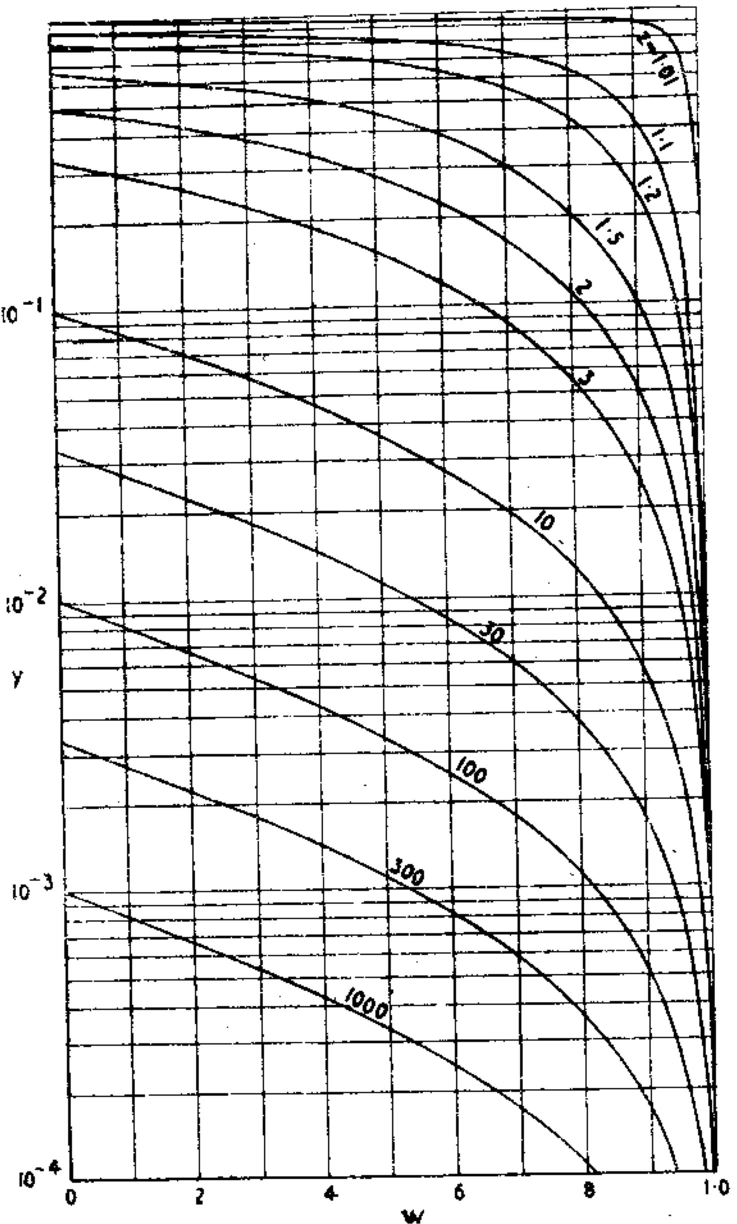
u_{max} given by eqn. (4.1.4). The carrier concentration in the space-charge region is always greater than that in the l-region: consequently any disturbance of u_{max} must be less than E_{0l}. The condition $E_{0l}\ll u_{max}$ will be satisfied in nearly all the cases which we shall consider, and hence the quasi-equilibrium approximation will be a useful one.

Fig 2



From the definition of a quasi Fermi level, it follows that $pn=\exp(\phi_p-\phi_n)$. If the subscripts 0h and 0l denote quantities measured just outside the space charge region, in the h- and l-regions

Fig 3



respectively, then in the present approximation we have the important continuity relation

$$p_{0h}n_{0h}=p_{0l}n_{0l} \qquad (4.2.1)$$

since $\phi_{p0h}=\phi_{p0l}$ and $\phi_{n0h}=\phi_{n0l}$.

Provided quasi-neutral conditions (§2) prevail in the h- and l-regions, the relations $p_{0h}-n_{0h}-\Omega_h\sim 0$ and $p_{0l}-n_{0l}-\Omega_l\sim 0$ permit eqn. (4.2.1) to be written, with p_∞=equilibrium value of p, in the form

$$(p_{0h}-p_{\infty h})(p_{0h}+p_{\infty h}-\Omega_h)=(p_{0l}-p_{\infty l})(p_{0l}+p_{\infty l}-\Omega_l). \qquad (4.2.2)$$

We have already assumed that p_{0h} and $p_{\infty h}$ are much less than Ω_h, so that the left-hand side of eqn. (4.2.2) is approximately $-\Omega_h(p_{0h}-p_{\infty h})$. If p_{0l} is comparable with $p_{\infty l}$, and the latter is negligible compared with $-\Omega_l$, then the equation is approximately

$$\Omega_h(p_{0h}-p_{\infty h})=\Omega_l(p_{0l}-p_{\infty l}) \text{ or } \Omega_h(N_{0h}-N_{\infty h})=\Omega_l(N_{0l}-N_{\infty l}). \qquad (4.2.3)$$

Alternatively, if p_{0l} is much greater than $p_{\infty l}$, we have, as an approximation,

$$-(b+1)\Omega_h(N_{0h}-N_{\infty h})=(N_{0l}-N_{\infty l})\{N_{0l}-N_{\infty l}-(b+1)\Omega_l\}. \qquad (4.2.4)$$

Finally, if $N_{0l}-N_{\infty l}$ becomes large compared even with $-\Omega_l$, the last relation tends to the quadratic form

$$-(b+1)\Omega_h(N_{0h}-N_{\infty h})=(N_{0l}-N_{\infty l})^2. \qquad (4.2.5)$$

It follows from the assumption of negligible recombination in the space-charge region that a hole current j_p must be the same on both sides of the region. The transport velocities v on the two sides of the junction are therefore related by equations of the same form as (4.2.3)–(4.2.5), with $N-N_\infty$ replaced by $(b+1)\,j_p/v$:

$$\Omega_h/v_{0h}=\Omega_l/v_{0l} \qquad (N_{0l}-N_{\infty l}\ll -\Omega_l) \qquad (4.2.6)$$

$$-\frac{\Omega_h}{v_{0h}}=\frac{1}{v_{0l}}\cdot\left(\frac{j_p}{v_{0l}}-\Omega_l\right) \qquad (N_{0l}-N_{\infty l}\gg N_{\infty l}+b\,\Omega_l) \qquad (4.2.7)$$

$$-\Omega_h/v_{0h}=j_p/v_{0l}^2 \qquad (N_{0l}-N_{\infty l}\gg -\Omega_l). \qquad (4.2.8)$$

Calculations of v_{0h} for various configurations of the l–h junction have been given in §3, and from these and the last three equations we can now calculate v_{0l} the transport velocity with which excess minority carriers are able to traverse the l–h junction. This velocity is, in general, quite small, whereas the junction offers no additional resistance to the passage of majority carriers. It can thus be regarded as a filter or semi-permeable membrane, which, ideally, is permeable only to majority carriers. A departure from ideality occurs because v_{0l} is only small, but not zero; we shall therefore call v_{0l} the *leakage velocity* of the junction. This behaviour may be contrasted to that of an ideal p–n junction which, as is well known, is permeable only to minority carriers.

§ 5. Density and Potential in the Lightly Doped Region ; Theory of the Generalized Conductance Transistor

An analysis of the conditions in the l region of the junction would start, ideally, from eqn. (2.11). The approximate form of this equation which was used in §3 is not applicable in the present case, however, since the restriction $N-N_\infty \ll N_\infty$ can no longer be made, and the non-linear nature of the equation is of primary importance. It turns out that the interesting solutions in the l region are those in which recombination effects are small, and as a first approximation we shall obtain solutions for the case of zero recombination. The necessary corrections for the case of small, but not negligible, recombination are discussed in §6. An exact solution for arbitrary recombination will not be attempted, as this can only be done numerically.

In a static one-dimensional system with zero recombination it is sufficient to use eqn. (2.10), since in this case j_p is a known function of position. The use of this equation under these conditions presents considerable advantages. Not only has it a relatively simple analytical solution, but it also permits the space coordinate to be eliminated from eqn. (2.9), so allowing a general solution for $\psi(N)$ to be obtained.

Let s be the space coordinate in a one-dimensional system. Then if J and j_p are solenoidal, they must both be described by the same function $S(s)$, which may for convenience be chosen in terms of the corresponding total currents I_p and i. Thus we have

$$J=J_s=IS, \qquad j_p=j_{ps}=i_p\,.\,s.$$

(Henceforth we write J for J_s and j_p for j_{ps}.) For brevity, put

$$\frac{2b}{b+1}=A, \qquad \frac{b(b-1)}{b+1}\Omega_1=B, \qquad I-(b+1)i_p=D, \qquad b\Omega_1 I=E.$$

Then eqn. (2.9) reads

$$\frac{d\psi}{ds}=\frac{b-1}{b+1}\frac{1}{N}\frac{dN}{dS}-\frac{IS}{N} \quad . \quad . \quad . \quad . \quad . \quad (5.1)$$

while eqn. (2.10) becomes

$$\frac{dN}{dS}=\frac{S(DN+E)}{AN+B} \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.2)$$

On dividing eqn. (5.1) by eqn. (5.2), we have

$$\frac{d\psi}{ds}\frac{ds}{dN}=\frac{d\psi}{dN}=\frac{b-1}{b+1}\frac{1}{N}-I\frac{AN+B}{N(DN+E)} \quad . \quad . \quad . \quad (5.3)$$

If A, B, D, E are constants, eqn. (5.3) has the solution

$$-\beta\psi+C_2=\ln\,(N-N_1) \quad . \quad . \quad . \quad . \quad . \quad (5.4)$$

where

$$\beta=\{1-(b+1)i_p/I\}/\{1+(b-1)i_p/I\}$$
$$N_1=-b\Omega_1/\{1-(b+1)i_p/I\}$$
$$C_2=\text{constant}.$$

This equation may be regarded as a modified form of the Boltzmann

equation. Indeed, when the minority carrier current is negligible compared with the total current, eqn. (5.4) does in fact reduce to the Boltzmann form of minority carriers. This may be seen by putting $(b+1)i_p/I \ll 1$ for n-type material, or $i_p/I \sim 1$ for p-type material. The resulting equations are then $-\psi+C=\ln p$, and $\psi+C=\ln n$, respectively.

In the case of the l region of an l–h junction, the constant of integration C_2 is determined by the boundary condition on the l side of the space-charge region. The leakage velocity v_{01} has already been calculated, so that, if ψ_{01} is the potential at this point, the condition is $\psi=\psi_{01}$, $S=S_0$, $N=N_{\infty 1}+(b+1)i_p(S_0/v_{01})$ when $s=s_0$. The potential is thus related to the carrier density by

$$-\beta(\psi-\psi_{01})=\ln\left\{\frac{N-N_1}{N_{\infty 1}-N_1+(b+1)i_p(S_0/v_{01})}\right\} \quad . \quad . \quad . \quad (5.5)$$

Turning now to the dependence of N on s, we see that this is obtained by integrating eqn. (5.2) and is

$$I\int S\,ds=N_1\beta^{-1}\ln\,(N-N_1)-\frac{2}{b+1}\frac{N_1}{\Omega_1}N. \quad . \quad . \quad (5.6)$$

If the l–h junction is placed at $s=s_0$, the integral in this equation may be made definite by using the boundary condition $s=s_0$, $S=S_0$ when $N=N_{\infty 1}+(b+1)i_p(S_0/v_{01})$. The equation thus obtained is

$$I\int_{s_0}^{s} S(s')\,ds'=N_1\beta^{-1}\ln\left\{\frac{N-N_1}{N_{\infty 1}-N_1+(b+1)i_p(S_0/v_{01})}\right\}$$
$$-\frac{2}{b+1}\frac{N_1}{\Omega_1}\{N-N_{\infty 1}-(b+1)i_p(S_0/v_{01})\}. \quad . \quad . \quad . \quad (5.7)$$

Finally, N may be eliminated from eqns. (5.5) and (5.7), yielding an equation between ψ and s:

$$\{I-(b+1)i_p\}\int_{s_0}^{s} S\,ds'=b\Omega_1(\psi-\psi_{01})$$
$$+\frac{2b}{b+1}\{N_{\infty 1}-N_1+(b+1)i_p(S_0/v_{01})\}$$
$$\times\{\exp\,[-\beta(\psi-\psi_{01})-1]\}. \qquad (5.8)$$

Equation (5.8) is the key equation of the present paper, and from it may be developed the theory of a certain general class of bi-polar transistors. This we shall term the class of *conductance transistors*; it is composed of those device in which a power gain is obtained from the modulation, by minority carriers, of the conductance of a quasi-neutral region of semi-conductor. (A typical conductance transistor is the filamentary transistor of Shockley *et al.* 1949.) The class is to be distinguished from a second, that of *rectifying junction* transistors, in which power gain is derived from the modulation by minority carriers of the reverse current of a rectifying junction (e.g. the p–n–p junction transistor). Such a junction is, by contrast, a region of space-charge.

The most important single parameter determining the properties of a conductance transistor is the leakage velocity v_{01} at the collector contact.

This may, as shown by the results of §3, be made to vary over a very wide range of values, depending on the nature of the contact. It is the consideration of an arbitary value of v_{01}, together with the use of equations which are valid when N is much greater than $N_{\infty 1}$, which distinguishes the general eqn. (5.8) from the theory of the elementary filamentary transistor (large values of N may occur when v_{01} is small). In the following subsections we shall apply the general equation to three examples, in each case calculating the current gain of the collector. The three examples have the following notation in common: I_0 is the value of I when $i_p=0$, and is just the product of the applied potential, and the conductance of the system due to electrons. s_e is the value of s at which the base connection to the system is made, and the potential ψ_e here is determined by external circuit conditions. The hole current enters the system at $s=s_e$. The internal current gain, α^* is equal to $\partial I/\partial i_p$, and is the current gain which would be exhibited by a transistor in which the given structure was used as collector and the effects of parasitic resistances and emitter inefficiency were negligible.

5.1. *Simple Filamentary Transistor*

As a first example of the application of eqn. (5.8), we consider the simplest possible case, namely, that of a filamentary transistor terminated in dumb-bell end connections at $s_0=x=0$ and at $s_e=x=x_e$. The hole current i_p enters at x_e and leaves at the origin. If the cross section of the filament is a, then S is simply a^{-1}, and

$$\int_{s_0}^{s_e} S\,ds' \quad \text{is} \quad x_e/a.$$

The dumb-bell ends of the filament ensure that the metallic connections are without effect on the filament proper, which behaves as though it were part of an infinite uniform filament. Under these conditions N is constant along the length of the filament, and equal† to N_1. Thus $N_{\infty 1}+(b+1)i_pS_0/v_{01}$, the value of N at $x=0$, is also equal to N_1, and the second term on the r.h.s. of eqn. (5.8) consequently vanishes.

If ψ_e is the potential at $x=x_e$, the total current is given by

$$\begin{aligned} I&=(\psi_e-\psi_{01})b\,\Omega_1 a/x_e+(b+1)i_p \\ &=I_0+(b+1)i_p. \qquad (5.1.1) \end{aligned}$$

The internal current gain α^* of the filament is thus equal to $(b+1)$. While this result is well known, it is not generally realized that its validity depends on the special value of v_{01} assumed here. Other values of v_{01} will lead to different values of α^*, since the second term on the r.h.s. of eqn. (5.8) will no longer vanish. The significance of N_1 is also brought out by the present treatment.

† This is seen by writing $i_p/I=p/(p+bn)=N+b\,\Omega_1/(b+1)N$. The physical significance of the constant N_1 is thus established.

5.2. *Filamentary Transistor with* l–h *Junction*

As a second example, a modification of the simple filament described above may be considered, in which the dumb-bell end at $x=0$ is replaced by an l–h junction having a low value of v_{01}. In particular, suppose $(b+1)i_p/av_{01}$ to be much greater than $N_{\infty 1}-N_1$, and $\beta(\psi_e-\psi_{01})$ much greater than unity. Equation (5.8) then becomes

$$I \sim I_0+\left(b+1-\frac{2b}{x_e v_{01}}\right)i_p. \qquad (5.2.1)$$

and large values of α^* can, in principle, be obtained (N.B. $v_{01}<0$). Such a structure has, however, the grave practical disadvantage that impracticably small values of x_e are necessary, for attainable values of v_{01}, if values of α^* significantly greater than $b+1$ are required. Not only are such short filaments difficult to manufacture, if planar flow is to be preserved, but there will also be serious power dissipation troubles due to I_0 if the requirement $\beta(\psi_e-\psi_{01})\gg 1$ is maintained.

5.3. *Radial System with* l–h *Junction*

The practical difficulties mentioned above may be largely overcome by the use of a system having spherical symmetry. For the purpose of analysis we consider a hemispherical junction of radius r_0. The space coordinate s is now the radius r, and the l–h junction occurs at $r=r_0$; outside this radius is the l-region, while within it is the h-region. The base contact is the surface $r=r_e$, at which radius the potential is ψ_e. If the solid angle occupied by the system is 2π, the function S is equal to $1/(2\pi r^2)$, and eqn. (5.8) reads

$$\frac{1}{2\pi}\cdot\left(\frac{1}{r_0}-\frac{1}{r_e}\right)\{I-(b+1)i_p\}=b\,\Omega_1(\psi_e-\psi_{01})$$

$$+\frac{2b}{b+1}\{N_{\infty 1}+N_1+(b+1)\,i_p/2\pi r_0^2\,v_{01}\}\{\exp[-\beta(\psi_e-\psi_{01})]-1\}. \quad (5.3.1)$$

In the special case where $\beta(\psi_e-\psi_{01})$ is much greater than unity, and $(b+1)(i_p/2\pi r_0^2)v_{01}$ is much greater than $N_{\infty 1}-N_1$, eqn. (5.3.1) becomes approximately

$$\begin{aligned} I&=2\pi r_0\frac{r_e}{r_e-r_0}\,b\,\Omega_1(\psi_e-\psi_{01}) \\ &\quad+\left\{(b+1)-\frac{r_e}{(r_e-r_0)}\cdot\frac{2b}{r_0 v_{01}}\right\}i_p \\ &=I_0+\left\{(b+1)-\frac{r_e}{(r_e-r_0)}\cdot\frac{2b}{r_0 v_{01}}\right\}i_p. \qquad (5.3.2) \end{aligned}$$

It will be seen that a large value of α^* can be obtained for any given v_{01} if r_0 alone is made small. This requires that the dimensions of the h-region to be kept small, but the difficulties in doing this are likely to be less than those involved in the corresponding filamentary case.

§ 6. Corrections for Recombination

The assumption of zero recombination in the l-region, forced upon us in §5 by the intractable nature of eqn. (2.11), is one which does not correspond very well to practical conditions. It is therefore of interest to make an approximate calculation of the effect of recombination on the results of §5, which may be used with confidence in the majority of practical cases.

The inclusion of the effect of recombination invalidates the method of integrating eqns. (5.2) and (5.3) because i_p, and hence D, are no longer constants. The solutions given (eqns. (5.4) and (5.6)), will, however, be approximately true provided the proportional variation D is small over the range of integration. This will be so, if either $(b+1)i_p/I$ is much less than unity, or the proportional variation of this quantity is small.

The first of these conditions is likely to be satisfied in any system in which α^* is large, whereas the second will apply when the leakage velocity of the junction is the dominant factor in determining the rate of hole loss from the l-region. In these cases, therefore, the function $N(s)$ and $\psi(s)$ can be taken to be those already calculated.

The hole current in the l-region obeys the equation $\nabla . j_p = G$ in the static state. In general, G is expected theoretically to be a function of $(N-N_{\infty l})$ of the form

$$G = -\frac{(N-N_{\infty l})}{(b+1)\tau_{0l}} \cdot \frac{1+c\,(N-N_{\infty l})}{1+d\,(N-N_{\infty l})},$$

and the lifetime appropriate to large values of $(N-N_{\infty l})$ may be different from τ_{0l}, that corresponding to small values (Shockley and Read 1952). In order to simplify the analysis, we shall neglect this refinement, and take G to be given by $G=-(N-N_{\infty l})/(b+1)\tau_l$ where τ_l is a constant 'average' lifetime for the l-region. On combining this relation with eqn. (2.3), and expressing $\nabla . j_p$ as $S(d/ds)(j_p/S)$, we obtain

$$\frac{di_p}{ds} = -\frac{N-N_{\infty l}}{(b+1)S\tau_l}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6.1)$$

If $i_p = i_{p0}, i_{pe}$, when $s=s_0, s_e$, integrating this equation gives

$$(b+1)\tau_l(i_{p_e}-i_{p0}) = -\int_{s_0}^{s_e} \frac{N-N_{\infty l}}{S}\,ds \quad . \quad . \quad . \quad (6.2)$$

but the integral cannot be evaluated directly because $N-N_{\infty l}$ cannot be expressed explicitly in terms of s. This difficulty can be overcome by changing the variable of integration from s to ψ by means of eqns. (5.1) and (5.2). When this is done, eqn. (6.2) becomes

$$(b+1)\tau_l(i_{p_e}-i_{p0}) = \int_{\psi_{0l}}^{\psi_e} \frac{\beta(N-N_{\infty l})(AN+B)}{DS^2}\,d\psi,$$

or, to the degree of approximation which has been assumed,

$$(b+1)\tau_l(i_{p_e}-i_{p0}) = \frac{B}{D}\int_{\psi_{0l}}^{\psi_e} \frac{(N-N_{\infty l})(AN+B)}{S^2}\,d\psi. \quad (6.3)$$

An expression for $N(\psi)$ can be obtained from eqn. (5.5) which, if $N'_{0l}=N_{\infty l}-N_1+(b+1)S_0 i_{p0}/v_{0l}$ is written for brevity, reads

$$N = N_1 + N'_{0l}\exp[-\beta(\psi-\psi_{0l})]. \quad . \quad . \quad . \quad . \quad (6.4)$$

Equation (5.8) gives only the implicit expression for $S(\psi)$,

$$\int_{s_0}^{s} S\,ds' = -\frac{N_1}{I}(\psi-\psi_{0l}) - \frac{2}{b+1}\frac{N_1}{\Omega_1}\frac{N'_{0l}}{I}\{\exp[-\beta(\psi-\psi_{0l})]-1\}. \quad (6.5)$$

The integral in eqn. (6.3) can be evaluated if eqn. (6.5) can be solved for S.

In dealing with eqn. (6.3), we shall not hesitate to make restrictive assumptions where these simplify the cumbersome forms representing the general case. The latter presents, however, no difficulty in principle.

6.1. *Planar Symmetry*

In a system with planar symmetry, eqn. (6.5) is redundant, since $S=a^{-1}$. The substitution of eqn. (6.4) into (6.3) gives

$$\frac{(b+1)D\tau_l}{\beta a^2}(i_{p_e}-i_{p0}) = (N_1-N_{\infty l})(AN_1+B)(\psi_e-\psi_{0l})$$
$$-\frac{1}{\beta}(2AN_1-AN_{\infty l}+B)N'_{0l}\{\exp[-\beta(\psi_e-\psi_{0l})]-1\}$$
$$-\frac{AN'^2_{0l}}{2\beta}\{\exp[-2\beta(\psi_e-\psi_{0l})]-1\}. \quad (6.1.1)$$

The implications of this equation are most clearly seen when $\beta(\psi_e-\psi_{0l})$ is much greater than unity, N'_{0l} much greater than $4N_1-2N_{\infty l}+2B/A$, and I much greater than I_0. The first two terms on the right-hand side can then be neglected. If i_{p0} is expressed in terms of D from eqn. (5.2.1) as

$$i_{p0} = -\frac{x_e v_{0l}}{2b}(D-I_0) \sim -\frac{x_e v_{0l}}{2b}D,$$

eqn. (6.1.1) leads to an approximate relation between i_{p_e} and I:

$$\tau_l\left(i_{p_e} + \frac{v_{0l}x_e}{2b}D\right) = \frac{x_e^2}{4b}D. \quad . \quad . \quad . \quad . \quad . \quad (6.1.2)$$

When $2\tau_l v_{0l}/x_e$ is much greater than unity, recombination can be neglected, and eqn. (6.1.2) is equivalent to eqn. (5.2.1). When this quantity is much less than 1, on the other hand, leakage through the junction is negligible, and most of the holes are lost by recombination. Under these conditions eqn. (6.1.2) reduces to

$$I \sim \left\{(b+1) + \frac{4b\tau_l}{x_e^2}\right\} i_{p_e}$$

or

$$\alpha^* = (b+1) + 4b\tau_l/x_e^2. \quad . \quad . \quad . \quad . \quad . \quad . \quad (6.1.3)$$

6.2. *Spherical Symmetry*

In a radial flow system of solid angle 2π, S is equal to $1/(2\pi r^2)$, and eqn. (6.5) gives

$$S=2\pi\left[-\frac{N_1}{I}(\psi-\psi_{01})-\frac{2}{b+1}\frac{N_1}{\Omega_1}\frac{N'_{01}}{I}\{\exp[-\beta(\psi-\psi_{01})]-1\}-\frac{1}{2\pi r_0}\right]^2.$$

From eqn. (5.3.1) we have

$$N'_{01}\{\exp[-\beta(\psi_e-\psi_{01})]-1\}=\frac{(b+1)}{4\pi b r_0}\frac{(r_e-r_0)}{r_e}\{I-I_0-(b+1)i_p\}$$

and hence

$$S=\frac{1}{2\pi r_0^2}\left[\frac{(r_e-r_0)}{r_e}\cdot\frac{I_0}{\{I-(b+1)i_p\}}\cdot\frac{(\psi-\psi_{01})}{(\psi_e-\psi_{01})}+\frac{(r_e-r_0)}{r_e}\frac{\{I-I_0-(b+1)i_p\}}{\{I-(b+1)i_p\}}\cdot\frac{\{\exp[-\beta(\psi-\psi_{01})]-1\}}{\{\exp[-\beta(\psi_e-\psi_{01})]-1\}}\right]^2.$$

If I_0 is restricted to be much less than $I-(b+1)i_p$, the first term on the r.h.s. of this equation, can be neglected. This gives an approximate expression for S,

$$S=\frac{1}{2\pi r_0^2}\cdot\left[\frac{(r_e-r_0)}{r_e}\cdot\frac{\{\exp[-\beta(\psi-\psi_{01})]-1\}}{\{\exp[-\beta(\psi_e-\psi_{01})]-1\}}-1\right]^2 \quad . \quad (6.2.1)$$

which, when substituted in eqn. (6.3), leads to an integral which can be evaluated analytically:

$$\frac{(b+1)D\tau_1(i_{p_e}-i_{p_0})}{4\pi^2\beta r_0^4}=\int_{v_{01}}^{v_e}\frac{\{N_1-N_{\infty 1}+N'_{01}\exp[-\beta(\psi-\psi_{01})]\}\times\{AN_1+B+AN'_{01}\exp[-\beta(\psi-\psi_{01})]\}d\psi}{\left[\frac{(r_e-r_0)}{r_e}\frac{\{\exp[-\beta(\psi-\psi_{01})]-1\}}{\{\exp[-\beta(\psi_e-\psi_{01})]-1\}}-1\right]^4}. \quad (6.2.2)$$

If, for brevity, we write

$$y=\exp[-\beta(\psi-\psi_{01})],\qquad g=\frac{N_1-N_{\infty 1}}{N'_{01}},$$

$$h=\frac{N_1+B/A}{N'_{01}},\qquad k=\frac{r_0-r_e\exp[-\beta(\psi_e-\psi_{01})]}{r_e-r_0}$$

and neglect I_0 by comparison with $I-(b+1)i_p$, eqn. (6.2.2) becomes approximately

$$-\frac{2b\tau_1(i_{p_e}-i_{p_0})}{\{\exp[-\beta(\psi_e-\psi_{01})]-1\}^2D}\cdot\frac{(r_e-r_0)^2}{r_0^2r_e^2}=\int_1^{\exp[-\beta(\psi_e-\psi_{01})]}\frac{(g+y)(h+y)}{y(k+y)^4}dy=H,\text{ say}. \quad . \quad . \quad . \quad (6.2.3)$$

The r.h.s. of eqn. (6.2.3) is

$$H=\left[\frac{1}{3(k+y)^3}\left(\frac{gh}{k}-g-h+k\right)+\frac{1}{2(k+y)^2}\left(\frac{gh}{k^2}-1\right)+\frac{1}{k+y}\cdot\frac{gh}{k^3}-\frac{gh}{k^4}\ln\left(\frac{k+y}{y}\right)\right]_1^{\exp-\beta(\psi_e-\psi_{01})}$$

The physical significance of this expression is best seen in the case where N'_{01} is very large, so that g and h can be neglected, and $\beta(\psi_e-\psi_{01})$ is, as usual, much greater than unity. It is then approximately equal to

$$\frac{(r_e-r_0)^2}{r_0^2}\left(\frac{1}{3}\frac{r_e^3-r_0^3}{r_e^3}-\frac{1}{2}\frac{r_e^2-r_0^2}{r_e^2}\right).$$

If i_{p_0} is expressed from eqn. (5.3.2) as $i_{p_0}\sim-(1-r_0/r_e)\cdot(r_0v_{01}/2b)\cdot D$, eqn. (6.2.2) is approximately

$$\tau_1\left\{i_{p_e}+\frac{(r_e-r_0)}{r_e}\cdot\frac{r_0v_{01}}{2b}\cdot D\right\}=\frac{re^2}{2b}\left\{\frac{1}{2}\frac{(r_e^2-r_0^2)}{r_e^2}-\frac{1}{3}\frac{(r_e^3-r_0^3)}{r_e^3}\right\}D. \quad . \quad . \quad (6.2.3)$$

This may be compared with eqn. (6.1.2), to which, indeed, it becomes equivalent when r_e-r_0 is much less than r_0. When r_e^2 is much greater than r_0^2, on the other hand, the r.h.s. of eqn. (6.2.3) becomes approximately $r_e^2/12b$. The condition for negligible leakage through the junction is now that $6r_0v_{01}\tau_1/r_e^2\cdot(r_e-r_0)/r_e$ should be much less then unity, and under these conditions α^* is given by

$$\alpha^*=(b+1)+\frac{12b\tau_1}{r_e^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (6.2.4)$$

(cf. eqn. (6.1.3)).

§ 7. Applications

We now turn to some additional consequences of practical interest which may be predicted from the results of previous sections. The experimental verification of some of these consequences has been discussed in other papers (Arthur *et al.* 1956 a, b).

7.1. *Measurement of Leakage Velocity*

Since the leakage velocity v_{01} of an l–h junction system plays such an important part in determining its properties, it will be important to have available a simple method of measuring this quantity. The method should preferably use simple apparatus likely to be already available in a transistor laboratory, and require the minimum of sample preparation. The following scheme meets these requirements quite well under conditions where recombination can be neglected.

A lamp, chopper disc, and lens system, such as are normally used for the measurement of minority carrier lifetime by the travelling light spot method, are arranged to project a small intensity-modulated spot of light on the surface of the l-region of a uniform filament containing the junction under investigation. A constant current I is passed through the specimen by means of a battery and series resistor, and the small fluctuating current $i_p\sim$ of holes excited by the light spot is swept towards the junction where it accumulates. If the spot is a distance x from the junction, the potential difference in the filament between the spot and a

point just to the l-side of the junction will contain an alternating component equal to

$$i_p\sim\frac{\partial}{\partial i_p}\{\psi(x)-\psi_{01}\}\equiv i_p\sim\psi'_{x01},\ \text{say}.$$

From eqn. (5.8). ψ'_{x01} is given by

$$\psi'_{x01}\left\{\frac{2b\beta}{(b+1)}N'_{01}\exp\left[-\beta(\psi-\psi_{01})\right]-b\Omega_1\right\}$$
$$=(b+1)\frac{x}{a}+2b\{\exp\left[-\beta(\psi-\psi_{01})\right]-1\}\left\{\frac{-N_1}{I-(b+1)i_p}+\frac{1}{av_{01}}\right\}$$
$$-\frac{4b^2}{b+1}\frac{N'_{01}I\psi}{\{I+(b-1)i_p\}^2}\exp\left[-\beta(\psi-\psi_{01})\right] \quad (7.1.1)$$

(assuming that v_{01} is independent of i_p).
When $\beta(\psi-\psi_{01})$ is much greater than unity, this takes the approximate form

$$\psi'_{x01}=-\frac{(b+1)}{ab\Omega_1}\cdot x+\frac{2}{\Omega_1}\left\{\frac{-N_1}{I-(b+1)i_p}+\frac{1}{av_{01}}\right\}. \quad . \quad . \quad (7.1.2)$$

In practice, it is not possible to measure ψ_{01}, and so ψ_{1x01}, directly, since any connection made near the junction will be in a region of large minority carrier density, and a floating potential will be developed whose magnitude will depend on the precise nature of the contact. It is better to measure $\psi_{\infty h}$, the potential in the h-region far away from the junction, and to make allowance for the difference between this and ψ_{01}. If the resistive potential drop across the h-region is small, and the electron density in the l-region near the junction is n_{01}, while that in the h-region far away from it is $n_{\infty h}$, $\psi_{01}-\psi_{\infty h}$ is approximately

$$\psi_{01}-\psi_{\infty h}=\ln\left(\frac{n_{01}}{n_{\infty h}}\right).$$

But n_{01} is equal to $-\Omega_1+p_{\infty 1}+(i_p/av_{01})$, while $n_{\infty h}$ is approximately $-\Omega_h$. Hence we have

$$\psi_{01}-\psi_{\infty h}=\ln\left\{\frac{-\Omega_1+p_{\infty 1}+(i_p/av_{01})}{-\Omega_h}\right\}.$$

and

$$\frac{\partial}{\partial i_p}(\psi_{01}-\psi_{\infty h}=\frac{1}{av_{01}\{-\Omega_1+p_{\infty 1}+(i_p/av_{01})\}}$$
$$\equiv\psi'_{01\infty h},\ \text{say} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.1.3)$$

(again assuming v_{01} independent of i_p). The alternating component $\psi\sim$ of the total potential difference $\psi(x)-\psi_{\infty h}$ is therefore

$$\psi\sim/i_p\sim=\frac{\partial}{\partial i_p}\{\psi(x)-\psi_{\infty h}\}=\psi'_{x01}+\psi'_{01\infty h}$$
$$=-\frac{(b+1)}{ab\Omega_1}\cdot x-\frac{2N_1}{\Omega_1}\cdot\frac{1}{\{I-(b+1)i_p\}}$$
$$+\frac{1}{av_{01}\Omega_1}\frac{\{\Omega_1-2(p_{\infty 1}+i_p/av_{01})\}}{\{\Omega_1-(p_{\infty 1}+i_p/av_{01})\}}. \quad (7.1.4)$$

Since the total current is constant, and the hole density fluctuation at x is small by virtue of the condition $\beta(\psi-\psi_{01})\gg 1$, the potential difference between x and the end connection to the l-section of the filament will be substantially independent of i_p. The alternating component of the potential difference between the end-connections to the filament will therefore be equal to $\psi\sim$, which can thus be measured by some suitable amplifier and meter. If $\psi\sim$ is plotted against x, eqn. (7.1.4) shows that a straight line should result for values of x such that $\beta(\psi-\psi_{01})\gg 1$. The ratio of the intercept $\psi\sim(0)$ of this line on the $\psi\sim$ axis to its slope $d\psi\sim/dx$ should be

$$\frac{\psi\sim(0)}{d\psi\sim/dx}=\frac{b}{(b+1)}\left[\frac{2aN_1}{I-(b+1)i_p}-\frac{1}{v_{01}}\cdot\frac{\{\Omega_1-2(p_{\infty 1}+i_p/av_{01})\}}{\{\Omega_1-(p_{\infty 1}+i_p/av_{01})\}}\right]. \quad (7.1.5)$$

If v_{01} is small compared with $\{I-(b+1)i_p\}/2aN_1$, and $(p_{\infty 1}+i_p/av_{01})$ small compared with Ω_1, eqn. (7.1.5) takes the simple form

$$\frac{\psi\sim(0)}{d\psi\sim/dx}\sim-\frac{b}{b+1}\cdot\frac{1}{v_{01}} \quad . \quad . \quad . \quad . \quad . \quad (7.1.6)$$

and hence v_{01} is easily measured.

7.2. *Transistor Properties*

In order to facilitate the discussion given in the next section, it will be convenient to collect together in a single section simplified forms of the equations given previously, together with certain additional results necessary for the estimation of transistor parameters. Attention will be confined mainly to spherically symmetrical systems of solid angle 2π, since these represent the most practically interesting type. They have been described by Granville (1956).

7.2.1. *Collector Properties*

Leakage velocity. The results given in §3 and 4 may be simplified somewhat when the h-region is made by the alloying technique. The thickness r_0-r_s is then usually small compared both with r_0 and $\tau_h^{1/2}$, so that eqn. (3.2.6) is applicable, and in addition the conductivity is so high that the term $I/(4\pi b\Omega_h r_0^2)$ can be neglected. In the alloying process, a small pellet of alloying material of volume U is placed on the semiconductor surface, heated, and allowed to come to equilibrium at some temperature above its melting point at which the volume solubility of the semiconductor in the molten alloying material is η_1. On cooling, dissolved material is precipitated on the inside of the approximately hemispherical cavity formed by the previous solution process until the remaining liquid freezes, at which temperature the solubility is η_0. The re-precipitated material, which is doped to an extent depending on the small solubility η_2 of the impurity component of the alloying material in the semiconductor, is then in the form of a hemispherical shell of volume $U(\eta_1-\eta_0)$. The alloying material may solidify in an approximately

spherical blob, and if so the radius of the junction is, if η_1 is small, approximately $\delta(3U/4\pi)^{1/3}$. Hence the thickness of the h-region is $r_0-r_s=\frac{2}{3}r_0=(\eta_1-\eta_0)r_0$, and the transport velocity is, from eqn. (3.2.6),

$$v_{0h}=\frac{-3}{2(\eta_1-\eta_0)r_0}. \qquad (7.2.1.1)$$

The quantity $(\eta_1-\eta_0)$ can usually be obtained, as a function of temperature, from the phase diagram of the system involved.

The leakage velocity v_{01} may be expressed, from eqn. (4.2.2), in a simple form by introducing the abbreviations

$$v_{010}=v_{0h}(\Omega_1-2p_{\infty 1})/\Omega_h \qquad (7.2.1.2)$$

and

$$i_{p_c}=-\frac{\Omega_h}{4S_0}\cdot\frac{v_{010}^2}{v_{0h}}. \qquad (7.2.1.3)$$

The leakage velocity is then

$$v_{01}=\tfrac{1}{2}v_{010}\{1+(1+i_{p_0}/i_{p_c})^{1/2}\}. \qquad (7.2.1.4)$$

For values of i_p much less than the critical hole current i_{pc}, v_{01} is equal to v_{010}, which may be termed the initial leakage velocity. For values of i_p much larger than i_{pc}, on the other hand, the leakage velocity is $\frac{1}{2}v_{010}(i_{p0}/i_{p_c})^{1/2}$.

It is evident from the foregoing equations that the maximum temperature of the alloying process will, by affecting both $(\eta_1-\eta_0)$ and η_2, affect both v_{0h} and Ω_h in an l–h junction. This will result in the maximum temperature having a marked effect on the leakage velocity, which is expected to decrease with increasing temperature. This behaviour is distinct from that of an alloyed p–n junction, in which the composition and thickness of the heavily doped region are irrelevant to its electrical properties, and the temperature is chosen from other considerations.

Collector resistance. The collector resistance, R_c, of conductance transistor is, formally, $1/R_c=\partial I/\partial(\psi_{\infty h}-\psi_c)$. In practice, however, $\psi_{\infty h}-\psi_{01}$ will vary negligibly if Ω_h is large, and $1/R_c$ is thus equal to $\partial I/\partial(\psi_{01}-\psi_e)$. Hence, for an ideal spherical system with $\beta(\psi_{01}-\psi_e)\gg 1$, R_c is from eqn. (5.3.1),

$$R_c=\frac{1}{2\pi b\Omega_1}\left(\frac{1}{r_0}-\frac{1}{r_c}\right). \qquad (7.2.1.5)$$

Provided that, as is usually the case in practice, the spreading resistance term $1/2\pi r_0 b\Omega_1$ is dominant, large departures of the equipotential surface ψ_e=const from sphericity will not affect the result.

The collector cut-off current I_0 is simply

$$I_0=(\psi_{01}-\psi_e)R_c. \qquad (7.2.1.6)$$

7.2.2. *Transfer Properties*

The interaction between emitter and collector terminals of a transistor is completely described by the current gain α and the base resistance R_b. The latter can in principle be reduced to a small value, but the resulting structure is impossible to manufacture. In practice, R_b is likely to depend so much on the exact type of construction employed that it is useless to calculate it theoretically. The current gain can for our purposes be set equal to α^*, the internal current gain at the collector, if a p–n junction emitter is used. This is so because the efficiency of such an emitter is negligibly different from unity by comparison with α^*. The emitter current I_e is then sensibly equal to i_{p_e}.

If it is assumed once again that $\beta(\psi_{01}-\psi_e)$ and r_e/r_0 are both much greater than unity, then in the case where recombination is unimportant $(6r_0v_{01}\tau_1/r_e^2\gg 1)$, eqns. (5.3.1) and (7.2.1.4) show that the collector current I is related to I_e by

$$I-(b+1)I_e=I_0+\frac{4\pi br_0}{(b+1)}\left\{N_{\infty 1}+\frac{b\Omega_1}{1-(b+1)I_e/I}\right\}-\frac{4bi_{p_c}}{r_0v_{010}}\{(I_e/i_{p_c}+1)^{1/2}-1\}. \qquad (7.2.2.1)$$

Hence the current gain α is given by

$$\{\alpha-(b+1)\}\left[1-\frac{4\pi b^2\Omega_1 r_0 I_e}{\{I-(b+1)I_e\}^2}\right]=\frac{4\pi b^2\Omega_1 r_0}{\{I-(b+1)I_e\}}-\frac{2b}{r_0v_{010}(I_e/i_{p_c}+1)^{1/2}}$$

The terms in $4\pi b^2\Omega_1 r_0/\{I-(b+1)I_e\}$ will usually be negligible in an l–h junction structure, and under these conditions the current gain may be approximated by

$$\alpha-(b+1)=\frac{-2b}{r_0v_{010}(I_e/i_{p_c}+1)^{1/2}}. \qquad (7.2.2.2)$$

The current gain has thus the following properties:

(*a*) The current gain α_0 for small signals $(I_e\ll i_{p_c})$ is given by

$$\alpha_0-(b+1)=-\frac{2b}{r_0v_{010}}, \qquad (7.2.2.3)$$

or, if eqn. (7.2.1.1) is applicable, by

$$\alpha_0-(b+1)=\frac{4b\Omega_h(\eta_1-\eta_0)}{3\Omega_1-2p_{\infty 1}}. \qquad (7.2.2.4)$$

Thus the current gain should be independent of the radius r_0. This result has been confirmed by Granville (1956).

(*b*) When I_e becomes comparable to the critical current i_{p_c}, the current gain begins to decrease, falling to the value α_1 when $I_e=i_{p_c}$. This is given by

$$\alpha_1-(b+1)=\frac{1}{\sqrt{2}}\{\alpha_0-(b+1)\}. \qquad (7.2.2.5)$$

(*c*) The critical current is

$$i_{p_c}=-\frac{\pi r_0^2}{2}\cdot\frac{(\Omega_1-2p_{\infty 1})^2}{\Omega_h}v_{0h} \qquad (7.2.2.6)$$

and for an alloy junction this is, by eqn. (7.2.1.1)

$$i_{p_c}=\frac{3r_0}{4}\cdot\frac{(\Omega_l-2p_{\infty l})^2}{\Omega_h(\eta_1-\eta_0)}.\qquad(7.2.2.7)$$

(*d*) For large emitter currents ($I_e \gg i_{p_c}$), the current gain tends to

$$\alpha-(b+1)=\left(\frac{i_{p_c}}{I_e}\right)^{1/2}\{\alpha_0-(b+1)\}.\qquad(7.2.2.8)$$

It will be seen from eqns. (7.2.2.4)–(7.2.2.8) that, for an alloy junction, the dimension r_0 of the alloy region enters only into the expressions for the behaviour at high emitter currents. Even so, the quantity $\{\alpha-(b+1)\}$ at a given large value of I_e is proportional only to $r_0^{1/2}$, and consequently the size of the alloy region is not likely to be a critical factor in determining α, by comparison with the firing temperature.

7.2.3. *Emitter Properties*

It might at first sight seem that the advantage of large current gain to be obtained from the use of an l–h junction in a conductance transistor would be offset by a concomitant increase in emitter resistance and decrease in emitter efficiency resulting from a greatly increase hole density in the l-region near the emitter. This could occur if the region of accumulation extended as far as the emitter. This cannot happen, however if $\beta(\psi_e-\psi_{0l})$ is much greater than unity. For, from eqn. (5.5), the density N_e near the emitter is, in the absence of recombination,

$$N_e-N_1=\{N_{\infty l}-N_1+(b+1)i_p\}\exp[-\beta(\psi_e-\psi_{0l})].\qquad(7.2.3.1)$$

In practice, the r.h.s. will be small even though v_{0l} be small. The emitter resistance may therefore be calculated on the assumption $N_e=N_1$. The calculation follows the method given by Shockley (1949). If ψ_E and $\Omega_E \gg 1$ are the potential and impurity density in the heavily-doped p-region of the emitter junction, then, if ϕ_p is continuous across the junction, we have

$$\frac{N_1+b\Omega_l}{b+1}=\Omega_E\exp(\psi_E-\psi_e).$$

Hence the emitter current is

$$I_e \sim i_p=I\left\{(b+1)-\frac{b\Omega_l}{\Omega_e}\exp(\psi_e-\psi_E)\right\}^{-1}.\qquad(7.2.3.2)$$

If I_e is much less than I, this is approximately

$$I_e=-\frac{b\Omega_l}{\Omega_E}I\exp(\psi_E-\psi_e)$$

and the emitter resistance $R_e=\partial(\psi_E-\psi_e)/\partial I_e$ is

$$R_e=1/I_e\qquad(7.2.3.3)$$

as for an ordinary p–n junction.

7.3. *Surface Recombination Velocity*

Consider a structure in which a thin, heavily doped 'skin' has been formed on the surface of a lightly doped specimen by, for example, the diffusion of suitable impurities into the surface. If this heavily-doped layer can be made thin compared with the diffusion length of minority carriers within it, eqn. (4.2.6) shows that the leakage velocity across the l–h junction may be smaller than the surface recombination velocity v_s at the outer surface of the h-region by a fraction of the order of Ω_l/Ω_h†. The specimen will then behave, in many ways, as though it were a simple l-type region with a surface recombination velocity equal to $(\Omega_l/\Omega_h)v_s$. Since Ω_l/Ω_h can easily be made very small, such an arrangement may provide a useful method of obtaining a very low value of apparent recombination velocity.

In order to calculate more generally the conditions under which an improvement may be made in this respect, we may make use of eqn. (3.1.5), which takes into account the thickness—x_s, minority carrier lifetime τ_h, and surface recombination velocity—v_s appropriate to the surface layer. On combining this equation with eqn. (4.2.6), we obtain, for small carrier density disturbances, the following expression for the apparent recombination velocity v_{0l}:

$$v_{0l}=\frac{\Omega_l}{\Omega_h}\tau_h^{-1/2}\frac{\exp(2x_s/\tau_h^{1/2})+(v_s\tau_h^{1/2}-1)/(v_s\tau_h^{1/2}+1)}{\exp(2x_s/\tau_h^{1/2})-(v_s\tau_h^{1/2}-1)/(v_s\tau_h^{1/2}+1)}.\qquad(7.3.1)$$

In terms of the variables $w\equiv -x_s/t_h^{1/2}$, $y\equiv -v_s\tau_h^{1/2}$, and $z\equiv v_{0l}\Omega_h/v_s\Omega_l$, this assumes the simpler form

$$\exp(-2w)=\frac{(y+1)}{(y-1)}\frac{(zy-1)}{(zy+1)}.\qquad(7.3.2)$$

The parameter z is simply the ratio of v_{0l} to the velocity which would be obtained with a surface layer of negligible thickness, and consequently it is a measure of the effect of the finite thickness and lifetime of the surface layer.

Figure 3 shows the general nature of the dependence of y on w, with z as a parameter. These curves may be used to determine the region of the yw-plane in which z is less than a given number, and consequently, to design a surface layer to have a given performance.

§ 8. Discussion

The theory given in this paper indicates that the phenomenon of carrier accumulation is likely to prove a useful tool in semiconductor physics, both as an experimental technique, permitting the observation of new effects such as those described in §7, and as a concept, giving a better insight into the mechanism of certain known effects. Progress in the former direction is bound to depend on the evolution of methods of constructing l–h junctions having suitable properties. This may prove slightly more difficult than the manufacture of equivalent p–n junctions.

† This possibility has been suggested independently by Moore and Webster (1955).

The properties of an l–h junction are sensitive to the nature of the material on both sides, whereas those of a p–n junction can be made to depend only on the material on one side by making the other side sufficiently impure. As examples of known phenomena in which accumulation may play a hitherto unsuspected part, we may cite the flow of forward current in the so-called 'p–i–n diode' and the large current gain of point-contact transistors which have been subjected to certain electro-thermal forming procedures.

The p–i–n diode is a structure consisting of an impure p-type region, an intermediate layer of relatively pure material of either type, and an impure n-type region. When connection is made to the p- and n-type impure regions, the device functions as a rectifier having a superior combination of properties to those of any conventional p–n junction. It is clear that, during the flow of forward current, minority carriers must be injected into the intermediate layer at the p–n junction formed between it and one of the impure regions. The junction with the remaining impure region is, however, of the l–h type, and the injected carriers must accordingly accumulate there to an extent depending on its properties. This accumulation can be important in determining the forward current of the diode. The complete analysis requires more detailed consideration of the properties of a p–n junction than can be given in the present paper.

Any theory purporting to explain the mechanism of the current multiplication at a formed point-contact collector electrode must take account of the following salient experimental results:

(1) Enhanced current gain after forming is observed only if the collector wire contains an impurity of the same (p or n) type as those in the bulk semiconductor.

(2) At room temperature, the current gain after appropriate forming is usually large for small values of emitter current, but always falls to a value of the order 1–3 at large currents.

(3) Those transistors which do not show effect (2) can often be made to do so by cooling to a temperature of the order of 0°c.

(4) After forming, the collector resistance is considerably decreased. These results will be interpreted in terms of the theory developed in this paper, but it is not considered worthwhile to attempt to find more than a qualitative agreement between theory and experiment. The characteristics of point contact structures are so variable that it is possible to choose a set of curves to fit almost any theory.

It is known that the action of forming produces intense local heating and results in the transfer of matter from the collector wire to the underlying semiconductor. Result (1) above indicates that the significant effect of this process is the introduction into the semiconductor of an impurity of the same type as that already present, resulting in the formation of a small, heavily-doped region surrounding the collector contact. The simplest hypothesis about the outer boundary of this region is that it forms a simple l–h junction with the bulk semiconductor† : the result is a radial structure, consisting of an approximately hemispherical metal surface, h-region in the form of a thin hemispherical shell, and surrounding l-region, so that all the results of the preceding sections can be applied. It follows that a high current gain is to be expected if the h-region has suitable properties, that this current gain should fall off with increasing emitter current according to eqn. (7.2.2.2) (result 2), and that the resistance of the contact should decrease with the decreased spreading resistance resulting from the increased conductivity in the neighbourhood of the contact (result 4). The third result may at first sight seem difficult to explain, but it should be remembered that in this case the radius of the junction is very small and that the critical current i_{p_c} may also be very small. It is possible, therefore, that if the lifetime in the l-region is rather short, the current of thermally excited minority carriers at room temperature may already exceed i_{p_c}, so that, even in the absence of emitter current, the current gain is drastically reduced below the value α_0.

Having verified that the proposed mechanism offers a convincing qualitative explanation, we need show only that the orders of magnitude involved are reasonable. In a particular fairly typical point contact n-type germanium transistor, α is equal to 15 at $I_e = -10^{-7}$ A, 11 at -2×10^{-6} A, and 2 at -10^{-3} A (data communicated privately by Mr. E. H. Cooke-Yarborough). The discrepancy between the last value and the theoretical value of $b+1=3{\cdot}0$ may reasonably be ascribed to recombination in the l-region and emitter inefficiency. On this assumption i_{p_c} is equal to $-1{\cdot}3\times10^{-6}$ A and α_0 (assuming negligible thermally excited current) and $\alpha_0{}^*$ to 23. From eqn. (7.2.2.3) it follows that $r_0 v_{010} = -0{\cdot}2$. If eqn. (3.2.6) is assumed to be applicable, this and eqn. (7.2.1.3) give a second relation

$$r_0{}^2 v_{010}{}^2 = \frac{2}{\pi}\,\frac{i_{p_c}}{\Omega_h(r_0 - r_s)},$$

whence $\Omega_h(r_0-r_s)$ is approximately -10^3. If (r_0-r_s) is of the order of one micron, Ω_h must correspond to an impurity density of the order of 10^{16} cm^{-3}. If, on the other hand, recombination in the h-region should determine v_{0h}, the same value of Ω_h would give a value for τ_h corresponding to a lifetime of about 10^{-10} sec. These figures do not seem implausible.

Acknowledgments

This paper is based upon a suggestion by Dr. A. F. Gibson that accumulation might lead to current gain. To him, and to the author's other colleagues at R.R.E. who have given help and encouragement, must

† If more than one impurity is concerned, the outer boundary may be a so-called 'p–n hook' (Shockley 1950), but precise assumptions about the relative diffusion constants are then necessary.

belong much of the credit for this paper but none of the responsibility. Thanks are also due to Mr. G. G. E. Low for permitting use to be made of his paper before publication, and to Mr. E. H. Cooke-Yarborough for data on point-contact transistors.

Acknowledgment is made to the Controller, H.M. Stationery Office, for permission to publish this paper.

References

Arthur, J. B., Gibson, A. F., and Gunn, J. B., 1956 a, *Proc. Phys. Soc. Lond.* B, **69**, 697; 1956 b, *Ibid.*, **69**,, 705.

Granville, J. W., 1956, *J. Electron.*, **1**, 565.

Gunn, J. B., 1954, *Proc. Phys. Soc. Lond.* B, **67**, 575.

Low, G. G. E., 1955, *Proc. Phys. Soc. Lond.* B, **68**, 310.

Moore, A. R., and Webster, W. M., 1955, *Proc. Inst. Radio Engrs., N.Y.*, **43**, 427.

Shockley, W., 1949, *Bell Syst tech. J.*, **28**, 435; 1950, *Phys. Rev.*, **78**, 294; 1951, *Bell Syst. tech. J.*, **30**, 990.

Shockley, W., Pearson, G. L., and Haynes, J. R., 1949, *Bell Syst. tech. J.*, **28**, 364.

Shockley, W., and Read, W. T., 1952, *Phys. Rev.*, **87**, 835, 1952.

van Roosebroeck, W., 1950, *Bell Syst. tech. J.*, **29**, 560; 1953, *Phys. Rev.*, **91**, 282.