P-N JUNCTIONS
IN
INTERMETALLIC SEMICONDUCTORS

Being the text of a thesis
submitted to the
University of London in the Faculty of Engineering
for the degree of
Doctor of Philosophy
by

Herbert Douglas Barber B.Sc.(Eng), M.Sc.
Electrical Engineering Dept.
Imperial College
University of London

1965

Abstract

A technique has been developed to prepare shallow penetration alloyed p-n junctions utilizing atomic or ionic hydrogen to achieve low temperature wetting. The electrical characteristics of diodes prepared by this technique on the intermetallic compound, indium antimonide, have been measured as a function of doping density, temperature, crystallographic orientation, dislocation density, surface treatment and metallurgical processing. The diode characteristics have been interpretted in terms of the properties of the base crystals and the junction regrowth.

An analytic expression for the current-voltage relationship of a diode in conditions of medium level injection has been developed. When combined with existing expressions for conditions of high and low level injection, the theory accurately describes the characteristics of InSb p^+n diodes.

Evidence has been obtained from measurements of space charge generation currents for a trapping level in InSb at 0.133 ev having a temperature dependence of -1.55×10^{-4} ev/ ^{o}K.

Metallurgical investigations have been shown that facets form on {111} and {100} surfaces in InSb and on {111} surfaces only in germanium. It has also been observed that A{111} surfaces dissolve more rapidly during alloying than do B{111} surfaces. Models for {111} and {100} surfaces are presented and used to explain these observations as well as metallurgical observations made by other workers.

Acknowledgements

I am indebted to Dr.J.C.Anderson and to Professor J.Lamb for providing the facilities required for this work and for supporting it throughout. The support and interest of Dr.E.L.Heasell is also gratefully acknowledge.

I would like to thank the members of the Materials Laboratory for their co-operation during the course of this work. In particular I would like to thank Dr.S.C.Choo and Mr.J.Hollis for valuable discussions on recombination in InSb, Dr.P.M.Gundry and Mr.P.Norgate for helpful discussions on the metallurgical and chemical aspects of the work, and Mr.J.Conradi and Mr.A.C.Papadakis who did a large part of the work in determining the In-InSb and Sn-InSb liquidus curves. Thanks are also due to members of the Transistor Laboratory for their help in the design of the low noise preamplifier and the broad band capacitor, and to the technical staff of the department for their assistance in constructing the many pieces of apparatus used in the work. It is also a pleasure to acknowledge the help of Mr.A.Willoughby who supplied the strained sample of InSb and information on unpublished results of his work on dislocations.

I am indebted to Dr.J.B.Mullin of the Royal Radar Establishment for providing most of the single crystals of InSb and to Dr.K.F.Hulme, Dr.D.T.Hurle and Dr.J.B.Mullin for their interest and stimulating discussions. I am also indebted to Dr.G.K.Teal of Texas Instruments who kindly provided me with a copy of an

unpublished report on some aspects of the investigations of InSb devices at the T.I.Laboratory in Dallas Texas. The development of the H100 etch used in the present work was a result of their investigation.

This work would not have been possible without the encouragement and understanding of my wife and the financial support of the Board of Trade in the form of an Athlone Fellowship and the National Research Council of Canada in the form of a NATO Special Scholarship.

INDEX

CHAPTER 1

Introduction

1.1 Purpose of the Study.

The metallurgy, electrical properties and methods of fabrication of p-n junctions in elemental semiconductors have been the subject of vast amounts of research since the discovery of transistor action[1] in 1948. Extremely rapid advances in semiconductor physics and devices have arisen from the reciprocal stimulation of pure research and technology. As a result a wealth of information is available which has made it possible to design devices to exacting standards in these materials. Without these developments many of the spectacular advances in other fields of research and endeavour would have been very seriously hampered, if not impossible.

Every material possesses different properties. Since design is the process of combining and compromising the properties of the components to approximate the desired properties of the device, it is not necessary to argue the value of a choice of materials for any application. It is for this reason that increasing resources are being channelled into the search for new materials and for an understanding of their properties. New families of semiconductors have been and are being discovered as a result. One of these families, first recognized as semiconductors by Welker[2] in 1952, is the family of IIIb-Vb (hereafter called III-V) compounds. Since

1952 these compounds have received increasing attention and although many of the properties of these materials pointed to unique and useful applications the technology and understanding, first of material preparation and then of device fabrication has caused much slower development than that achieved in elemental semiconductors.

The preparation of one of these compounds, indium antimonide, has been extensively studied[3] and is not hampered by the severe difficulties, such as the high vapor pressure of one component, encountered in other III-V compounds. Indium antimonide is therefore available in single crystals of comparable perfection and purity to the elemental semiconductors, silicon and germanium. It is largely for this reason that indium antimonide was chosen for this work. The purpose of this work was to begin a study of p-n junctions which would yield fundamental and technological information for device design in indium antimonide, similar to that already available for silicon and germanium. This is, of course, a vast project and hence the following limited objectives were envisaged:

(a) The necessary techniques for p-n junction fabrication in indium antimonide were to be developed and understood.

(b) Diodes formed by these techniques were to be characterized as far as possible in terms of the bulk properties of the material.

(c) The effects of temperature, doping density, dislocation density, crystallographic orientation and surface treatment on both the metallurgical and electrical properties of the junctions were to be determined.

(d) Diodes were to be used to study carrier recombination in an effort to substantiate the results of other experiments.[4][5]

(e) The possibility of transistor fabrication in indium antimonide was to be assessed.

The following chapters of this thesis describe how these objectives have been met and what information has been acquired as a result. In the remaining sections of this chapter the properties of indium antimonide and their implications for p-n junctions will be discussed, diffusion in III-V compounds and its applications to junction formation will be surveyed, and finally the work on p-n junctions in indium antimonide and other III-V compounds which preceded and paralleled this work will be reviewed.

1.2 Properties of Indium Antimonide.

1.2.1. Physical and Metallurgical Properties.

Indium antimonide is generally prepared in single crystal form by the method of melt growth. The techniques used for this purpose[3] are similar to those used for germanium. Measurements to establish the In-Sb phase diagram[6] have shown that InSb is highly stoichiometric. This has since been confirmed by electrical measurements and by the large negative entropy of formation found by Schottky and Bever[7]. At the melting point of InSb, 525°C [6], the vapor pressure of either

component is low and as a result the loss of one component at the melting point can be easily prevented and no elaborate apparatus is required for purification or crystal growth. The technical problems involved in the growth of single crystal InSb are, therefore, not large and provided precautions are taken properly to purify the starting material[8] and to use suitably oriented seeds[3], production of high purity single crystals is relatively easy. At present InSb is the only III-V compound which can be grown with purity and perfection comparable to silicon or germanium. It is, therefore, the obvious starting point for an investigation of p-n junctions in III-V compounds.

Although the low melting point and high stoichiometry of InSb are conducive to good crystal growth, they indicate possible difficulties in junction formation. Substitutional diffusion via vacancies could be a slow process in such a material. Diffusion to form p-n junctions will be discussed in more detail in section 1.3. Alloying to form junctions must be done at temperatures much lower than the melting point of InSb if large junction penetrations are to be avoided. This necessitates the use of a low melting point alloy and a procedure for achieving good wetting at relatively low temperatures.

Measurements of the thermal expansion of InSb[9][10] have shown it to vary from 1.5×10^{-6} per ^{o}C at $80^{o}K$ to 5.04×10^{-6} per ^{o}C at $300^{o}K$. The high temperature value lies between that of germanium and silicon[11]. Since the expansion

coefficient is almost identical to that of Ge from 80°K - 300°K, mounting devices on headers designed for germanium should avoid difficulties arising from thermal stress particularly if low yield strength solders are used.

In general the heat dissipation of a p-n junction is determined by the thermal conductivity of the semiconductor. The thermal conductivity of InSb has been measured[12][13][14] and shown to vary from 1.0 watt/cm°K at 77°K to 0.15 watt/cm°K at 300°K. The 77°K value is not substantially different from that for Si or Ge[15] at 300°K.

Indium antimonide is brittle and relatively hard[16] and it is particularly subject to a form of damage called "dislocation cracks"[17][18]. Damage has been shown to affect the electrical properties of InSb[19] and therefore cutting and handling must be done with great care.

As early as 1929 Goldschmidt[20] reported the crystal structure of InSb to be that of zinc-blende. More recently X-ray measurements by Warekois et al[21] have unequivocally established this to be true. The zinc-blende structure is polar in $\langle 111 \rangle$ directions as shown in Figure 1 and the consequences of this polarity will be discussed in section 1.2.2.

A further effect of the crystal structure is the tendency for the densely packed $\{111\}$ planes to form facets during crystal growth or dissolution. This tendency in the diamond structures of Si and Ge has been exploited to form flat alloyed junctions[22]. However, it has recently been observed in

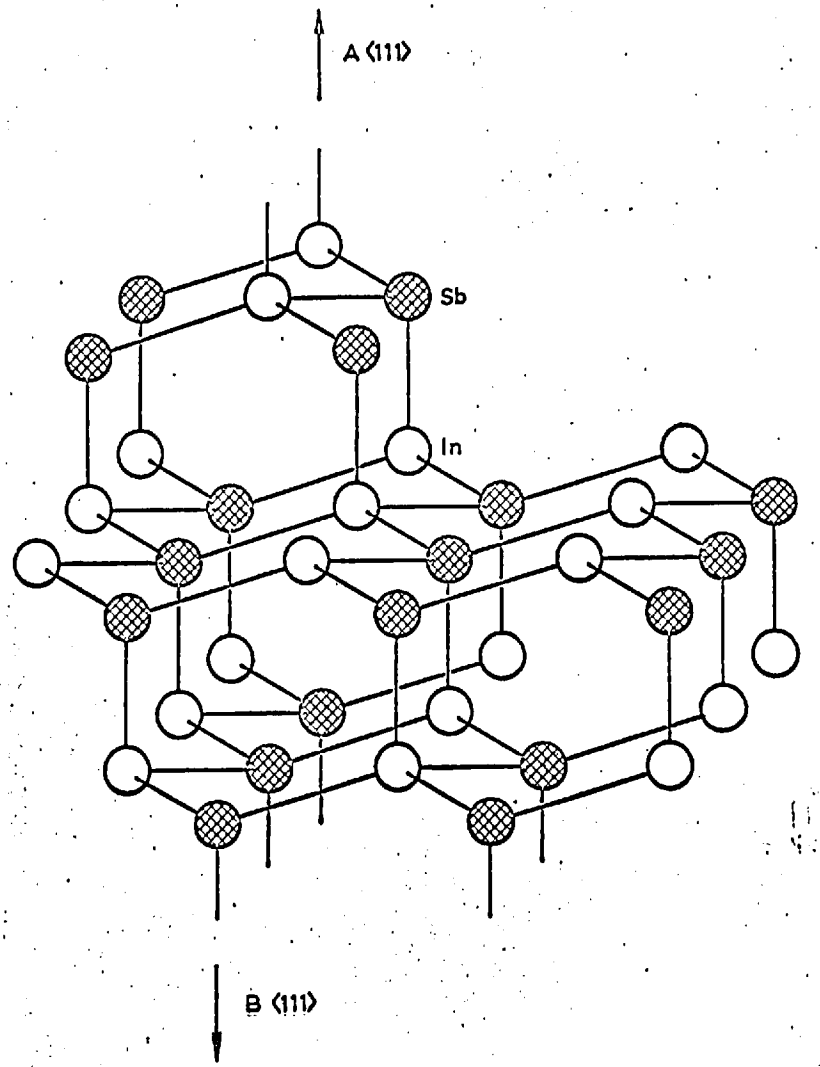


Figure 1. The crystal structure of InSb showing the convention used for labelling polar directions throughout this thesis. Since it will not be necessary to distinguish between members of a family of surfaces (i.e. $[111] = [\bar{1}11] = [1\bar{1}1] = [11\bar{1}]$) the $\{111\}$ surface composed of triply bonded indium atoms (i.e. uppermost surface) will be designated A$\{111\}$ and that composed of triply bonded antimony atoms will be designated B$\{111\}$. When referring to other III-V compounds the A and B will apply in the same way to the group III and group V atoms respectively.

InSb[23] that the segregation coefficient for some impurities is very much larger for growth occuring on a facet. This anomalous segregation could have a marked effect on the concentration and uniformity of impurities in the regrowth region of alloyed p-n junctions. The segregation coefficients and facet ratios for a number of impurities in InSb have been summarized by Hulme and Mullin[3]. In alloyed p-n junctions it is desirable to choose impurities with small facet ratios and a distribution coefficient less than one[3].

1.2.2. Polar Properties.

In this section the observed chemical, physical and metallurgical effects of the polar nature of the zinc-blende structure will be reviewed briefly. The main emphasis will be on effects observed in InSb but observations of similar effects in other III-V compounds will also be mentioned. The section will demonstrate the large macroscopic effects which can arise from the polarity of a crystal and will indicate the possible ways in which this polarity could influence the preparation and properties of p-n junctions. The conventions used throughout this thesis to designate each of the polar surfaces and directions are indicated in Figure 1.

(a) Chemical Polar Effects.

Experiments by Dewald[24] and others[25][26] have shown that during anodic oxidation under low field conditions the B surface oxidized more rapidly than the A surface. Gatos and Lavine[26][27][28] have found that the rate of attack by

oxidizing chemical etches was greater on B surfaces than on A for temperatures below 25°C. In addition it has been found (28)(29)(30) that unless inhibiters were used, dislocation etch pits developed only on A surfaces. (This effect is now generally accepted as basis for identification of surface polarity). The chemical effects may be related by noting that electrode potential measurements[27][31] indicate that the A surface is more noble or less chemically reactive than the B surface.

(b) Physical Polar Effects.

Measurements by both X-ray[32] and chemical etching[33] techniques have shown with good agreement that B surfaces suffer considerably greater damage during abrasive treatment than do A surfaces. More recently Pugh and Samuels[34] have carried out experiments using a low angle sectioning technique which are reported to show no difference in the maximum depth of damage on either surface. These results appear contradictory but in comparing them it must be noted that the density and distribution of damage with depth was measured in the early experiments while in the latter only the maximum depth of damage was measured. Also a sectioning technique to observe damage in InSb must be treated with some caution[18]. Warekois et al[32] have shown from X-ray analysis after chemical etching that A surfaces appear to be considerably more disordered than B surfaces. This has been taken to indicate a higher degree of strain in the bonding on A surfaces. Experiments by Hanneman, Finn and Gatos[35][36] have shown a difference in strain present between

A and B surfaces. It was found that specimens ground to thicknesses of about 10 microns bent spontaneously leaving the A surface convex while germanium control specimens remained flat. Haneman[37] has recently reported results of similar experiments which show that this bending effect cannot be attributed directly to a difference in the nature of the surface bonds.

Haneman et al.[38] have reported studies on the decomposition temperatures of InSb and GaAs on A{111}, B{111}, {100} and {110} surfaces. The decomposition temperatures for A{111} surfaces were considerably higher than those for B{111} surfaces. Those for {100} and {110} were nearly identical and lay between B{111} and A{111} but closer to B{111}. This has been taken to indicate stronger bonding of A atoms in the A{111} surfaces. Vacuum evaporation of GaAs[39] has shown that the dissociation rate is higher on B surfaces than on A for temperatures below 770°C. Above 800°C the rates become equal indicating molecular evaporation. Eisen[40] has found that the threshold displacement energy under electron bombardment is 2-3 ev less for indium atoms displaced in the B⟨111⟩ direction than for antimony atoms displaced in the A⟨111⟩ direction. Slow electron diffraction studies[41] on A and B surfaces of InSb have revealed no substantial differences although the experiments indicated that there could be some variation in spacing between first and second atomic layers of the two surfaces.

The polar nature of the zinc-blende structure permits two types of 60°-dislocation[42]. The α dislocations have a row

of A atoms with dangling bonds and the β dislocations have a similar row of B atoms. Experiments have shown[43][44] that these dislocations have different properties. The α dislocations have a higher mobility and lower activation energy than β dislocations.

(c) Metallurgical Polar Effects.

Although there has been conflicting evidence on the metallurgical effects of crystal polarity, experimental observations indicate an influence on crystal growth which cannot be explained on the basis of environmental conditions alone. Moody et al[45] found that of twenty GaAs crystals grown in the A direction none was single crystal. All had dislocation densities two to four times higher than crystals grown in the B direction and the dislocation density increased with the length of crystal grown. Good crystals of uniformly low dislocation density could be obtained by growth on the B surface. Steinemann and Zimmerli[46] also found that crystals grown on A surfaces tended to twin easily on oblique B surfaces and that the probability of twinning was sensitive to the stoichiometry of the melt. Similar growth peculiarities have been observed[3][47][48] in InSb. However, Richards[49] growing GaAs in horizontal boats obtained good single crystals in either growth direction when the shape of the freezing interface was correctly controlled. Observations of dendritic growth in InSb[50] gave no evidence for differences in the growth properties of A or B surfaces but similar observations in GaAs[51] showed evidence of differences in step

nucleation and propagation on the two surfaces. Both Ellis[52] and Steinemann[46] have found that under the same conditions A surfaces grow more quickly than B. Recent observations of epitaxial growth on GaAs[53][54] have confirmed this. Mullin[23] has shown that during crystal growth facets tend to form more easily on B surfaces than on A surfaces. He has also shown that for some impurities the segregation coefficients are increased on the facets with the values possibly somewhat higher on B facets. Williams[54] has observed large differences in the impurity concentrations of expitaxial layers grown on A or B surfaces. No polarity dependence has been observed in diffusion experiments.[31]

(d) Possible Effects on p-n junctions.

The metallurgical polarity phenomena could easily affect alloyed p-n junctions and it might be expected that junctions formed on A surfaces would have poor regrowth regions. It is also possible that the anomalous segregation coefficients on facets could lead to differing impurity concentrations in A or B regrowth regions. During the course of this work, Minamoto[55][56] reported observing both of these effects in alloyed junctions. Because strain and damage affect the wetting and alloying characteristics of semiconductors[57] a difference in these processes might be expected on A and B surfaces. Also it is well known that dislocations affect the electrical properties of p-n junctions[58] and therefore in compound semiconductors it is possible that a difference in these effects might be observed depending on the type of dislocation which was predominant.

The chemical treatment of p-n junctions to eliminate surface effects is extremely important and it would be most fortuitous if the differences in the chemical properties of A and B surfaces did not give rise to different surface effects for the same chemical treatment. It is therefore apparent that the polar nature of III-V compounds could have important influences on the preparation and properties of p-n junctions.

1.2.3. Electronic and Optical Properties.

It is the properties of electrons in InSb coupled with its ease of preparation which have made it one of the most studied semiconductors. InSb appears to be the only well established semiconductor with valence band maximum and conduction band minimum both at k = 0 and both bands spherically symmetric as a good approximation.[18] The energy gap obtained by extrapolating the results of high temperature optical absorption[58] and intrinsic Hall[60] measurements to $0^{\circ}K$ is 0.25 ev. The energy gap at $300^{\circ}K$ is 0.165 ev. This small energy gap results in an optical absorption edge at a wave length of about 7 microns. InSb is therefore useful both as an infra-red detector and filter. Another effect of this small energy gap is the extremely high intrinsic carrier concentration ($n_i = 2 \times 10^{16} cm^{-3}$) at $300^{\circ}K$. Except for p-n junctions where tunneling predominates, that is with carrier concentrations in both regions greater than, say, $10^{17} cm^{-3}$, electrical characteristics could not be expected to show diode action at room temperature. Junctions in InSb must therefore

be cooled to liquid oxygen (90.2°K) or liquid nitrogen (77.4°K) temperatures to observe efficient diode action. Refrigerated shielding must also be provided since such devices would be extremely sensitive to any radiation beyond the absorption edge.

The effective mass of electrons in InSb ($m_e = 0.0145\ m_o$) is now well established.[61][62][63] However, since the conduction band is not parabolic[63] this value is an average which adequately represents the band for carrier concentrations up to about $10^{17} cm^{-3}$. Because of this low effective mass the Landau splitting in the presence of a magnetic field is greatly increased and a variety of magneto-optic effects are well defined in InSb[64]. The low effective mass also results in a very high electron mobility, typically $7.8 \times 10^4 cm^2/v$ sec at 300°K. This is some twenty times higher than the electron mobility in germanium. In pure InSb cooled to 77°K it is possible for this value to increase by a factor of ten. Such a high mobility gives rise to high conductivities and to large galvanomagnetic effects such as magneto-resistance and Hall effect. The increased conductivity of n-type InSb suggests that the series resistance effects in p-n junction devices could be very considerably less than comparable Ge or Si devices. The combined effect of increased conductivity and decreased transit times should considerably improve the response times of transistors and junction photodetectors. In fact, Henneke[65] has calculated that an InSb transistor should have a cut off frequency five to twenty-five times higher than that

of a Ge transistor of similar geometry. The low effective mass, however, also implies a large curvature in the conduction band at k = 0 and hence a low density of states. Degeneracy in the conduction band with the ensuing changes in conduction phenomena occurs therefore at relatively low concentrations.

Until recently the values reported for the effective mass of holes in InSb were extremely varied[15][18]. The value now accepted ($m_h = 0.41_{mo}$) is in agreement with cyclotron resonance[66], density of states and theoretical[67] determinations. This effective mass results in more usual mobilities and density of states for holes.[18]

Carrier lifetimes are also a very important factor affecting the properties of p-n junctions. In InSb[4][5][68] the lifetime of electrons is very short, typically less than 5×10^{-9} sec at 77°K, while that of holes is typically 5×10^{-7} sec. In spite of the high electron mobility the very short lifetime causes the diffusion length for electrons to be very small ($\sim 10^{-3}$cm). It is in fact, almost identical with the hole diffusion length. These short diffusion lengths require that, for efficient operation, the base regions of transistor type devices and the sensitive regions of junction photodetectors must be very narrow.

In InSb a large amount of research has been done in high field conductivity and bulk breakdown[69]. This work has been stimulated by the unique properties of the electron-hole plasmas which arise after the onset of impact ionization.

The self pinch effect [70][71] of the plasma is likely to be observed in few other semiconductors. Plasma oscillations[72] and associated phenomena such as E-field hysteresis[73] have also generated both theoretical and experimental interest. The possibility of establishing a plasma at low fields by electron injection[73] and the ensuing magnetic properties of the negative resistance characteristic have already found application in madistor devices.[74] The value of these measurements to simple p-n junctions, however, is their application to the estimation of junction breakdown voltage. It is quite well established that the field for appreciable impact ionization in n-type InSb[71][75] is 180 ± 25 v/cm while that in p-type[76] is about 700 v/cm. These values, if realized in p-n junctions, would give breakdown voltages of a few tenths of a volt at most. Unlike Ge junctions, however, where the impact ionization field[77] gives a good estimate of breakdown voltage[78] particularly in lightly doped junctions, breakdown voltages in InSb junctions are several orders of magnitude higher than expected. This was first observed by Lee and Kaminsky[79] in 1960 but no reason has yet been advanced for this very large discrepancy.

The dielectric constant which is a determining factor in the capacitive effects of p-n junctions is 16.8[69] for InSb.

1.3 Diffusion in III-V Compounds.

Diffusion[80] of impurities into semiconductors is, without doubt, the most exact and reproducible method of forming p-n

junctions. The use of diffusion in producing very high frequency devices and integrated circuitry, where dimensions and impurity concentrations must be controlled within very fine limits, is a well known example. It is also necessary in forming alloyed p-n junctions where the abruptness of the junction is important, as in tunnel devices, to know the nature of impurity diffusion in order that the alloying cycle can be designed to achieve the necessary abruptness. It is important, therefore, in preparing any junction whether alloyed, grown or diffused to have a knowledge of impurity diffusion.

In diffusion studies two techniques are commonly used. The technique which provides the most information uses a radio-active isotope of the impurity for diffusion. The complete diffusion profile can then be obtained by careful grinding and counting. The p-n junction technique[81] measures the depth at which the junction is formed. (See Appendix A) It is necessary therefore in using this technique to diffuse an impurity of the opposite type to that already present in the crystal. To interpret the results of such measurements a distribution in the crystal for the diffused impurity must be assumed.

Prior to the commencement of this work a limited number of investigations had been carried out on diffusion in InAs, GaSb, GaAs and InSb. The earliest work was done in 1956 by Schillmann[82] on InAs. He investigated eight impurities using the p-n junction technique. Eisen and Birchenall[83] using radio-tracers measured self diffusion in InSb and GaSb.

Boltaks et al also used radio-tracers to investigate diffusion of indium, antimony and tellurium in InSb[84] and of indium, tin, antimony and tellerium in GaSb.[85] Hulme and Kemp[86] using the p-n junction technique and Goldstein[87] using radio tracers both investigated zinc diffusion in InSb and obtained widely divergent results. The first clear evidence of complications in the diffusion process in III-V compounds was the observation[88][89] that zinc diffusion from the vapor phase did not give a complementary error function distribution in GaAs.

At the beginning of this work a considerable amount was done to establish thecause of thermal conversion (i.e. the complete conversion from n-type to p-type during heating) and the nature of zinc and cadmium diffusion in InSb, using the p-n junction technique. These experiments are summarized in Appendix A. The primary interest of the work at that time was the possibility of fabricating an alloy-diffused transistor to examine the effects of the high electron mobility on transistor cut-off frequency. The results of the diffusion experiments showed that (a) unless the dilution of the diffusant was carefully controlled the penetration depth of the diffused layer was not reproducible, (b) diffusion could not be described in terms of error functions unless other mechanisms[90] such as surface rate limitation or out-diffusion were invoked, and (c) surface concentrations were so high that unless the diffusant was introduced in the form of a highly dilute alloy, surface

alloying could not be avoided. On the basis of this last observation the prospects for an alloy-diffused transistor were considered remote and attention was focused on alloyed junctions. It is ironical to note that about this time Henneke[65] reported an InSb alloy-diffused transistor which performed as well as similar germanium transistors but did not exhibit the increase in cut off frequency expected.

Recently work on diffusion in III-V compounds has centred mainly on GaAs and InSb. The work in GaAs has continued to show discrepancies and anomalies such as "uphill diffusion"[91]. The most successful treatments[92][93][69] of possible diffusion mechanisms indicate a considerable complexity. Radio tracer techniques have been used to investigate diffusion of zinc[31][94] cadmium[31][95][96], tin[94], gold[97], iron, cobalt and nickel[98] in InSb. The results on zinc and cadmium have continued to vary but the work by Kendall[31] has shown that the diffusion coefficient for these is extremely concentration-dependant. The diffusion coefficient for zinc, for example, is reported to vary by five orders of magnitude for concentrations ranging from $10^{18}cm^{-3}$ to $2 \times 10^{20}cm^{-3}$. It is thought that this phenomenon could explain the differences between the results of different workers. Kendall has also shown, in agreement with the present work, that for diffusion from the vapor phase surface concentrations are high ($>10^{18}cm^{-3}$) and the distribution is not a complementary error function. This last observation is also supported by the measurements of Minamoto[99] which indicated that zinc diffused junctions are abrupt.

1.4 p-n Junctions in III-V Compounds.

It is necessary before proceding to describe the present work to place it in the context of junction studies which have both preceded and parallelled it. To do this the general development of p-n junctions in III-V compounds will first be reviewed and then the specific developments relating to InSb junctions will be considered in more detail.

1.4.1 The Development of p-n junctions in III-V Compounds.

Except for InSb, where uncompensated impurity concentrations of $10^{14}cm^{-3}$ are easily achieved, III-V compounds cannot yet be prepared with impurity concentrations much less than $10^{16}cm^{-3}$[69]. Exceptions brought about by recent developments are GaAs grown epitaxially[100] ($\sim 2 \times 10^{15}cm^{-3}$) and InAs grown by a liquid encapsulation technique[101] ($\sim 10^{16}cm^{-3}$). GaSb has not yet been purified to better than $10^{17}cm^{-3}$. There are two main reasons for this rather slow advance since 1952.[2] One is the severe technological difficulties which arise in compound systems where one component has a high vapor pressure or is highly reactive. The other reason stems from the lack of vital application which so stimulated the advances in silicon and germanium. Initially a very considerable amount of work was done preparing and measuring the properties of III-V compounds. These measurements soon made it clearthat for many devices the advantages which III-V compounds might have over silicon and germanium did not merit the effort required to overcome the associated technical difficulties. For example, Jenny[102]

has examined both bipolar and unipolar transistors and found that the main advantages of III-V compounds arise from the possibility of higher operating temperatures. Because of the relatively poor material available and, to some extent, poor junction techniques, the investigations of p-n junctions in III-V compounds have been limited to effects such as photo response, tunneling and injection luminescence which are superimposed on, or external to, the primary electrical characteristic of the junction. On poor quality or heavily doped materials junction studies are of necessity limited to the above effects since some photo effects can be observed without rectification while tunneling and injection luminescence occur in heavily doped materials at current desities so high that they are not particularly sensitive to imperfections.

Many of the first junctions formed were used to study the photo response. [InSb[103], InAs[104], GaAs[105], InP[106]] The application of photo diodes as solar cells and the advantages of high energy gap III-V compounds have been discussed by Rappaport[107]. Since diodes with large rectification ratios can be produced in GaAs[108][109][110] and in InP[111] these have been used in solar cells and GaAs photo diodes giving 0.9 volt[107] have in this sense out-performed silicon. However, the maximum efficiency of silicon at 15% has not yet been reached by either GaAs[107] at 12% or InP at 2%[107].

Point contact diodes were also investigated in AlSb[112], GaP[113], GaAs[114] and InSb[3]. Some of the interest in

point contact diodes stemmed from the possibility of obtaining improved microwave rectifiers. Only GaAs has shown any promise as yet and in general the characteristics of the diodes are not understood.

Other investigations have centred on tunnel diodes, using them largely as a tool to investigate the fine structure of the tunneling processes. Hall et al[115][116][117] have studied tunneling in InSb, InAs, InP, GaSb, GaAs and GaP. Many other investigations have been carried out particularly on InSb and GaAs. A good review of tunneling in p-n junctions of III-V compounds is given by Madelung.[69]

Heterojunctions or junctions between two different semiconductors have also received considerable attention recently. In heterojunctions not only is the conduction process different in the two regions but the energy gap is also different. These junctions are particularly suited for transistor emitters and for opto-electronic devices. Ge- GaAs[118][119] heterojunctions have been studied in some detail. Others[120][121] have also been investigated and the technology of these junctions is developing.

Perhaps the most significant advance in III-V compounds was the discovery of stimulated emission in forward biased GaAs diodes in 1962.[122][123][124] The importance of the p-n junction laser both as a research tool and as a possible device has given fresh impetus to the search for techniques to overcome the difficulties associated with III-V compounds. Laser action

has since been reported in InSb[125], InP[126] and InAs[127] junctions and in InGa-As[128], InP-As[129] and GaAs-P[130] mixed junctions. The reports on p-n junction lasers, and GaAs lasers in particular, are voluminous. A comprehensive review has been published recently by Burns and Nathan.[131]

Transistor action has been observed in InP[102][132], InSb[65] and GaAs[133][134]. Because of the short carrier lifetimes transistors in these materials must have very narrow base regions and therefore the fabrication of transistors represents a considerable technological achievement. Both double diffused[133] and alloy-diffused[134] transistors with base widths less than one micron have been reported in GaAs. These transistors have shown current gains from 6 to 15, alpha cut off frequencies in excess of 1 Gc/s and switching times shorter than those of silicon devices designed for high speed switching. Reynolds and Lilburne[132] reported power gains of 20 in point contact InP transistors. Jenny[102] formed transistors in InP by a surface diffusion technique. Using a point contact emitter on a 0.3 micron base layer a current gain of 13 and a power gain of 36 db were observed. Henneke's[65] alloy-diffused InSb transistors will be discussed in section 1.4.2.

GaAs diodes which have lower reverse currents than silicon diodes and extremely short recovery times ($< 10^{-9}$ sec)[18][108] have found commercial application as switching diodes.

Except for the tunnel diodes and point contact diodes most of the junction devices described above were formed by diffusion

techniques. For the reasons already described, few attempts have been made to relate the primary electrical characteristics to the bulk properties of the materials and where such attempts have been made they have in general been more speculative than definitive.

1.4.2. Junctions in InSb.

The early work on InSb p-n junctions centred largely on photodetectors and tunnel diodes. In 1955 Mitchell et al[103] studied the photo-voltaic effect in grown junctions. Later Avery et al[135] examined junctions formed by a remelt technique. The sensitivity of these junctions as infra-red detectors was shown to be comparable to that of PbTe cells and the response time was less than 2×10^{-7} sec. Further work was also reported on alloyed junctions[136] where the maximum observed photo-voltage was 60 mv and on diffused junctions [137] which were shown to be extremely sensitive infra-red detectors. With one exception[136] no rectification properties were reported in any of this work. Rectifying junctions were first reported by Lee and Kaminsky[138] in 1959. They prepared junctions by either alloying or diffusion. Rectification ratios of about 2×10^{3} at 0.25 volts and a maximum breakdown of about 7 volts were observed. The same authors[139] also reported anomalous capacitance in alloyed junctions and were the first to point out the unexpectedly high breakdown voltages observed. Marfaing[140] reported very poor electrical characteristics on diffused diodes but made considerable effort to interpret them.

The characteristics of narrow junctions and tunnel diodes were first reported in 1960. Internal field emission was investigated by Chynoweth and Logan[141] and very fast tunnel diodes were reported by Batdorf et al.[142] Reports of magnetic effects on tunneling[143][144], of low temperature electric field effects on tunneling[145] and of the tunneling fine structure[115] soon followed.

This was the state of p-n junction development in InSb when the work to be described in this thesis began. Since good InSb was available and the diodes reported gave reverse currents four to six orders of magnitude higher than could reasonably have been expected, it was clear that good junction preparation techniques were required. For a study of the electrical characteristics of a p-n junction to be meaningful it must be possible to relate them to the properties of the materials used. Therefore a study of junctions whose characteristics result from unpredictable and uncontrollable imperfections or poor fabrication techniques is of little value.

During the course of this work further investigations of InSb junctions in other laboratories were reported. Some further work on tunnel diodes[146][147] was published but Hulme[148] has shown that, because of the cooling required, the possible applications of InSb tunnel diodes are limited. Marfaing[149] published results on surface effects in photodetectors and recently[150] on the characteristics of diffused diodes. Minamoto[99] reported normal capacitance in diffused abrupt diodes and suggested that

the anomalous capacity observed by Lee and Kaminsky[139] was due to an inversion layer which proper etching would prevent. Minamoto[55] also showed that diodes formed on A {111} surfaces had inferior regrowth and electrical properties to those formed on B {111} surfaces. The electrical characteristics of grown junctions were measured by Gorton et al[151] and it was demonstrated that the reverse currents were dominated by surface leakage at low temperatures. Electrical characteristics have also been reported by Galavanov.[152]

A group working at M.I.T. have also reported a considerable amount of work on InSb diodes. They have been concerned largely with forward bias phenomena such as tunneling[143] and the negative resistance in n^+p diodes[153] which forms the basis of madistor[74] devices. They have also reported laser action in InSb diodes[125] and, perhaps even more significant, coherent emission in bulk InSb.[154]

Allen et al[155] have reported InSb varactor diodes with cut off frequencies as high as 400 Gc/s. When incorporated in a 5 Gc/s parametric amplifier a noise temperature approaching the theoretical optimum ($\simeq 8^{\circ}K$) was realised.

Recently Mueller and Jacobson[156] have shown that an anodized surface layer on InSb can be changed from n to p by illumination. This has been used to obtain neutral surface layers in InSb diodes[157] and the resulting characteristics have been interpreted in terms of the theory of Sah et al[158] at low forward bias.

A group at Texas Instruments has made what would seem to be the most fruitful study of InSb junction devices. They have reported both diffused and alloyed diodes[159] with rectification ratios of $\sim 10^6$ at 0.25 volts. This, apart from present work[160] is about three orders of magnitude better than any other InSb diodes where full electrical characteristics are reported. At low forward bias the electrical characteristics were interpreted as the sum of leakage, recombination and diffusion currents. Shortly after Stocker reported these diodes Henneke[65] published results on alloy-diffused transistors with current gains as high as 500 and cut off frequencies up to 300 Mc/s. The transistor was followed by a report on InSb diffused photo-diodes[161][162] with noise figures so low that the performance of an ideal detector was closely approached. This work indicated that the group at Texas Instruments had achieved a very advanced state of technology in diffusion, alloying and chemical treatment of InSb. However, no technical details have been published and the above reports give virtually no information on the materials used or the observed characteristics.

From the above discussion it is apparent that although some of the effects observed in InSb p-n junctions are qualitatively if not quantitatively understood and very promising devices have been made, the nature of some of the electrical characteristics is still not known. Also reliable, controlled junction formation techniques are not general knowledge and where such techniques have been developed they are guarded. The objectives set down early in this work as described in section 1.1 appear, therefore, to be valid and desirable.

CHAPTER 2

Theoretical Considerations.

The theories required to interpret the results presented in Chapter 5 will be considered in this chapter. Theories for the behaviour of p-n junctions and refinements of such theories are numerous, but concise statements of the assumptions used in these theories are not often made. For the most part existing diode theories will be presented in this chapter but the implications and limitations of the assumptions used will be carefully assessed. One exception is the treatment of medium level injection given in section 2.1.5. No previously derived analytic expression for diode currents in this injection region has been found.

2.1. Direct Current Characteristics of p-n Junctions.

2.1.1. Formulation of p-n Junction Theory.

The following equations govern steady state one dimensional current flow in semiconductors[163][164][165]:

$$J_n = q\left(n\mu_n \frac{d\psi}{dx} + D_n \frac{dn}{dx}\right) = q\mu_n n \frac{d\psi_n}{dx} \quad \ldots\ (2.1)$$

$$J_p = q\left(p\mu_p \frac{d\psi}{dx} - D_p \frac{dp}{dx}\right) = q\mu_p p \frac{d\psi_p}{dx} \quad \ldots\ (2.2)$$

$$\frac{dn}{dt} = (g-r) + \frac{1}{q}\frac{dJ_n}{dx} = 0 \quad \ldots\ (2.3)$$

$$\frac{dp}{dt} = (g-r) - \frac{1}{q}\frac{dJ_p}{dx} = 0 \quad \ldots\ (2.4)$$

$$\frac{d^2\psi}{dx^2} = \frac{q}{\varepsilon}(p-n+N_D-N_A) \quad \ldots\ (2.5)$$

where the symbols have the following meanings:

ψ_n - the electrochemical or quasi-Fermi potential for electrons.

ψ_p - the quasi-Fermi potential for holes.

ψ - the electrical potential corresponding to the intrinsic Fermi level.

g - generation rate of holes and electrons.

r - recombination rate of holes and electrons.

N_D - ionized donor concentration.

N_A - ionized acceptor concentration.

J_n - electron current density.

J_p - hole current density.

μ_n - electron mobility.

μ_p - hole mobility.

D_n - electron diffusivity.

D_p - hole diffusivity.

n - electron concentration.

p - hole concentration.

q - electronic charge.

ε - dielectric constant.

t - time.

x - position in one dimension.

Further symbols will be defined as they are introduced in the text.

It is possible to obtain a set of auxiliary functions which reduce the number of unknowns in equations (2.1) - (2.5). The

carrier concentrations, mobilities and diffusivities and the number of ionized impurities can be expressed as functions of the potentials ψ_n, ψ_p and ψ, the temperature T and the total donor and acceptor concentrations N_{DT} and N_{AT}. The net generation rate (g-r) may be given as a function of the same potentials, the temperature and the densities and energy levels of the recombination centres in the material. If the recombination process and the spacial variation of impurities are known and the system is isothermal, the above relationships can be used to reduce equations (2.1) - (2.5) to three simultaneous nonlinear differential equations involving three dependent variables ψ_n, ψ_p and ψ. In theory it is possible to obtain solutions for these variables and hence for current flow for any specified set of auxiliary functions. In practise, however, it is not possible in the case of p-n junctions to obtain analytic solutions even when these functions are made as simple as realistically possible.

For example, in the case of an abrupt p-n junction the following assumptions are made in order to obtain a simple set of auxiliary functions:

(1) For both the conduction band and valence band the constant energy surfaces are spherically symmetric and parabolically distributed in momentum space and at no position in real space does n or p become degenerate. Then the carrier concentrations can be described by Boltzmann distributions and

$$n = n_i \exp \frac{q(\psi - \psi_n)}{kT} \quad \ldots\ldots (2.6a)$$

$$p = n_i \exp \frac{q(\psi_p - \psi)}{kT} \quad \ldots\ldots (2.6b)$$

where n_i is the intrinsic carrier concentration and k is Boltzmann's constant.

(2) The mobilities and diffusivities are constants.

(3) The net rate of generation over recombination is determined by a single level of trapping or recombination centres[166][167].

Then:

$$(g-r) = \frac{np - n_i^2}{(n+n_t)\,\tau_p + (p+p_t)\,\tau_n} \quad \ldots\ldots (2.7)$$

where n_t and p_t are the number of electrons in the conduction band and holes in the valence band when the Fermi level lies at the trapping level E_t, and τ_p and τ_n are the typical capture times for a hole or electron when the trapping centres contain no carrier of the kind in question.

(4) All donors or acceptors are ionized and the p-n junction is an abrupt transition between two regions of constant impurity concentration.

Then:

$$N_D - N_A = a \qquad x > 0 \quad \ldots\ldots (2.8a)$$

$$N_D - N_A = -b \qquad x < 0 \quad \ldots\ldots (2.8b)$$

These four assumptions with further simplification of assumption (3) have been used [165][168] in attempts to obtain

solutions from equations (2.1) - (2.5) but even in the zero bias case, where $\psi_n = \psi_p$ = constant, the solution cannot be expressed in analytic form. These assumptions can be used, however, to reduce the problem to a single integral equation which can be solved numerically. This has been done[169][170][171][172] in some instances to examine the validity of the more simplified models which have been used to derive analytic expressions for current flow in p-n junctions. In comparison to any real situation the conditions described above are, of course, a considerable simplification. Any real p-n junction is a three dimensional device involving surfaces which can be at least as complex as the junction itself and often much less reproducible. In alloyed junctions where assumption (4) is most likely to be satisfactory, assumption (1) is most in doubt. Assumption (2) certainly does not correspond to any real situation although errors arising from the assumption may be small. Assumption (3) is probably quite good for p-n junctions within specified limits of injection but, at least in some semiconductors[4][5][68] it has not been possible to interpret recombination observations in terms of a single level model.

In view of this, where real results are to be interpreted, there seems little justification in pursuing the very difficult problem of applying equations (2.1) - (2.5) rigorously to such a model. In this discussion it is intended to follow the self consistent approach where additional reasonable, but not necessarily justifiable, assumptions are made which permit one to obtain

analytic solutions. These solutions are then substituted back into the basic equations to determine the validity of the initial assumptions. Since such assumptions must be checked for each particular situation it will be assumed that they are valid throughout the theoretical development. In a later chapter when the theory is applied these assumptions will be tested.

In the case of the abrupt p-n junction assumptions (1) - (4) are retained and the following further assumptions are made:

(5) The diode can be treated as two well defined neutral regions separated by a narrow space charge region. This implies that, outside the space charge region, equation (2.5) is approximately zero.

(6) The carriers at the two space charge region boundaries are in equilibrium with each other. That is to say that when an external voltage is applied the excess carrier concentration in the space charge region is essentially constant throughout and not significantly affected by recombination. In mathematical terms it implies that ψ_n and ψ_p are essentially constant through the space charge region or that changes in ψ_n or ψ_p are small compared to $\psi_n - \psi_p$.

These six assumptions are, with slight variations particularly in assumption (4), the basis of previous analytic treatments of p-n junctions. In the following sections assumptions (1) to (4) will be critically assessed and then used in conjunction with assumptions (5) and (6) to solve equations (2.1) - (2.5). The diffusion characteristics for p-n junctions

for both low and high level conditions will be determined and with one further assumption the current due to recombination in the space charge region will be considered. Such an analysis results in three expressions involving three currents which must be combined to give the overall current characteristic. A method for combining these and the assumptions involved are treated in section 2.1.5. Finally reverse currents and breakdown are considered briefly. It will be assumed throughout this treatment that surface effects can be neglected.

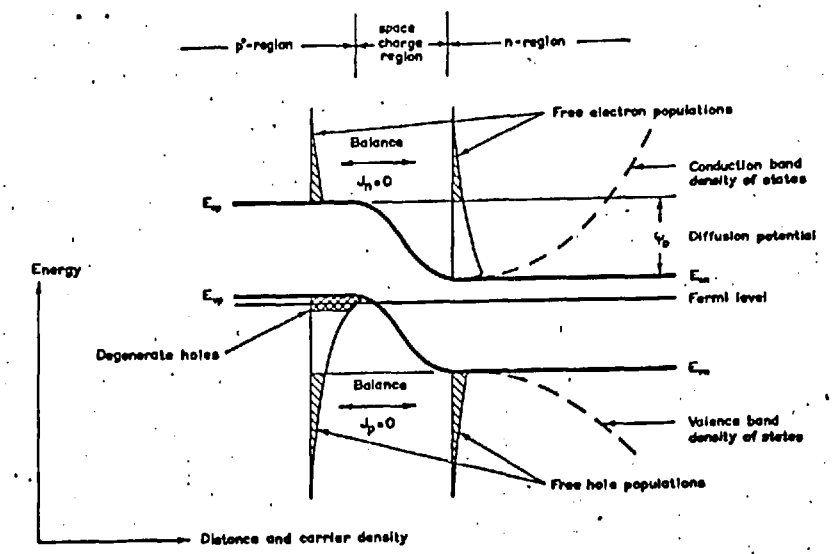
2.1.2. Low Level Diffusion Currents.

The current of carriers which diffuse from the n-region into the p-region and vice versa when a voltage is applied across a p-n junction are considered in this section under conditions where the number of carriers entering either region is not sufficient to significantly perturb the majority carrier concentration. Assumptions (1) - (4) from section 2.1.1. are then applied as follows:

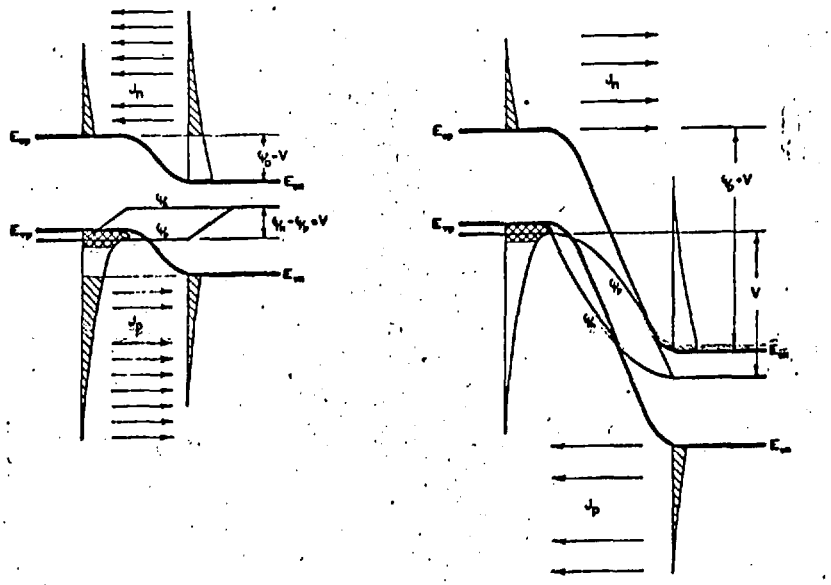
(1) The energy surfaces are spherically symmetric and parabolically distributed in k-space. The density of states effective mass will be used in determining the intrinsic carrier concentration and thus this part of the assumption can be considered valid over the range of carrier concentrations for which this effective mass accurately describes the density of states. The added stipulation that neither n nor p are anywhere degenerate is often not true for at least one region in alloyed p-n junctions. The validity of Boltzmann boundary conditions

for degenerate regions has been the subject of argument[173][174] and has been proved untrue. It will be shown, however, first intuitively and then analytically, that with certain modifications the Boltzmann boundary conditions can be used even in cases where both regions are degenerate.

The hole current in a p-n junction arises from an imbalance between the hole population on the n-type side of the junction and that part of the hole population on the p-type side with hole energies higher than E_{vn}, the lowest energy state in the valence band of the n-type region. The electron current arises from a similar imbalance between the electron population of the p-type region and that part of the electron population in the n-type region having electron energies in excess of E_{cp}, the lowest energy state in the conduction band of the p-type region. The junction currents, except for those in large reverse bias, are determined therefore by parts only of the majority carrier populations. This is best illustrated by referring to Figure 2 where the horizontal axis represents both distance and carrier density thus permitting the carrier populations as a function of energy to be plotted at any position inside or outside of the junction. The parts of the carrier populations involved in the balance at zero bias or the imbalance for current flow are distinguished by cross-hatching. When no external bias is applied no current flows and these populations exactly balance each other as shown in Figure 2(a). This situation implies a built in potential, ψ_D, which displaces the bands in the two



(a) Zero Bias



(b) Forward Bias (c) Reverse Bias

Figure 2. The carrier populations on either side of the space charge region of a p^+n junction plotted on the potential energy diagram and showing the balance (a) at zero bias and the imbalance (b) in forward bias and (c) in reverse bias between nondegenerate sections of the populations indicated by cross-hatching.

regions with respect to each other to achieve balance. Any applied potential, V, disturbs this balance and causes one population to shift with respect to the other. The resulting currents are illustrated for forward and reverse bias in Figure 2(b) and 2(c). It is obvious from the p^+n junction illustrated in Figure 2 that, although the p^+ region is degenerate, the current which flows when a bias is applied is caused by an imbalance between sections of the carrier populations which are nondegenerate. This remains true except at a forward bias large enough to cause the Fermi level in the p^+ region to come within say kT of E_{vn}. Thus the exponential form of assumption (1) can be considered valid for p-n junctions even in cases where n or/and p are degenerate provided the applied voltage does not cause the Fermi level of either degenerate band to come within kT of the lowest energy state in the same band of the opposite material. In an asymmetric junction with similar effective masses this makes the form of assumption (1) generally valid for both low and high level conditions. In very high level conditions where it is not valid a detailed balance integral would be required to determine the junction currents.

This intuitive argument has established that, even for high level currents, the distribution of carriers can be considered exponential (i.e. Boltzmann distributions) in the degenerate regions of a p-n junction. It does not, however, establish the nature of the magnitude of these functions. It is on this point that error could arise since it is well known[175] that

$n_o p_o \neq n_i^2$ in degenerate material. In fact, if the energy gap is greater than several kT

$$n_o p_o = (n_{i\ deg})^2 = \frac{n_i^2 F_{\frac{1}{2}} (\partial/kt)}{\exp (\partial/kT)} \lessapprox n_i^2 \qquad \ldots . (2.9)$$

where $F_{\frac{1}{2}}$ is the Fermi-Dirac integral of order one half, ∂ is the energy difference between the degenerate band edge and the Fermi level and n_o and p_o are the equilibrium electron and hole concentrations. For instance, if the Fermi level is 2kT inside the band.

$$n_o p_o = (n_{i\ deg})^2 = (\frac{n_i}{1.6})^2$$

If a degenerate p^+ region with hole density p_{po} and electron density n_{po} is considered, as illustrated in Figure 2, and it is assumed that the injected electron density $\Delta n = \Delta p \ll p_{po}$, then it is obvious that, since $p_{po} \gg n_{no}$,

$$p_p = p_{po} + \Delta p \simeq p_{po} \qquad \ldots . (2.10a)$$

$$n_p \simeq \frac{(n_{i\ deg})^2}{p_{po}} \exp \frac{(\psi_n - \psi_p)q}{kT} \qquad \ldots . (2.10b)$$

These are the modified Boltzmann boundary conditions for the degenerate p^+ region. It should be noted that medium level injection ($\Delta n \simeq p_{po}$) cannot be realized in the degenerate p^+ region of a p^+n junction because there are not sufficient electrons in the n region to significantly perturb the hole population in

the p^+ region even if all of the electrons were injected. Therefore the low level Boltzmann boundary conditions described above can be considered to apply in the heavily doped region of an asymmetric junction for all current levels.

If now the nondegenerate n-type region of such a junction is considered it is obvious that, short of hole injection levels which cause degeneracy in either the conduction or the valence band, the Boltzmann boundary conditions as stated in equations (2.6) apply. In the low level injection case where $\Delta p = \Delta n \ll n_{no}$, the equilibrium electron density, equations (2.6) can be arranged to give:

$$n_n \simeq n_{no} \qquad \ldots\ldots (2.11a)$$

and

$$p_n \simeq \frac{(n_i)^2}{n_{no}} \exp \frac{(\psi_n - \psi_p)q}{kT} \qquad \ldots\ldots (2.11b)$$

These can be seen to be exactly analagous to the low level boundary conditions prevailing in the degenerate p^+ region as given by equations (2.10). For the high level injection case where

$\Delta p = \Delta n \gg n_{no}$ equations (2.6) give:

$$p_n = n_n = n_i \exp \frac{(\psi_n - \psi_p)q}{2kT} \qquad \ldots\ldots (2.12)$$

provided $\Delta p < n_{deg}$ or p_{deg}, the carrier concentrations required to cause degeneracy in the conduction band or the valence band. The condition of injection where $\Delta p \geqslant n_{deg}$ or p_{deg} will be termed very high level injection.

The analytic considerations above substantiate the intuitive conclusions drawn earlier about the Boltzmann boundary conditions continuing to apply to an asymmetric junction even in conditions of high level injection.

(2) The mobilities and diffusivities are constants. The use of assumptions (5) and (6) which divide a p-n junction into two separate regions make it possible to modify this assumption so that these constants need not necessarily be the same for both regions. For the sake of notational simplicity the difference will not be designated in this treatment but it will be used in applying the expressions in a later chapter.

It is necessary in obtaining solutions to equations (2.1) - (2.5) also to know the relationship between mobilities and diffusivities. For carriers in a nondegenerate band this ratio is given by the Einstein relation ($D = \mu \frac{kT}{q}$). The relation is changed in a degenerate band[15] and is given by:

$$D = \mu \frac{kT}{q} \frac{F_{\frac{1}{2}}(\partial / kT)}{F_{-\frac{1}{2}}(\partial / kT)} \times 2. \qquad \ldots . \quad (2.13)$$

where ∂ is the energy difference between the band edge and the Fermi level. It is clear that this reduces to the Einstein relation for a nondegenerate band. For a degenerate band where the Fermi level is 2kT inside the band:

$$D \simeq \mu \left(\frac{2kT}{q}\right)$$

Since such a correction can only be made for a specific case of degeneracy, the Einstein relation will be used in this

treatment. However, in applying any expression derived using the Einstein relation it is possible to correct the value of diffusivity for degeneracy by the use of equation (2.13).

In this connection it is important to note that in low level conditions the diffusion characteristic of a p-n junction is entirely determined by minority carrier flow and therefore the use of the Einstein relation is justified even in degenerate regions. Furthermore, it has already been pointed out that medium level injection is not likely to be realised in the heavily doped region of an asymmetric junction and therefore the Einstein relation will continue to apply to the governing minority carrier for all currents even if the region is degenerate. In the lightly doped region when medium or high level injection is realised the current flow is determined by the ambipolar[164][178] mobilities and diffusivities but the Einstein relation continues to apply. Only in very high level injection conditions where the injected carrier density causes degeneracy would the Einstein relation cease to apply. This is interesting because it implies that in spite of the inapplicability of Boltzmann statistics and the Einstein relation to degenerate regions their use introduces negligible error into the derived current-voltage relationship for a p-n junction.

(3) The net rate of generation over recombination is given by equation (2.7) assuming low injection levels:

i.e. that $\Delta n = \Delta p \ll n_{no} > n_t$ in the n-region

and $\Delta n = \Delta p \ll p_{po} > p_t$ in the p-region

then

$$(g-r) = \frac{\Delta p}{\tau_p} = \frac{p-p_{no}}{\tau_p} \quad \text{in the n-region} \qquad \ldots\ldots (2.14a)$$

and

$$(g-r) = \frac{\Delta n}{\tau_n} = \frac{n-n_{po}}{\tau_n} \quad \text{in the p-region} \qquad \ldots\ldots (2.14b)$$

Most recombination processes give this form of behaviour at low levels of injection so equations (2.14) must be acceptable. (4) The justification for assuming an abrupt p-n junction can only be ascertained for specific cases. In general no significant error is involved in assuming that all donors or acceptors are ionized unless it can be shown that very deep donor or acceptor levels are present.

Having determined the validity of assumptions (1) - (4) for this case it is possible with the aid of the additional assumptions (5) and (6) to solve equations (2.1) to (2.5). This is discussed in more detail in Appendix B.2. The resultant low level current voltage expression is the ideal diode equation given by Shockley[168]

$$J_{DL} = q\left[\sqrt{\frac{D_p}{\tau_p}}\, p_{no} + \sqrt{\frac{D_n}{\tau_n}}\, n_{po}\right] \left(\exp\left(\frac{qV}{kT}\right) - 1\right) \qquad \ldots\ldots (2.15)$$

where V is the voltage across the space charge region and is approximately equal to the applied voltage for low level conditions.

2.1.3. High level diffusion currents.

High level currents result when the injected level of carriers greatly exceeds the majority carrier concentration in

either region of a p-n junction. It is not possible to obtain high level currents in a symmetrical junction. For this reason, only the asymmetric case will be considered where high level injection is realized only in the region of low carrier concentration. This case has been considered by a number of authors (158)(165)(176)(177) using slightly differing assumptions but in each case obtaining the same voltage dependence for the current.

Before treating this case the validity of assumptions (1)-(4) must again be examined.

(1) The treatment of the Boltzmann boundary conditions in section 2.1.2. established that they are valid for high level injection provided the injected carrier densities do not cause degeneracy in either minority carrier band.

(2) The assumption of constant mobilities and diffusivities remains valid although as shown in the derivation (Appendix B.3) the carrier flow in the high resistivity region is determined by the ambipolar mobility and diffusivity at these levels of injection. Furthermore the Einstein relation remains valid as discussed in section 2.1.2.

(3) Since the region of high carrier concentration remains in low level conditions the assumption as applied in section 2.1.2. remains valid. In the region of low carrier concentration, however, high level injection implies that:

$$\Delta n = \Delta p \gg n_{no} > n_t \quad \text{for an n-type region}$$

and

$$\Delta n = \Delta p \gg p_{po} > p_t \quad \text{for a p-type region.}$$

Thus equation (2.7) reduces to:

$$(g-r) = \frac{\Delta n}{\tau_n + \tau_p} = \frac{\Delta p}{\tau_n + \tau_p} \quad \text{for either case} \quad \ldots . \quad (2.16)$$

It is probable that other recombinations processes become important at these high injection levels and thus the validity of this assumption must be placed in doubt. Other processes such as Auger or radiative recombination[175] which could be significant are related in a nonlinear fashion to the excess carrier densities for high level injection. Recombination which is related in such a way to excess densities makes analytic solutions of equations (2.1) - (2.5) impossible. In this derivation therefore, equation (2.16) will be considered valid but the limitations of this assumption must be carefully considered in any application of the resulting expression.

(4) The assumption of an abrupt junction where donors and acceptors are fully ionized is valid as discussed in section 2.1.2.

Using these assumptions combined with assumptions (5) and (6) equations (2.1) - (2.5) can be solved as shown in Appendix B.3 to give the following expression for a p^+n junction:

$$J_{DH} = q\sqrt{\frac{2D_nD_p}{(D_n+D_p)(\tau_n+\tau_p)}}\, n_i \exp\left(\frac{qV}{2kT}\right) + q\sqrt{\frac{D_n}{\tau_n}}\, n_{po}\left(\exp\left(\frac{qV}{kT}\right)-1\right) \quad (2.17)$$

or for an n^+p junction

$$J_{DH} = q\sqrt{\frac{2D_nD_p}{(D_n+D_p)(\tau_n+\tau_p)}}\, n_i \exp\left(\frac{qV}{2kT}\right) + q\sqrt{\frac{D_p}{\tau_p}}\, p_{no}\left(\exp\left(\frac{qV}{kT}\right)-1\right) \quad (2.18)$$

A comparison of these equations with equation (2.15) shows that for either case the nature of the diffusion current in the region of low carrier concentration changes as the injection level changes.

2.1.4. Recombination Currents in Forward Bias.

This current arises from the recombination of electrons and holes as they traverse the space charge region. It has been treated in considerable detail by Sah, Noyce and Shockley[158] and is determined by integrating the net recombination rate over the space charge region. To perform the integration the following assumptions are required:

(1) The carrier concentrations in the space charge region are given by equations (2.6). It is obvious from Figure 2(b) that for small forward bias this is true without qualification over the greater part of the space charge region even when one region is degenerate. It can be shown, however, that the recombination current is dependent on the excess carrier concentrations in the space charge region and the treatment in section 2.1.2. established that these continue to be governed by Boltzmann statistics even in large forward bias.

(3) The net recombination rate is given by equation (2.7). This has also been established with certain reservations for conditions of high forward bias.

These two assumptions coupled with assumption (6) which establishes the variation of the quasi-Fermi potentials in the space charge region are not sufficient to permit the integration

of equation (2.7). A further assumption must be made with regard to the electrostatic potential.

(7) The electrostatic potential is assumed to vary linearly with distance through the space charge region. This assumption cannot, of course, be correct for it implies that $\frac{d^2\psi}{dx^2} = 0$ in the space charge region, and this is only true for a neutral region. However, the error is not as serious as this would suggest since assumption (6) establishes the primary character of the integral. Assumption (7) is necessary to permit integration and it allows an estimation of the variation of recombination rate through the junction. The term whicharises as a result of assumption (7) is only slightly voltage dependent in forward bias and hence, except for a possible small error in an essentially constant term, the resulting expression should be reasonably accurate. In fact, using this approximation for an abrupt junction is much better than using it in the linear graded junction treated by Sah et al.[158]

With these assumptions the integral of equation (2.7) can be evaluated. If the trapping level is near the centre of the energy gap the resulting recombination current is.[158]

$$J_{rg} = \frac{n_i \pi}{\sqrt{\tau_n \tau_p}} kTW \, (\psi_D - V)^{-\frac{1}{2}} \sin h\left(\frac{qV}{2kT}\right)$$

where W is the width of the space charge region. A value for W can be determined by combining assumption (4) and (5) and integrating equation (2.5). Then

$$W = \left(\frac{2\varepsilon}{q}\right)^{\frac{1}{2}} \left(\frac{1}{N_D} + \frac{1}{N_A}\right)^{\frac{1}{2}} (\psi_D - V)^{\frac{1}{2}} \quad \ldots\ldots (2.19)$$

where N_D is the net donor concentration in the n-region and N_A is the net acceptor concentration in the p-region. Thus:

$$J_{rg} = \frac{\pi n_i kT}{\sqrt{\tau_n \tau_p}} \left(\frac{2\varepsilon}{q}\right)^{\frac{1}{2}} \left(\frac{1}{N_D} + \frac{1}{N_A}\right)^{\frac{1}{2}} (\psi_D - V)^{-\frac{1}{2}} \sinh\left(\frac{qV}{2kT}\right) \quad \ldots\ldots (2.20)$$

It should be noted that in determining the value of ψ_D for given values of N_D and N_A it is important if either region is degernerate to use Fermi statistics.

2.1.5. Complete Forward Bias Current Dependence.

The preceding sections have made it clear that in any p-n junction three currents must be considered; the diffusion current of electrons, the diffusion current of holes and the current due to recombination or generation in the space charge region. It is obvious from the form of the solutions for both high level and low level injection that the electron and hole currents through the space charge region can be considered independent. For this reason they are treated as parallel currents in equations (2.15), (2.17) and (2.18). If assumption (6) is valid in the presence of recombination, and it has been demonstrated to be valid for silicon junctions[158], then the recombination current can be considered independent of the diffusion currents. It is added to the diffusion currents therefore, as a parallel current. Low level current flow is thus described by the sum of equations (2.15) and (2.20), while high level current flow

is described by the sum of equations (2.20) and (2.17) or (2.18).

The absence of any expression which describes the important region of medium level injection and which would therefore bridge the gap between the expressions for low level and high level injection is immediately obvious. Medium level injection has in general been avoided in all treatments of p-n junctions except in cases where the mathematical difficulties arising from its consideration have been described. Even if assumptions (5) and (6) are used, analytic solutions to the transport equations cannot be obtained in medium level injection where the Boltzmann boundary conditions[179][180] carrier mobilities, diffusivities and lifetimes[177] all become complicated functions of the injection level. Because of the lack of any theory for medium level injection, attempts to interpret measurements on real diodes have been limited to low level conditions. Where characteristics have exhibited high level injection[165][176][177] interpretation has been in essence simply to point out that the characteristic tended to exhibit the expected form of voltage dependence.

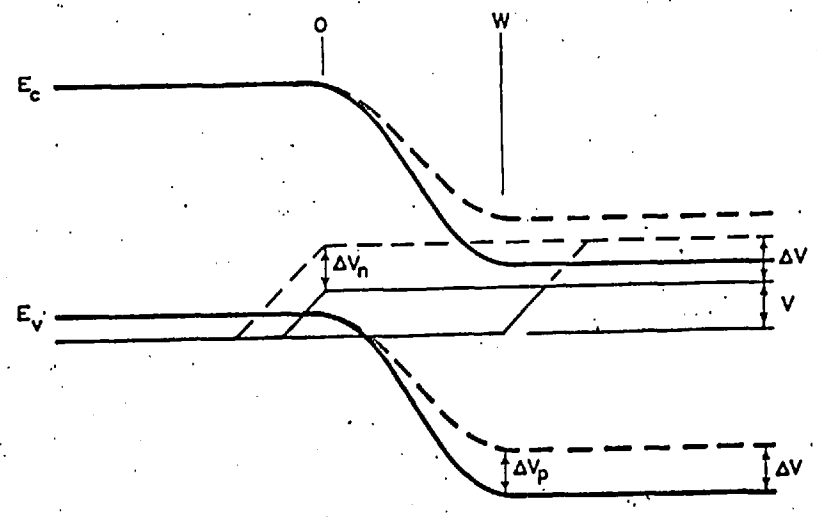
In this section the physical processes responsible for high and medium level injection will be considered. On the basis of these considerations one further assumption similar in nature to assumptions (5) and (6) will be made and used to obtain an expression for medium level injection current. The resulting expression willbe compared to numerical solutions of the transport equations to establish its validity.

If a p^+n junction is considered similar to that in Figure 2, it is apparent from the boundary conditions in the n-region for low and high level injection that two processes are affected by a change in voltage across the space charge region:

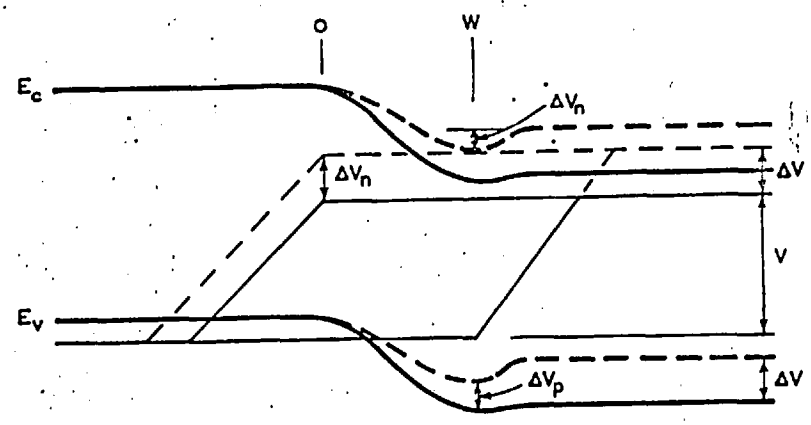
(1) Minority carriers are injected into both the n and p regions

(2) Majority carriers are accumulated in the injection regions to maintain neutrality. At all current levels part of the applied voltage is absorbed by each of these processes and the total junction voltage is therefore the sum of the two components.

This is best illustrated in the diagrams of Figure 3, where the Fermi level in the p-type region is chosen as the reference potential. At all low injection levels, Figure 3(a), an extremely small increase in the majority carrier concentration is required to maintain neutrality. Thus in the n-region at $x = W$, the change in potential for hole injection $\Delta V_p = \Delta V$, the change in applied potential. The same applies for electrons in the p-region at $x = 0$. This implies that in low level injection the portion of the applied voltage required to neutralize the injected carriers is negligible, in agreement with previous considerations. In medium level injection, Figure 3(b), a portion ΔV_n of the change in applied potential ΔV is required to accumulate electrons to neutralize the injected holes at $x = W$. It is obvious from Figure 3(b) that the change of injection potential $\Delta V_p = \Delta V - \Delta V_n$. The solutions for low and high level injection suggest that ΔV_n varies from zero at low injection to $\frac{\Delta V}{2}$ in high injection.



(a) Low Level Injection



(b) Medium Level Injection

Figure 3. The division of a change of applied potential, ΔV, between the establishment of a level of injected carriers and the accumulation of carriers to maintain neutrality (a) in low level conditions where ΔV equals the change in injection potential for holes at W and for electrons at O, and (b) in medium level where ΔV equals the change in injection potential for electrons at O but at W, $\Delta V = \Delta V_p + \Delta V_n$ the sum of the change in injection potential for holes and the accumulation potential for electrons.

These observations are important since they imply that the current flow in a p-n junction is a function of two potentials which add to give the applied potential. Such a situation is analagous to two series conductances. It is therefore reasonable to assume that each of the diffusion currents in a p-n junction is determined by two voltage dependent series conductances. In a p^+n junction the hole current at low injection is determined entirely by what might be considered the simple injection conductivity, g_a. At high injection it is determined entirely by what can be termed the charge neutrality conductivity, g_b. The overall conductance g is therefore the series combination of g_a and g_b.

$$g = \frac{g_a g_b}{g_a + g_b}$$

Using this form of expression to combine the high and low level solutions for the hole current in a p^+n junction:

$$J_{pD} = \frac{J_{Lp}\, J_{Hp} \left[\exp\left(\frac{3qV}{2kT}\right) - \exp\left(\frac{qV}{2kT}\right) \right]}{J_{Lp} \left[\exp\left(\frac{qV}{kT}\right) - 1 \right] + J_{Hp} \exp\left(\frac{qV}{2kT}\right)} \qquad \ldots\ (2.21)$$

where $$J_{Lp} = q \sqrt{\frac{D_p}{\tau_p}}\ p_{no}$$

and $$J_{Hp} = q \sqrt{\frac{2 D_n D_p}{(D_n + D_p)(\tau_n + \tau_p)}}\ n_i$$

Since low level conditions always prevail in the p^+ region the electron current is given by the electron component of

equation (2.15). The hole current of equation (2.21), the electron current from equation (2.15) and the recombination current of equation (2.20) can therefore be combined as follows:

$$J = J_{Ln}\left[\exp\left(\frac{qV}{kT}\right) - 1\right] + J_{rgo}(\psi_D - V)^{-\frac{1}{2}} \operatorname{Sinh}\left(\frac{qV}{2kT}\right) + \frac{J_{Lp}\, J_{Hp}\left[\exp\frac{3qV}{2kT} - \exp\frac{qV}{2kT}\right]}{J_{Lp}\left[\exp\left(\frac{qV}{kT}\right) - 1\right] + J_{Hp}\exp\frac{qV}{2kT}} \quad \ldots\ldots (2.22)$$

where J_{Lp} and J_{Hp} are defined above,

$$J_{Ln} = q\sqrt{\frac{D_n}{\tau_n}}\, n_{po}$$

and $$J_{rgo} = \frac{\pi n_i kT}{\sqrt{\tau_n \tau_p}}\left(\frac{2\varepsilon}{q}\right)^{\frac{1}{2}}\left(\frac{1}{N_A} + \frac{1}{N_D}\right)^{\frac{1}{2}}$$

This gives the total diode current for any injection level except very high level conditions.

It is important to realize that in the derivation of equations (2.15), (2.17), (2.18) and (2.20) it was not necessary to assume the electric field zero in the neutral regions. It was sufficient to assume it constant as implied by the term "neutral regions" and that the voltage V, applied to the space charge region only. In order to relate the voltage across the space charge region to the applied voltage V_a, it is necessary to add the integral of electric field over the neutral regions to the space charge voltage. This can be done simply by assuming

an effective value of resistance R for the neutral regions so that if the diode area is A:

$$Va - V = JRA \qquad \ldots\ldots (2.23)$$

where J is the current given by equation (2.22). If the neutral regions are long in comparison to the diffusion length of minority carriers, this approximation is acceptable.

In Figures 2 and 3 the effect of electric field in the neutral regions was not included for the sake of simplicity. However, in connection with Figure 3(b) it can be seen that in medium and high level injection strict space charge neutrality cannot be maintained in the injection region. If a field exists in the neutral region the departure from neutrality is decreased. This further emphasizes the necessity to check the validity of assumption (5) for each case considered.

Figure 4 demonstrates how equations (2.22) and (2.23) combine the currents and current limitations of a p^+n junction to give an expression for current at all levels of injection. To establish the validity of such an expression, however, the accuracy of equation (2.21) must be examined. To do this, equations (2.1) - (2.5) were solved numerically in the medium level injection region using assumptions (1) - (6). The values obtained were compared to those given by equation (2.21) for the same level of injection. Equation (2.21) underestimates the current throughout the medium level region by the amounts shown in Figure 5. The maxium error is about 20% but over the entire region the average error is less than 10%. In most practical

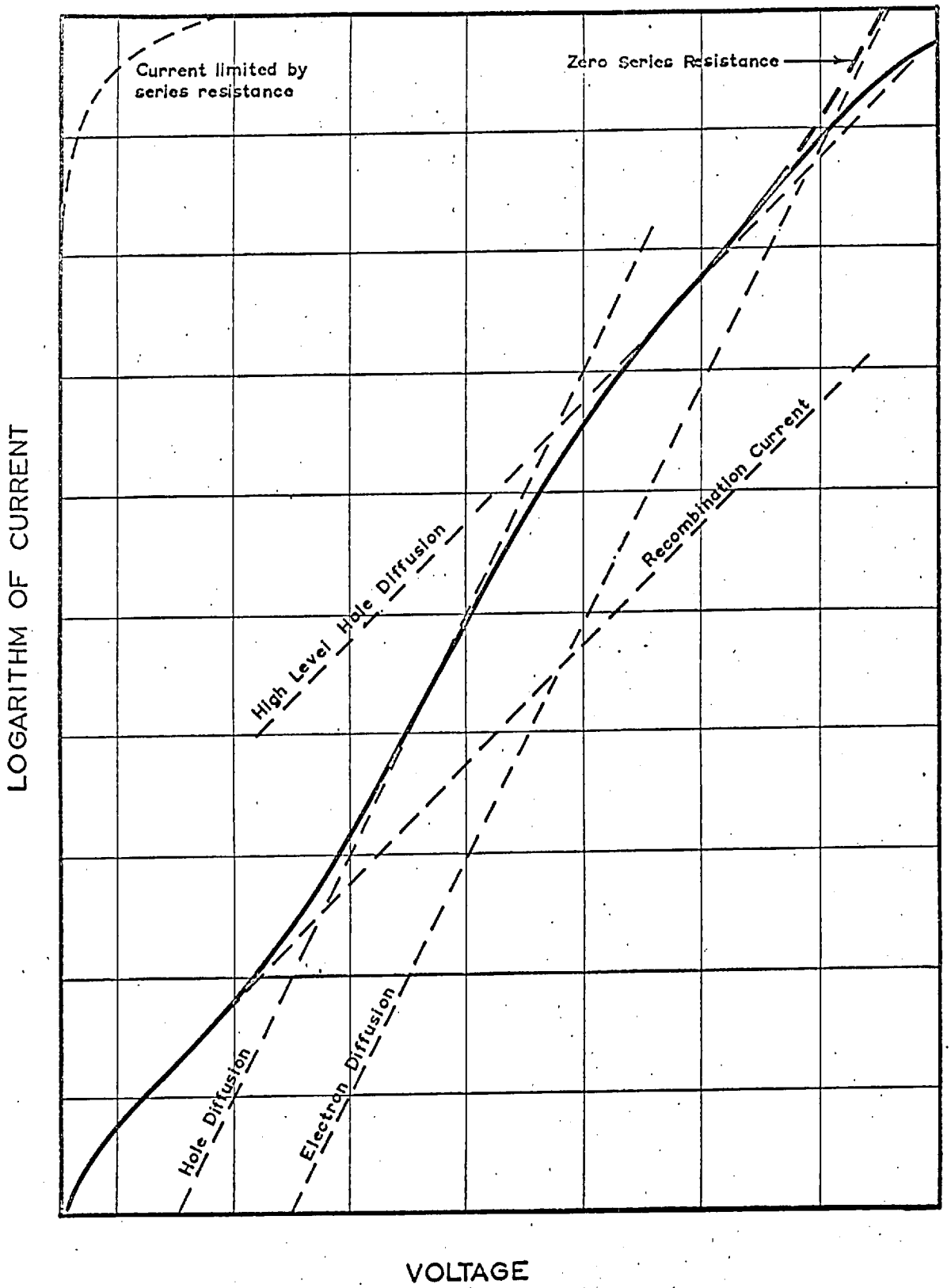


Figure 4. The current components of a p^+n junction and their combination.

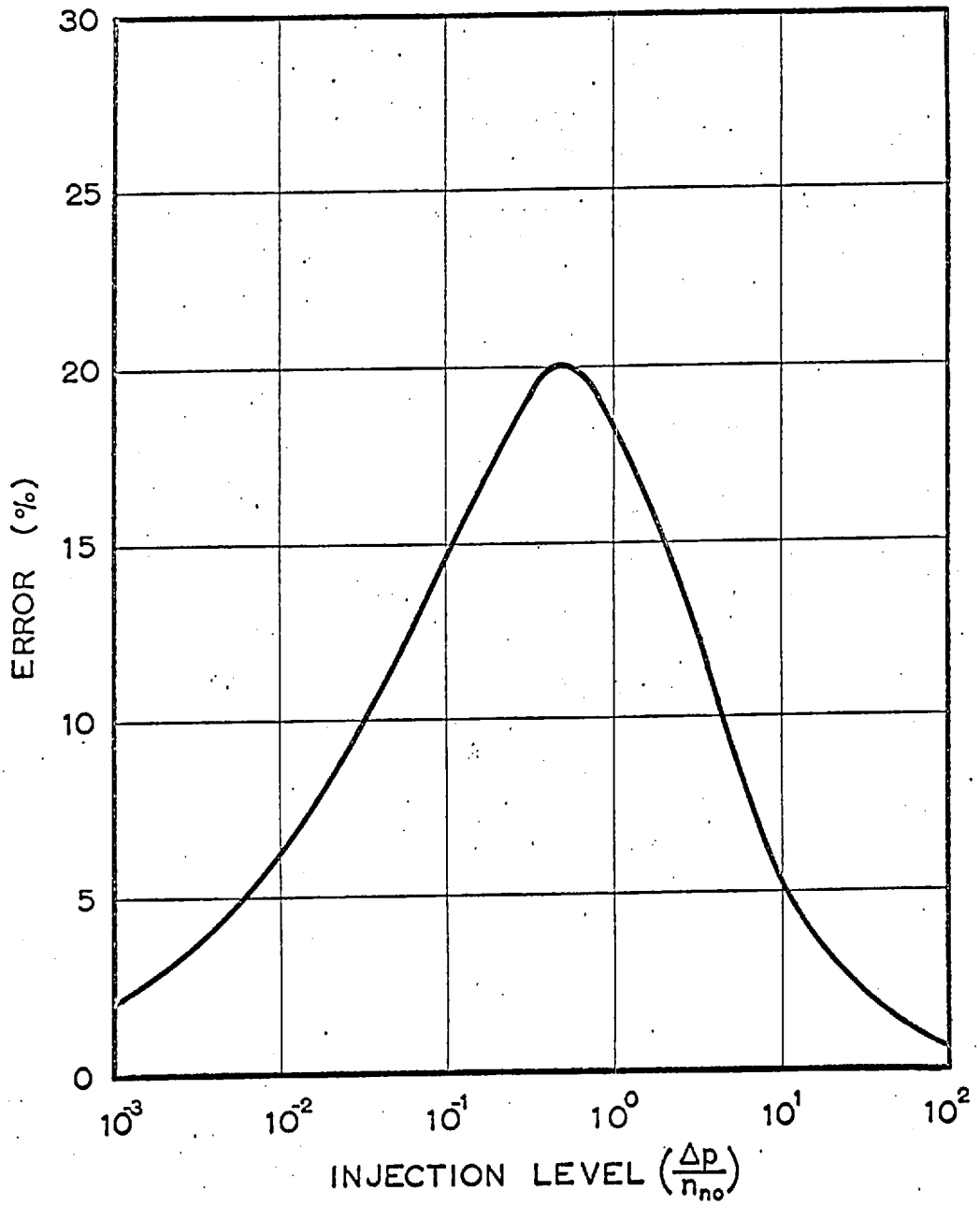


Figure 5. The error in the expression for medium level currents (equation (2.21)) as a function of injection level.

applications the advantages of an analytic expression for medium level injection is of greater importance than errors of this magnitude. Where it is possible to justify a more accurate interpretation of experimental results the graph of Figure 5 could be used to correct equation (2.21).

2.1.6. Reverse Currents.

When a reverse bias is applied to a p-n junction, currents flow due to electron and hole diffusion and to generation in the depleted space charge region. The independence of these currents implies that they add to give the total reverse current. Provided the voltage is not sufficient to cause multiplication or tunneling, equation (2.15) is also valid for negative voltages and gives the reverse diffusion currents. All the assumptions used to derive equation (2.15) do not remain valid, however, for assumption (6) is not correct at large reverse bias. However, in reverse bias ψ_n and ψ_p need not be constant through the space charge region for equation (2.15) to apply. It is sufficient if everywhere in the space charge region $|\psi_n - \psi_p| >$ several kT, since this implies that the region is essentially depleted of carriers. This modification to assumption (6) is also used [158] to integrate the generation rate over the space charge region. The resulting current is:

$$J_{rg} = \frac{qn_i W}{2\sqrt{\tau_n \tau_p}} \left[\cosh \left(\frac{E_t - E_i}{kT} + \tfrac{1}{2} \ln \frac{\tau_p}{\tau_n} \right) \right]^{-1} \qquad \ldots\ldots (2.24)$$

$$\text{if} \quad \left| \frac{E_t - E_i}{kT} + \tfrac{1}{2} \ln \frac{\tau_p}{\tau_n} \right| < \left| \frac{qV}{2kT} \right|$$

where E_t is the trap energy level, E_i the intrinsic Fermi level and W the width of the space charge region given by equation(2.19) The expression for total reverse current can be obtained by adding equations (2.15) and (2.24).

Assuming $|V| > 2kT$,

$$J_{rev} = J_{Ln} + J_{Lp} + \frac{J_{rgo} C}{2kT} (\psi_D - V)^{\frac{1}{2}} \quad \ldots\ (2.25)$$

where J_{Ln}, J_{Lp} and J_{rgo} are defined in section 2.1.5. and

$$C = \left[\cosh \left(\frac{E_t - E_i}{kT} + \tfrac{1}{2} \ln \frac{\tau_p}{\tau_n} \right) \right]^{-1}$$

For temperatures where both the n and the p regions are extrinsic, the temperature dependence of both J_{Ln} and J_{Lp} is largely that of n_i^2. The generation term has a temperature dependence determined largely by that of n_i and C. Therefore at low temperatures J_{Ln} and J_{Lp} become negligible in comparison to the generation term, equation (2.24). Substituting for n_i in equation (2.24),

$$J_{rg} = AT^{3/2} \exp\left(\frac{E_g}{2kT}\right) \left[\cosh \left(\frac{E_t - E_i}{kT} + \tfrac{1}{2} \ln \frac{\tau_p}{\tau_n} \right) \right]^{-1} \quad \ldots\ (2.26)$$

where $A = \frac{qW}{\sqrt{\tau_n \tau_p}} \left(\frac{2\pi k}{h^2}\right)^{3/2} (m_e m_h)^{\frac{3}{4}}$ and E_g is the energy gap.

By differentiating equation (2.26) two possible values for the trap energy can be obtained.

$$k \frac{d \ln(J_{rg}/T^{3/2})}{d(1/T)} = E_A = \frac{E_{go}}{2} \pm (E_{to}-E_{io})C' \qquad \ldots\ldots (2.27)$$

where $C' = \tanh \left| \frac{E_{to}-E_{io}}{kT} + \frac{1}{2} \ln \frac{\tau_p}{\tau_n} \right|$ and the subscript "o" indicates the extrapolated $0^{\circ}K$ value.

Thus reverse current measurements at low temperatures can yield two possible values of trap energy. The ambiguity arises since either the edge of the conduction band or valence band can be used as reference. A separate independent measurement would be required to eliminate this ambiguity.

2.1.7. Reverse breakdown.

There are two breakdown mechanisms which can cause a voltage saturated reverse characteristic in p-n junctions, Zener breakdown and avalanche breakdown. Zener breakdown occurs when the electric field in the space charge region is sufficiently high to cause direct excitation of electrons from the valence band to the conduction band. Avalanche breakdown occurs when carriers can gain enough kinetic energy from the electric field to cause ionizing collisions. Since the hole and electron produced by such a collision each have a finite probability of making a further ionizing collision this type of breakdown is often referred to as avalanche multiplication.

Avalanche breakdown in p-n junctions was first postulated by McKay and McAfee[181]. The theory has since received considerable attention[78][171][182] but it is difficult to

apply to breakdown measurements. The difficulty is apparent from the relationship for the multiplication factor M,

$$M = \frac{J}{J_{rev}} = \frac{1}{1 - \int_d^W \alpha_i dx} \qquad \ldots\ldots (2.28)$$

where J is the total reverse current, J_{rev} is the value of reverse current given by equation (2.25), W is the width of the space charge region and α_i is the ionization rate which is a function of electric field and is different for electrons and holes. In order to circumvent the complexity of the integral in the denominator of equation (2.28) it has been found[183] that the multiplication factor obeys the following empirical relation.

$$M = \frac{J}{J_{rev}} = \frac{1}{1-(V/V_B)^m} \qquad \ldots\ldots (2.29)$$

where V is applied voltage, V_B is the breakdown voltage defined as $V = V_B$ when $M = \infty$ and m is an experimentally determined coefficient.

The dependence of the avalanche breakdown voltage on carrier concentration is determined by two opposing factors. Firstly, the breakdown is dependent on the maximum field in the junction which is related inversely to the width of the space charge region given as a fuction of carrier concentration by equation (2.19). Secondly, it is dependent on the number of mean free paths in which a carrier could make ionizing collisions and thus directly on the width of the space charge region.

Combining these two factors for an abrupt junction gives

$$(\psi_D - V_B) \propto \left(\frac{1}{N_D} + \frac{1}{N_A}\right)^r \quad \ldots . \quad (2.30)$$

where r can take on values between one, for complete electric field dominance and minus one, for complete dominance by the width of the space charge region. Since breakdown cannot be independent of electric field and the field dependence of the ionization rate suggests electric field to be the dominant factor, the value of r is likely to lie between one and zero for any real junction.

The theory of Zener breakdown was first given by Zener[184] and later extended to semiconductors[185]. Zener breakdown occurs at much higher fields than that required for avalanche multiplication and as a result it is only observed in highly doped junctions where the space charge region is too narrow for avalanche multiplication to occur. An intermediate range of doping levels must exist, of course, in which both Zener and avalanche processes contribute to the breakdown. Since Zener breakdown does not depend on a path length the value of r in equation (2.30) might be expected to change from the value it assumes for avalanche, to r = 1 for Zener breakdown. In very highly doped junctions, however, a breakdown voltage is difficult to define and breakdown involves not only the electric field but a detailed balance between energy states in the conduction and valence bands in the high field region. The total breakdown voltage $(\psi_D - V_B)$ probably tends to a constant, E_g/q, in such cases.

2.2. Transient Response of p-n Junctions.

2.2.1. Capacity.

Capacity measurements on p-n junctions as a function of applied potential can be extremely useful since they can be used to determine the nature of the impurity distribution in the junction region and to detect the presence of surface inversion layers. Since the capacity of p-n junctions is treated in almost every text on semiconductor devices (see for instance[163][170][186]) the particular law pertinent to this thesis will simply be stated here.

For an abrupt p-n junction as described by assumption (4) if assumption (5) is valid, the unit area depletion capacity is:

$$C = \left[\frac{q\epsilon \ N_D N_A}{2(N_D+N_A)}\right]^{\frac{1}{2}} (\psi_D - V)^{-\frac{1}{2}} \qquad \ldots\ (2.31)$$

2.2.2. Step Response and Recombination.

The response of a forward biased p-n junction to a reverse voltage step can be used to measure the lifetime of excess carriers near the junction. The theory for the response of a junction to such a step has been considered by many authors. A particularly useful treatement has been given by Lax and Neustadter[187] for conditions of low level injection. The system used for transient measurements in this work has in general permitted measurements only in the medium to high level injection region and therefore reconsideration of the theory is necessary.

If a p^+n junction is considered and the following assumptions made:

(a) Space charge neutrality exists outside the space charge region,

(b) Recombination is determined by a single level of recombination centres,

(c) That the diffusivity can be considered a constant given by the ambipolar diffusivity for a particular injection level,

then the continuity equation reduces to the following (see Appendix C)

$$\frac{dp}{dt} = \frac{p}{\tau^*} + D^* \frac{d^2p}{dx^2} \qquad \ldots\ldots (2.32)$$

where

$$\tau^* = \frac{\tau_n + \tau_p(K+1)}{K+1}$$

$$D^* = \frac{D_p(K+2)}{K+1}$$

and $K = \frac{\Delta p}{n_{no}}$ the injection level ratio.

This equation is subject to the same treatment as that used by Lax and Neustadter[187] and gives the following expression for the current dependence of the recovery phase, that is the period in which the junction remains forward biased after the reverse bias step is applied.

$$\text{erf}\left(\frac{t_r}{\tau^*}\right)^{\frac{1}{2}} = \frac{J_f}{J_r+J_f} \qquad \ldots\ldots (2.33)$$

where t_r is the length of the recovery period and J_r and J_f are the reverse and forward currents immediately after and before the step.

It is obvious that D^* and τ^* are both functions of carrier density and therefore any solution of equation (2.32) which assumes these constant must be considered only as a first approximation. However, D^* and τ^* are essentially constant for $K > 1$ or $K < 1$ and hence the approximation is only invalid for carrier densities where $K \simeq 1$. If it can be shown that the variation of τ^* and D^* is small between the two constant regions then the solution can be considered accurate in the intermediate region.

In the reverse recovery phase τ^* can be determined independently of D^* and for this reason only the reverse recovery phase is considered. The value of τ^* obtained from equation (2.33) must, of course, represent a mean carrier lifetime which corresponds to a carrier density somewhat less than that indicated by the inital injection level. If, however, assumption (b) is accurate for high level and $\tau_n < \tau_p$ then $\tau^* \simeq \tau_p$ and the value of τ^* obtained from equation (2.33) is correct. The nature of τ^* must therefore be assessed for each particular case.

C H A P T E R 3

Experimental Apparatus and Procedures.

The techniques used to make alloyed p-n junctions in InSb and to measure the properties of such junctions will be described in this chapter. Only those techniques which required special development or modification will be discussed in any detail. The methods used for electrical measurements are, of course, applicable to any p-n junction. The description of alloying procedures applies to indium-InSb junctions although it has been shown that the alloying technique which resulted from this work can be successfully applied to other systems such as In-Ge (section 4.2.2.) and Sn-InSb (section 5.7.)

3.1. Crystal Preparation and Evaluation.

Since crystallographic direction is an important factor in alloying processes, all crystals were oriented before slicing. The orientation was determined using back reflection X-ray photographs. The crystals were mounted by either wax or shellac on a goniometer which could be transferred from a known orientation on the X-ray apparatus to a known position on the cutting machine. Slices aligned to within 1^{o} of a desired orientation were easily obtained by this method.

Because InSb is so subject to damage it was necessary to exercise great care both in cutting and polishing. Slices were

cut by means of a high speed carborundum impregnated cutting wheel, lubricated with water and driven at slow feed rates (i.e. 2 cm/hr.) The use of water soluble lubricants was found undesirable because deposits formed on the cutting wheel which caused wedging and breakage. For polishing, the slices were mounted on a piston type polishing jig which maintained the correct crystal alignment and a constant small pressure on the sample. Rough polishing was done on glass plates using a slurry of alumina or carborundum powder. The final polishing was done on $\frac{1}{4}$ micron diamond laps. The slice, with a final thickness of about 0.2mm, was cut into 1.5mm x 1.5mm dice and a 2mm x $\simeq$ 10mm Hall specimen using an air abrasive cutting tool. This method of cutting was very fast and introduced a minimum of damage. Handling was done wherever possible by means of vacuum tweezers, again, to minimize the possibility of crystal damage.

In general, part of a slice from each crystal was etched briefly in CP4A (5;5;3, HNO_3: CH_3COOH: HF) to permit an etch pit count and hence an estimation of the dislocation density. In diode fabrication each die was also etched briefly to clean it and to ensure that no damage was present.

To determine the carrier concentrations and mobilities standard Hall and conductivity measurements were performed over the temperature range from 300°K to 77°K on a sample from each slice of material used.

Spherical pellets of tin or of cadmium-doped indium, ranging in diameter from 40 to 500 microns, were used for alloying.

They were formed by a shot tower technique described in Appendix D.

3.2. Development of Alloying Techniques.

3.2.1. Methods of Junction Evaluation.

During wetting and alloying experiments it is necessary to have methods of evaluating the degree and nature of wetting or alloying achieved. The following methods were employed in these experiments:

(a) The wetted or alloyed area on the surface of the crystal was microscopically examined for any irregularities or for any dependence on orientation, damage or experimental conditions.

(b) The indium pellet was removed in warm, dilute HCl and the regrowth surface examined for any unwetted areas or imperfections.

(c) Junctions were potted in an epoxy resin (Araldite MY750 and hardener HY951), sectioned and etched two to three seconds in CP4A to reveal the shape and perfection of the regrowth and the extent of alloy penetration. Since InSb is normally intrinsic at room temperatures and is easily damaged the following procedure was necessary to obtain useful sections. The initial sectioning was done on a water-lubricated pregrinder using abrasives ranging from 120 to 600 mesh. Once the specimen was reached it was important that further grinding be done on abrasives not coarser than 600 mesh with medium pressure and steady movement to avoid cracking or deep scratches. After the required position had been reached it was also important to completely remove the damaged

surface layer resulting from the sectioning by polishing on graded diamond laps working down to a final polish on $\frac{1}{4}$ micron diamond. Since the etch was very brief and junction delineation was critical under ideal conditions, it was necessary to ensure that the sample surface was free from any contamination before etching. If this precaution was not observed mottled surfaces were obtained which made assessment of the regrowth difficult. In order to prevent oxidation the specimen was covered in etch for the required time and then the etch was rapidly diluted. At no time before complete dilution was the surface exposed to the atmosphere.

(d) Junctions were mounted and etched as described in section 3.3. and the electrical characteristics were measured as described in section 3.4.

3.2.2. Alloying Difficulties in InSb.

The preparation of alloyed p-n junctions dates back to 1953 when Saby and Dunlap[188] studied the capacitance of contacts prepared by fusing metals such as indium or aluminium to germanium wafers. The techniques of alloying to form p-n junctions have since become commonplace and the physical processes involved are relatively well understood. However, each new semiconductor produced brings with it new problems in the techniques of contact and junction formation. The difficulties arising from attempts to use established alloying techniques on InSb are described in this section.

Alloyed junctions are prepared by heating and subsequent cooling of a wafer of semiconductor in contact with a suitable

impurity metal or alloy containing a doping impurity. (For InSb, indium is particularly suitable and the required doping can be achieved by adding a few per cent of zinc or cadmium or selenium, etc.) As the temperature is raised the alloy first melts, then wets the crystal and as the temperature is raised further the molten alloy dissolves increasing amounts of the semiconductor wafer. When the crystal cools the melt supercools and simple freezing begins causing the dissolved crystal to precipitate out of the melt. If cooling occurs slowly the solid-liquid interface provided by the crystal will nucleate the precipitation and the crystal will grow back epitaxially, trapping some of the impurity atoms and forming the required junction.

Since alloying only occurs if and where wetting occurs, the importance of achieving uniform wetting cannot be over-emphasized. However, it is often necessary, particularly in transistor structures, to make alloyed junctions with shallow or at least accurately controlled penetrations. In such cases it is important that uniform wetting take place at temperatures where the crystal solubility is low and that the spreading which occurs during wetting and subsequent alloying be controlled or reproducible. In Ge and Si, which have relatively high melting points, it is possible to obtain good wetting at temperatures where the solubility of either Ge or Si is low provided both the alloy and the crystal are carefully cleaned and alloying is carried out in a reducing atmosphere such as hydrogen. In InSb where desirable alloying temperatures are likely to lie between 200°C and 300°C this is not the case.

Early experiments, using standard kanthal strip heaters to alloy indium pellets to InSb in hydrogen, showed that wetting could not be achieved much below 300°C and then only if the heating time was short. At these temperatures wetted areas were small and since the solubility is high large penetrations were unavoidable. It was also observed that during slow alloying cycles, which are conducive to good junction formation, the indium pellets tended to oxidize even in purified hydrogen atmospheres and the wetting temperatures were accordingly higher. It was apparent that hydrogen was not an effective reducing agent at these temperatures and neither variations in crystal or pellet cleaning procedures seemed to have any effect.

In an attempt to obtain uniform wetting at lower temperatures a number of fluxes were tried. Among these were resin; combinations of $ZnCl_2$, $InCl_3$ and NH_4Cl; $AlCl_3$; a commercial flux paste and a number of the fatty acids. Most of these fluxes seemed to improve the wetting in varying degrees, but with the exception of resin all had undesirable effects either on the wetting and alloying or on the electrical properties of the junction. Even the use of resin was restricted to a narrow temperature range where it promoted wetting but did not polymerize and results using resin were not entirely reproducible. Since in semiconductor systems, where minute amounts of impurity can have large effects, it is desirable to avoid the use of fluxes and since none of the fluxes tried seemed entirely satisfactory, the idea of flux-assisted wetting was rejected.

If the mass of indium is restricted it is possible to alloy at high temperatures and still achieve shallow, controlled penetrations. To obtain a limited mass of indium over a relatively large area, indium was evaporated on to InSb substrates in a vacuum of about 10^{-5} **Torr.** Alloying during evaporation or subsequent to evaporation both caused the indium film to form into globules and thus no satisfactory wetting or alloying was achieved.

Capillary alloying[189][190][191], a recent development, appeared to be the only remaining established technique which could possibly give uniform wetting at low temperatures. In this technique the semiconductor wafer is held immediately above a fine capillary. Molten indium at the desired temperature is forced up the capillary so that a fresh, pure surface of indium is continually exposed until it touches the wafer. A brief description of an extensive series of experiments using such a system is given in Appendix E. From these experiments it became clear that oxide on the indium had a serious effect on its wetting properties and it was possible that indium held sufficient oxygen in solution to cause poor wetting. It was therefore obvious that some new technique which would either cause reduction or solution of the oxide and yet not give rise to large penetrations was required.

Since alloying can be limited by the rate at which atoms are taken into solution, a fast alloying system was designed to see if a very quick, hot cycle would cause wetting before

significant penetration could occur. The system was capable of heating from the melting point of indium to 500°C and recooling to 150°C in less than 14 seconds. The wetting achieved was excellent but the times were not sufficiently short to prevent large penetrations. However, the result suggested that inversion of the heating arrangement might be more successful. The resultant radiation wetting system is described in the following section.

3.2.3. Radiation Wetting and Controlled Alloying.

The fast alloying cycle demonstrated that if high temperatures are reached even for short periods, uniform wetting can be achieved. However, it appeared that large penetrations could only be prevented if the crystal were kept cool while the pellet temperature was raised sufficiently to cause wetting. The system shown in Figure 6 was designed to do just this and in spite of its simplicity, its success was phenomonal. In this system a crystal supporting a pellet of indium is mounted on a polished heat sink immediately under a tungsten filament. The filament is heated to white heat for a period ranging from one to five seconds. The heating current is accurately set by means of a variable transformer and the heating time is controlled to better than 1% by an automatic timer. Under such radiation the pellet melts instantly and surface tension pulls it up into a perfect sphere. Momentarily the sphere becomes very shiny and then collapses on to the crystal and solidifies.

This technique generally produces uniform wetting over the entire area under the dot. This is demonstrated in Figs.7(a) & (b)

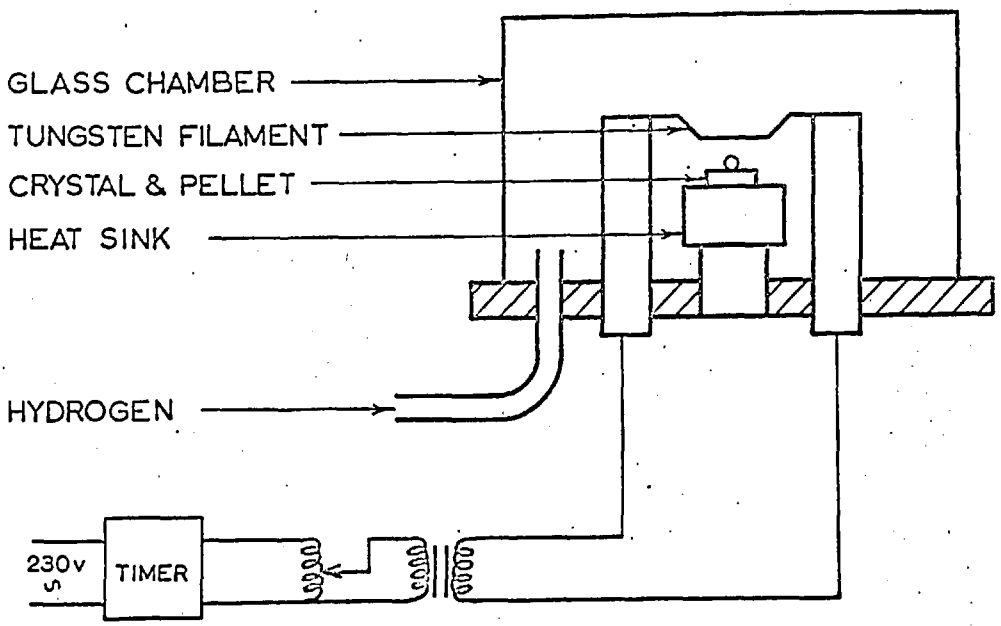
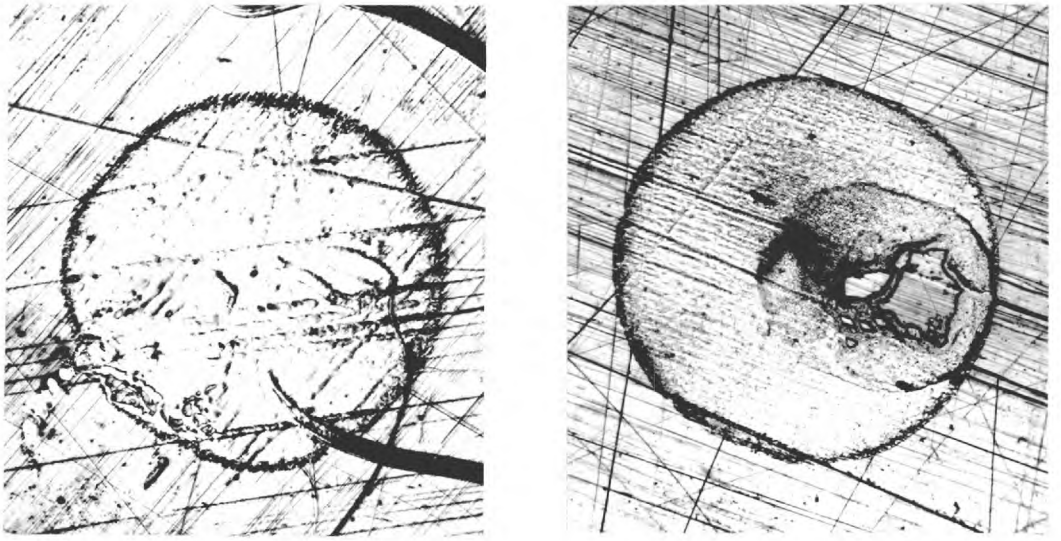


Figure 6. The radiation wetting system.

which show wetted areas after removal of the dot in HCl. Since surface scratches make comparison between the wetted region and the original surface easier, these crystals were not etched before wetting. Figure 7(a) shows that uniform wetting has occurred; however, it is not uncommon to find a spot or cluster of spots as shown in Figure 7(b) which have not wet during radiation. These spots have been related to the point of contact between the pellet and the crystal before wetting. When these spots occur they are, in general, small and thus are quickly and easily dissolved during subsequent alloying. Measurements of pellet penetration after radiation wetting have shown that it is limited to a few microns at most. This is demonstrated in Figures 7(a) and (b) by the fact that some of the small scratches (1 micron) and all of the larger scratches ($\simeq$ 3 microns) can be traced through the wetted area. It is clear, however, that, except for the spots already referred to in connection with Figure 7(b), none of the original crystal surface is visible indicating that uniform wetting has occurred.

Because this technique is so successful in obtaining low temperature wetting and since it can be applied with small variations to a great number of situations, it is important to understand the principles involved. During radiation, the pellet becomes hotter than the crystal because it is closer to the heater and because it makes poor thermal contact to the crystal. When it melts surface tension pulls it into a sphere reducing even further the small contact area on the crystal. At the resulting elevated pellet temperatures it is possible that some of the



(a) (b)

Figure 7. Wetted areas after removal of the indium pellets in dilute HCl (a) uniformly wetted (x150) (b) small unwetted areas remaining (x100).

Figure 8. Surface of regrowth after alloying at 250°C and removing the pellet in dilute HCl (x110).

surface oxides are taken into solution. However, more important is the fact that a hot tungsten filament in hydrogen generates considerable amounts of atomic hydrogen[193] which probably acts as an effective reducing agent at temperatures as low as 200°C. It is apparent, therefore, that two processes which promote wetting are proceeding simultaneously; the surface oxide is being removed and the pellet temperature is rising. Because these processes are proceeding quickly, the last point on the pellet where the critical wetting conditions are reached is at the point of contact between the pellet and the crystal. This gives rise to a highly unstable situation and thus when wetting starts the pellet literally collapses on to the crystal. The thermal contact to the crystal is then very good and the pellet temperature drops to the crystal temperature. It is now clear why a small unwetted area of crystal may be left at the original point of contact since not all of the contact area may have reached the critical wetting condition when the wetting action is nucleated.

During the course of other experiments it was found that the radiation wetting system of Figure 6 could be made more versatile by using a substrate heater of relatively large thermal mass in place of the heat sink used for InSb. Since wetting depends both on crystal and pellet temperatures it is possible with such a system to obtain optimum conditions of crystal temperature and radiation cycle for good wetting. The system can then be used effectively on higher melting point semiconductors such as Ge and GaAs.

Additional evidence was also found for the value of atomic

hydrogen as a low temperature reducing agent. Small spheres of InSb were placed with indium on a strip heater in a vacuum system. The system was evacuated, flushed with hydrogen several times and then filled with hydrogen at a pressure of about 5×10^{-2} Torr. The indium and InSb were heated considerably above the indium melting point but no wetting occurred and oxidation of the indium was observed as in previous experiments. The hydrogen was then ionized by a high electric field. In a very short time the indium became clean and shiny and wetting occurred.

Junctions formed by radiation wetting only, gave essentially ohmic electrical characteristics. This is not surprising since the times involved in the cycle are not sufficiently long to permit any significant amount of diffusion or regrowth and junctions are probably highly imperfect, as regions in Figure 7(b) would suggest. To form p-n junctions it was necessary, therefore, to submit the wetted junctions to a further alloying cycle. It is here that the great advantages of prewetting arise since in subsequent alloying no extreme precautions with regard to atmosphere are necessary and any temperature variation during the cycle can be specified. By further alloying of radiation-wetted junctions in a simple tunnel furnace using slow heating and cooling it has been possible to obtain good p-n junctions with alloying temperatures as low as 200°C. This substantiates the conclusion that uniform wetting and shallow penetration result from radiation wetting. The uniform regrowth achieved by slow alloying is demonstrated in Figure 8 by a regrowth surface exposed in a similar fashion to those of Figure 7. The results of further

metallurgical experiments using radiation wetting and controlled alloying are described in Chapter 4.

3.3. Diode Fabrication.

For the diodes described in Chapter 5 indium-1% cadmium pellets were radiation wetted to n-type InSb dice. The dice were then placed on thin wafers of pure tin on gold plated TO5 headers. The tin wafers had previously been coated with a very thin layer of resin to promote wetting during the alloying cycle. The whole assembly was subsequently transferred into a tunnel furnace and alloyed in air. The rate of heating and cooling was not allowed to exceed $5^{\circ}C$/min and the maximum temperature of the cycle was either $250^{\circ}C$ or $275^{\circ}C$.

After alloying the assembly was placed on a strip heater and heated to just above the indium melting point. A micro-manipulator was used to insert a 0.002" gold wire into the molten indium dot which was then allowed to freeze. The gold wire was soldered to the lead wire using indium 5% silver. A photograph of a completed diode is shown in Figure 9.

Before chemical treatment the tin and indium solders and part of the indium pellet were masked with 'Detel', chromium plating grade, masking paint. In order to minimize any possibility of contamination the etching and first stage washing were carried out in a jet etching system shown in Figure 10. The first stage of etching was done with 5-10cc of CP4A followed immediately by a rinse of about 150cc of 20MΩ -cm deionized water. Etching was then continued with 15-20cc of H100, an etch designed to

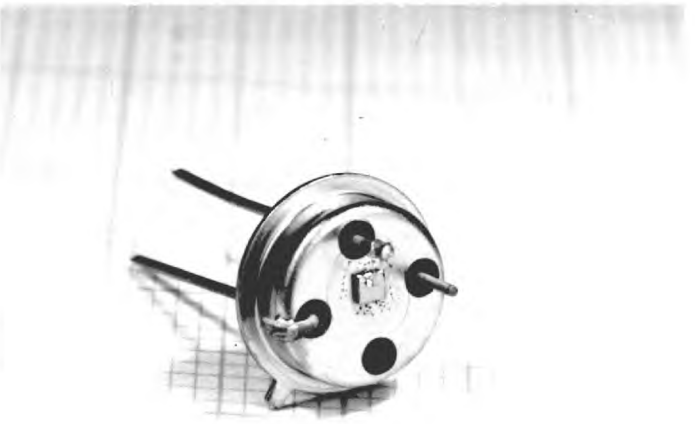


Figure 9. A completely assembled diode (x3).

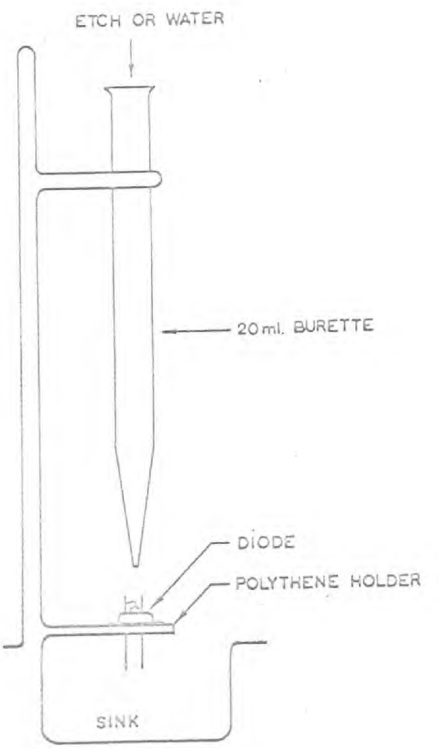


Figure 10. The arrangement for jet etching diodes.

leave the surface stoichiometric and free from any metallic contamination. This etch was followed immediately by a preliminary wash in about 200cc of deionized water. The diode was then transferred to a recirculating washing system and washed for about five minutes in freely flowing 20MΩ -cm deionized water. After washing, the excess water was shaken from the diode, it was dried in air at 100°C for five minutes and transferred immediately to the cryostat.

It was found that good results could be achieved only if the chemical treatment described above was carried out meticulously. Any contamination of the etches or carelessness in the washing or drying procedures could result in degraded electrical characteristics as described in section 5.5.1.

3.4. Electrical Measurements.

3.4.1. Cryostat.

In section 1.2.3. it was pointed out that, except for tunnel devices, InSb p-n junctions would not exhibit diode action at room temperature. The requirement that InSb junction devices must be cooled to liquid air temperatures is not only a serious handicap for possible devices but it presents technical difficulties, in the measurement of device properties. It would be ideal, since a cryostat is necessary, if it could be designed for one type of measurement only. However, it is important that all diode measurements be made under the same surface conditions and therefore it is not desirable to transfer a diode from one cryostat to another for different measurements.

In these experiments electrical measurements from D.C.

covering a band width of about 1Gc/s with temperatures ranging from $300^{o}K$ to $63^{o}K$ were envisaged. Accurate measurements of DC currents from $10^{-10}A$ to 0.5A, of DC voltages as low as one millivolt and of capacities as low as one picofarad were required. The cryostat designed to permit such measurements to be made reliably is shown in Figure 11.

To minimize any mismatch at high frequencies the electrical leads were constructed as 50 ohm lines and all connections to the diode holder were kept short. The response of the system is discussed in section 3.4.4. The outer conductors of the lead wires were specially reamed, thin walled, tubes of German silver to reduce the thermal conductivity. The leads were made as short as possible to reduce the undesirable electrical resistance but still keep the thermal resistance relatively high. The series resistance obtained was about 10 milliohms while the parallel leakage resistance was greater than 10^{12} ohms.

A third thin walled German silver tube with two external connections was used to control the atmosphere in the cryostat and to insert a fine copper-constantan thermocouple for temperature measurement. The cryostat chamber could be evacuated through this tube using a backing pump with a liquid nitrogen cold trap or filled with dry hydrogen. The thermocouple was attached to a brass washer and held firmly by a nylon screw to the copper diode holder which also acted as a cooling fin and heat sink. The thermocouple was electrically isolated from the holder by a thin film of P.T.F.E. When a diode was placed in the cryostat it was firmly clamped to the holder to provide good electrical and thermal contact.

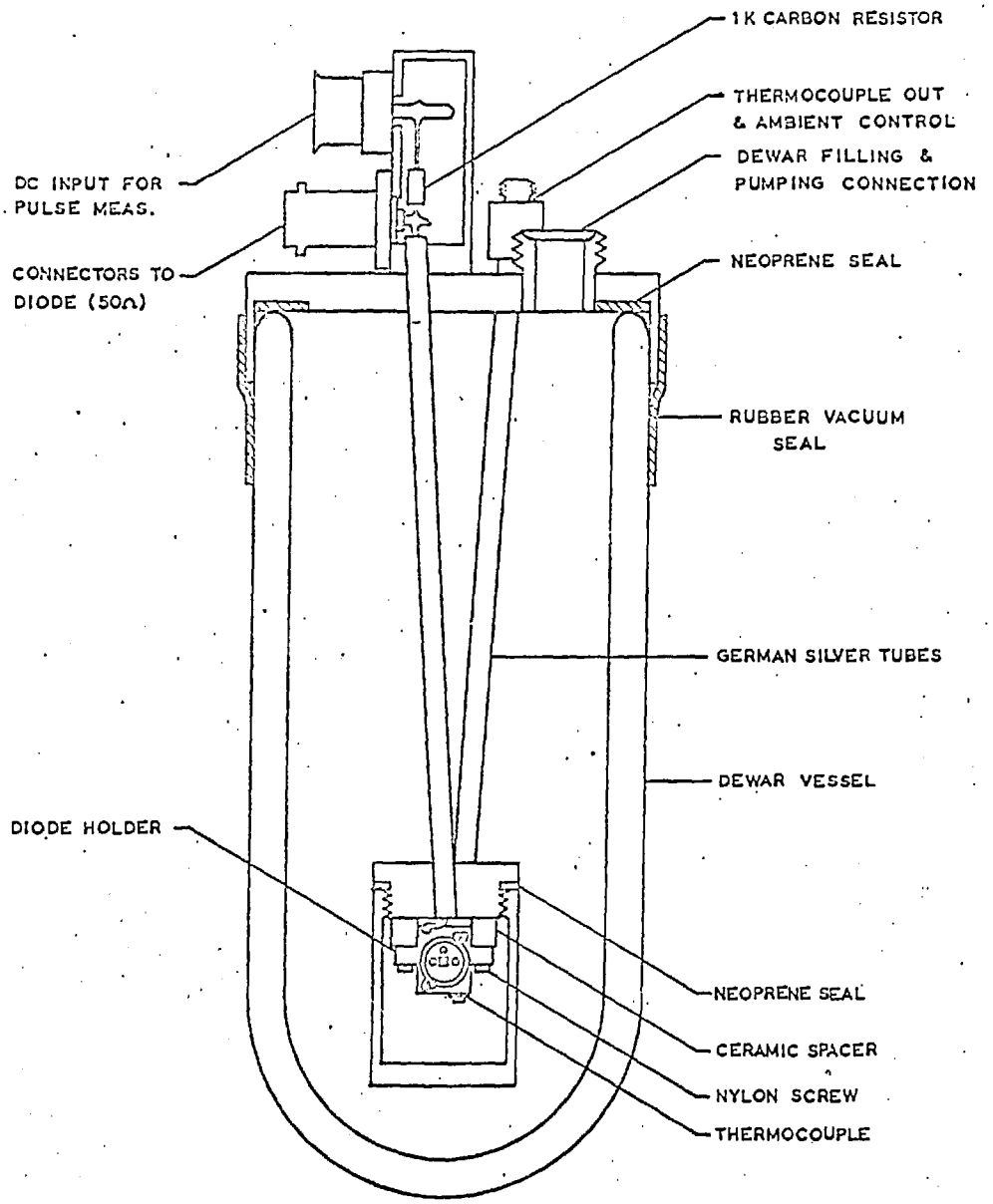


Figure 11. The cryostat.

The top flange of the cryostat was designed to fit tightly over a standard dewar vessel so that, when sealed with a rubber band, the dewar could be pumped. Solid nitrogen temperature ($63.3^{\circ}K$) was easily reached with this arrangement.

3.4.2. Direct Current Measurements.

Detailed measurements of the electrical characteristics of a diode can be very time consuming. It was necessary, therefore, to have a method by which a rapid initial assessment of the success of alloying and chemical etching could be made. Since properly etched InSb diodes give very low reverse currents it was necessary to design and build a high sensitivity curve tracer for this purpose. The curve tracer shown in Figure 12 utilized a doubly screened, low interwinding capacity transformer and a differential amplifier to reduce the effective capacity across the diode to less than 3 picofarads. The maximum useful sensitivity was therefore limited by the trace loop resulting from the diode capacity itself. The curve tracer shown gave a maximum sensitivity of 5×10^{-7}A/cm vertical and 50mv/cm horizontal. Typical traces are shown in Figure 13.

Detailed measurements of the DC characteristics were carried out using the circuit shown in Figure 14. A DC micro-voltmeter was used with calibrated input resistors for current measurements. The voltages across the current meter and diode were measured by a digital voltmeter with an input impedance greater than 10^9 ohms. Since current readings were made on a voltmeter, the voltage readings were easily corrected to give the diode voltage. With

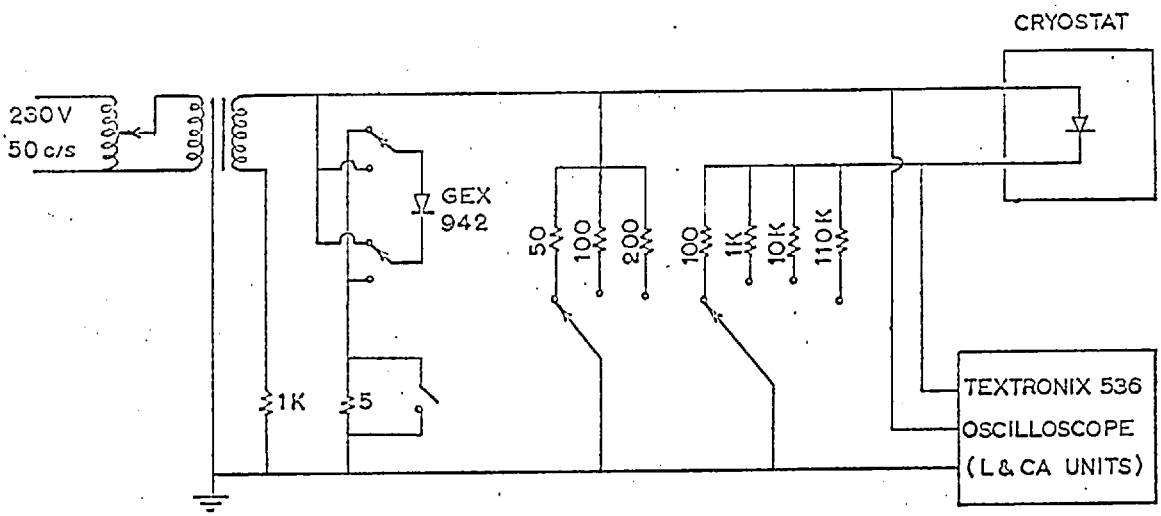


Figure 12. Circuit diagram of the high sensitivity curve tracer.

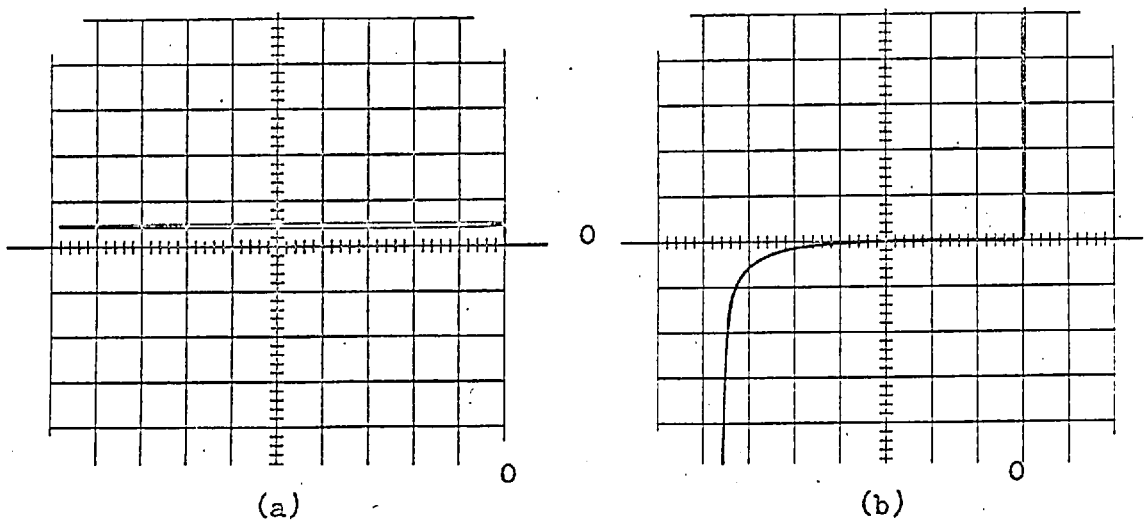


Figure 13. Oscilloscope traces obtained using the curve tracer of Figure 12 (a) open circuit trace (with zero shifted up) at 5 x 10^{-7}A/div vertical and 5 v/div horizontal (b) an InSb diode trace at 2 x 10^{-6}A/div vertical and 2 v/div horizontal.

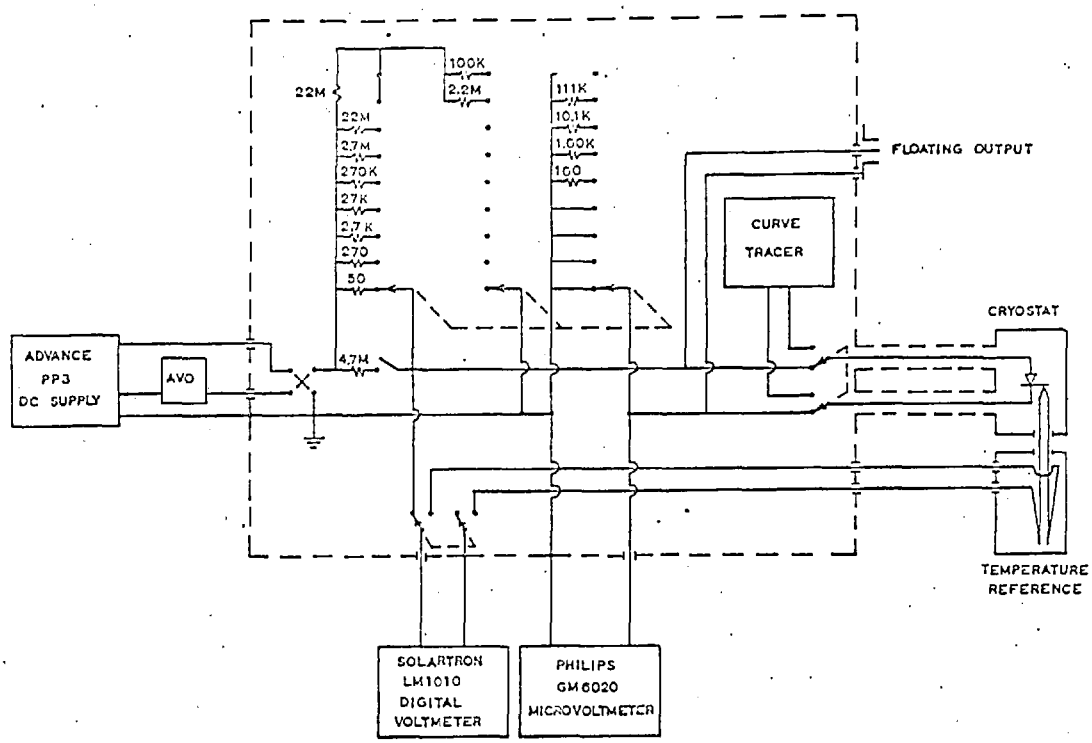


Figure 14. The circuit used for DC measurements and for DC biasing in capacity and pulse measurements.

all instruments placed on an aluminium ground plane and all instrument earths connected to this plane by short leads it was possible to read currdnts as low as 10^{-11}A and voltages as low as 50 microvolts. The errors in current measurements above 10^{-9}A were less than 5% and those in voltage for voltages above 20mv less than 0.5%.

The circuit for DC measurements and the curve tracer were arranged to permit the diode to be switched from one to the other. The overall characteristic could be easily rechecked at any stage in the measurement using this facility. The DC measuring circuit was also designed to provide a floating output for the DC biasing in capacity and pulse measurements.

To check the maximum dissipation which could be tolerated in a typical InSb diode a pulsed high current measuring system, similar to that described by Jonscher[194] was built. The forward characteristic of a diode was measured at high currents using this system and compared to that obtained by DC measurements. It was concluded that for a diode of $10^{-3}cm^2$ area formed on a wafer 2×10^{-2}cm thick no significant difference could be detected at $77^{\circ}K$ until 20mW was dissipated. Calculations have shown that a change in junction temperature of about $0.1^{\circ}C$ might be expected at 10mW dissipation. This would give rise to about 3% change in current and is therefore in good agreement with the observations.

3.4.3. Capacity Measurements.

The diode capacity measurements were made using a standard Wayne Kerr B601 transformer bridge. A signal generator was used

to supply a 500 kc/s voltage to the bridge. Since diode capacity is voltage dependent it was necessary to restrict the input voltage to less than 30 mv and when measuring around the zero bias point it was reduced to 5 mv. With such low input voltages a very sensitive detector was required to determine the null point of the bridge accurately. It was found that even with a narrow band-width receiver the sensitivity was too low and the noise figure much too high to obtain a sharp null. To overcome this difficulty the low noise preamplifier shown in Figure 15 was built. This preamplifier gave 50db gain at 500 kc/s and 11 db improvement in the signal to noise ratio. The noise figure of the pre-amplifier was 3 db.

A further problem was encountered in supplying DC bias to the diode and at the same time eliminating stray capacitance. By careful screening and appropriate earthing the bridge zero was made independent of the biasing and stable to better than 0.1 pf. The complete circuit shown in Figure 16 permitted measurement of capacities up to 22 pf accurate to $\pm$ 0.1 pf and above 22 pf accurate to 0.5 pf.

3.4.4. Pulse measurements.

When a diode which is forward biased is suddenly switched to a reverse biased condition the reverse current which flows initially arises largely from the need to extract the minority carriers injected into both n and p regions during forward bias. Since these carriers recombine with characteristic lifetimes the reverse current decays toward its saturation value in a manner

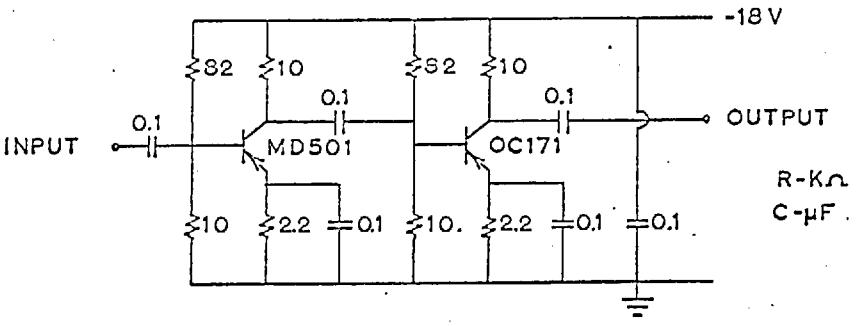


Figure 15. Low noise preamplifier.

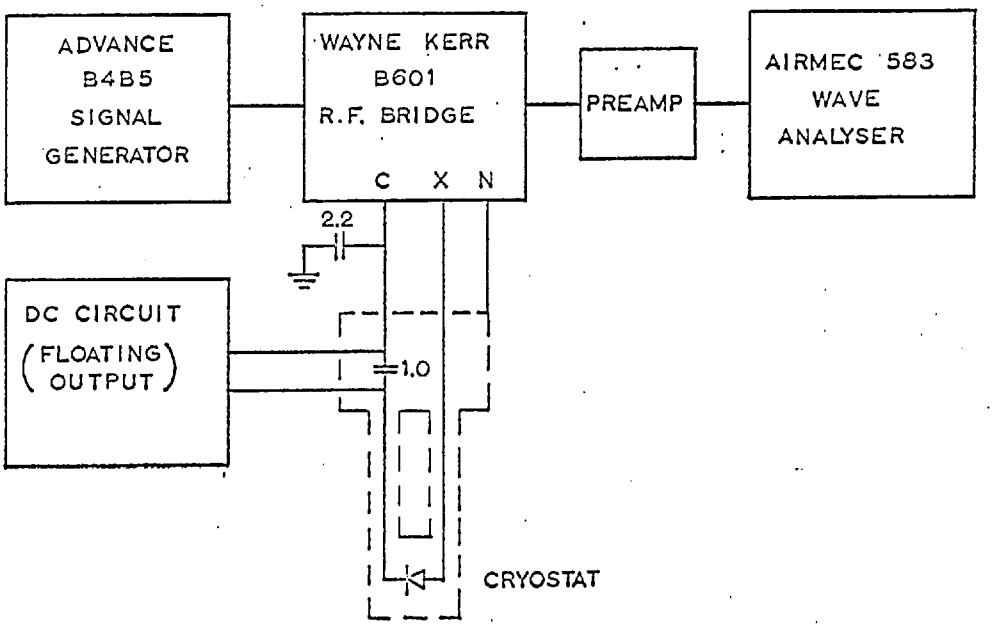


Figure 16. The complete system for diode capacity measurements.
(indicated capacities are in μF.)

determined by these lifetimes. There are two distinct phases in the response of a forward biased diode to a reverse biasing voltage step: One is the constant current or recovery phase, discussed in section 2.2.2. During this period the reverse current is limited by the series resistance of the circuit and for this reason the injected charge near the junction cannot be reduced immediately to zero. The diode continues, therefore, to have a forward bias voltage during this period. The other is the reverse phase. During this period the diode becomes reverse biased and the reverse current decays as the injected minority carriers in the junction continue to recombine. Either or both of these phases can be used to determine minority carrier lifetime.(187) However, if accurate measurements are to be made using the recovery phase, the risetime of the reverse pulse or step must be at least 100 times less than the carrier lifetime while simple interpretation of the reverse phase requires that the pulse be at least 10 times longer than the lifetime. For InSb p^+n diodes on good material a pulse with a risetime of about 10^{-9} sec and a length greater than 10^{-6} sec is required. Electronic pulse generators with the required risetime do not provide pulses much longer than 10^{-7} sec. Mercury relay pulse generators using a coaxial pulse-forming line can be used to generate pulses up to about 5×10^{-7} sec before either attenuation causes sag in the pulse or the volume of coaxial line becomes prohibitive. To obtain the desired pulse length with a short risetime the broad band capacitor shown in Figure 17 was made. The capacitor was tested and shown to give a standing wave ratio of less than 1.01 at 1.7 Gc/s. Using this capacitor on a mercury relay pulse generator and DC biasing it as shown in Figure 18, it was possible to generate pulses with a risetime of

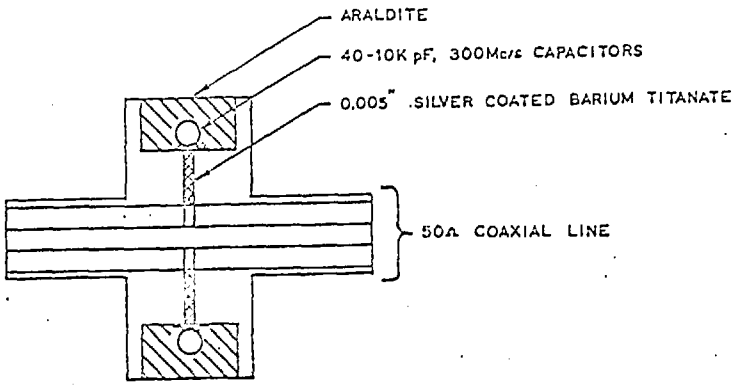


Figure 17. Sectional view of the high frequency, broad band capacitor.

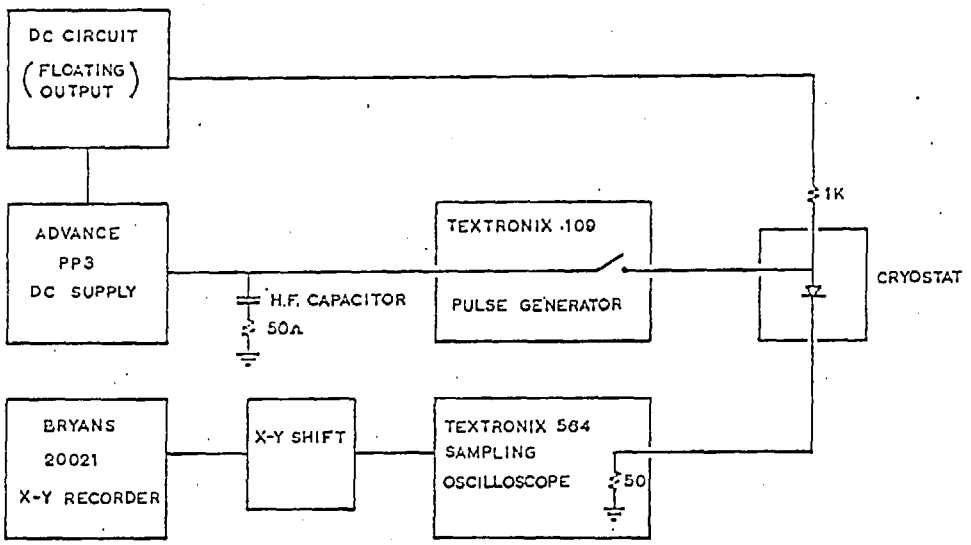


Figure 18. The complete system for diode reverse response measurements.

less than 2.5×10^{-10} sec and a length equal to the contact time of the mercury relay ($\simeq 10^{-3}$ sec). Pulses derived in this way were applied to the diodes and the responses were observed on a Textronix sampling and storage oscilloscope. This oscilloscope, which has a risetime of less than 4×10^{-10} sec, was used to drive an x-y recorder and thus accurate graphs of the response were available for analysis. The full system is shown schematically in Figure 18. A typical diode reverse response curve is shown in Figure 19. The rapid rise and small amplitude, short duration ringing at the beginning of the pulse indicate that the mismatch introduced by the diode and cryostat was small.

3.5. Area Measurements.

The electrical measurements described above can only be related to diode theory or to bulk properties if they are converted to densities, for example current density and capacity per unit area. To do this the diode junction areas must be known. In these experiments every junction was photomicrographed and the junction area was measured by planimeter on the photograph. An accurate calibration of the microscope magnification was obtained by photomicrographing a standard scale and subjecting this photograph to exactly the same processing as that used for the area photographs. The junction area thus obtained was corrected for the shape of the regrowth area, as outlined in Appendix F. After the first etch the forward current characteristic of each diode was carefully measured to establish an area reference for each subsequent etch. It was estimated that about 10 microns of crystal surface was removed in each etch. The diode area after the first etch was also corrected for this amount of surface removal.

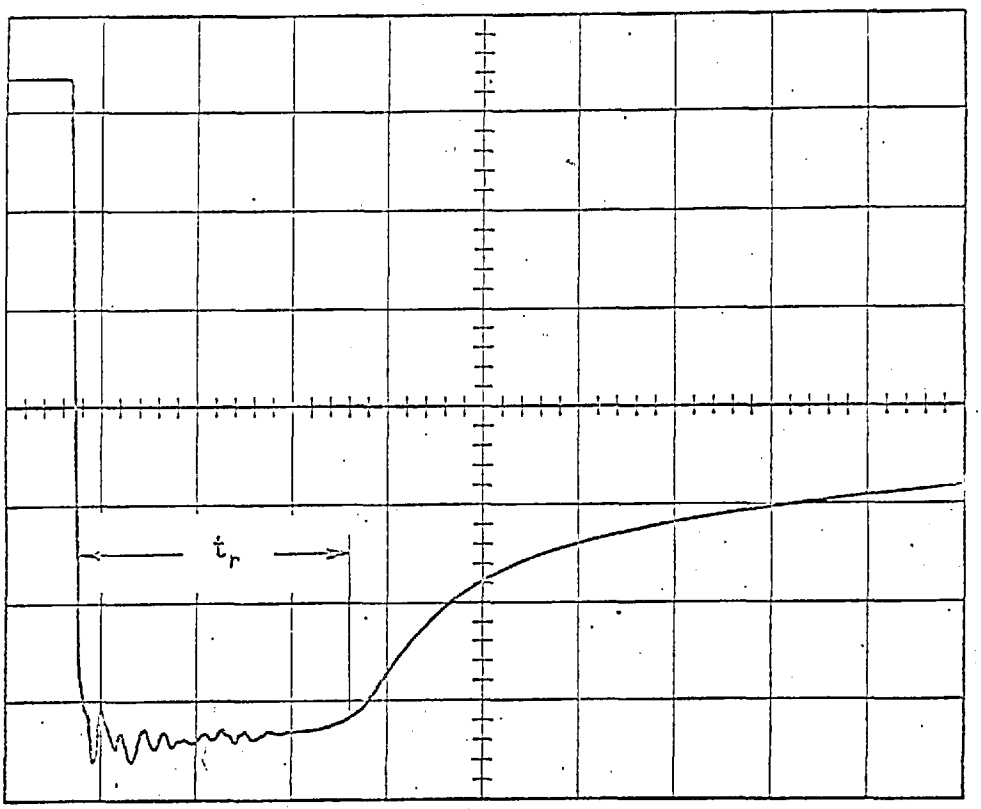


Figure 19. The typical response of a forward biased InSb diode to a reverse biasing step. The reverse recovery period t_r is indicated. (The diagram is reduced by one half from a typical x-y trace at 2 mA/div vertical and 10^{-8} sec/div horizontal.)

The measured diode currents were converted to densities using the areas determined in this manner and for as many as five diodes formed on the same material the resulting current densities differed by a maximum of about 10%. The agreement with theory discussed in Chapter 5 and the estimated errors in the area determination would suggest that these areas were also accurate to within 10%.

By converting all measurements to property densities a very large consolidation of results was achieved. Something like 250 sets of measurements on 80 diodes were reduced to 10 sets of graphs. The results discussed in Chapter 5 are the consolidated results where most graphs represent measurements on at least three diodes. The extremely good agreement among measurements on several diodes formed on the same material indicated that the techniques of diode fabrication and of diode measurement described in this chapter were very reliable.

CHAPTER 4

Crystallographic Effects on Alloying.

Part of the objective of this work as set out in section 1.1. was to examine the effects of polarity on both the metallurgical and electrical properties of p-n junctions. In section 1.2.2. it was pointed out that no polar effects have been observed in diffusion. It can further be stated that no polar effects on the electrical properties of diffused junctions have been reported. In this chapter the effects of polarity on the formation of alloyed p-n junctions will be described. The chapter will be devoted entirely to metallurgical observations and their interpretation. However, all four methods of junction evaluation (see section 3.2.1.) have been used and therefore electrical characteristics which contribute to an understanding of the metallurgical effects will also be included.

4.1. Alloying on A {111} and B {111} Surfaces.

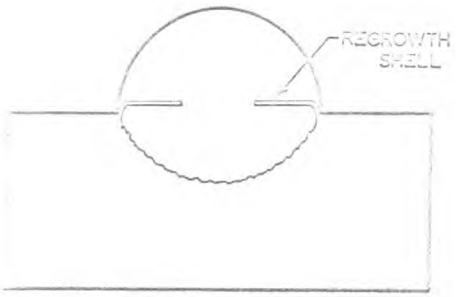
For the experiments on {111} surfaces, dice were cut as described in section 3.1. from InSb having donor concentrations less than $10^{15} cm^{-3}$ and dislocation densities less than $10^{4} cm^{-2}$. To determine the alloying conditions under which good junctions could be formed, pellets of indium 1% cadmium were alloyed on both A and B surfaces of a large number of dice under conditions where the rate of temperature rise and fall could be controlled at any value less than $70^{\circ}C/sec$. The very fast alloy cycles were carried out in hydrogen on a strip heater of low thermal capacity. The maximum temperatures used varied from $300^{\circ}C$ to $400^{\circ}C$

and the rate of temperature rise was controlled at values between 5^{o} and 70^{o}C/sec. For the slower alloying cycles the pellet was first wet on to the crystal using the radiation wetting technique described in section 3.2.3. Alloying was then carried out either on a strip heater or in a suitably programmed furnace at rates between 5^{o}C/min and 10^{o}C/sec.

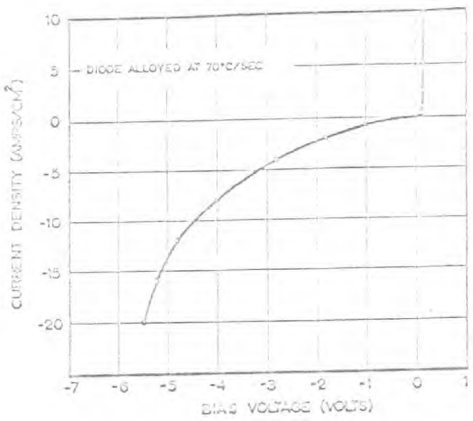
About 200 junctions were formed in the ways described. All were examined for any shape dependence of the surface wetted area, about 70 were sectioned, the pellets were etched from about 40 to permit examination of the regrowth surfaces and the remainder were tested electrically as described in section 3.4. For convenience the results will be described under three conditions of alloying.

(a) Highly nonequilibrium alloying, $(dT/dt) > 50^{o}$C/sec.

The regrowth regions produced under these conditions were round on both A and B surfaces as illustrated in Figure 20(a). The regrowth was limited and the regrowth surface was very uneven. In every case a single crystal shell was formed above the original crystal surface, Figure 20(a), and if this shell did not completely cover the regrowth region the open portions were, in general, bounded by edges identified as A {211} surfaces, Figure 20(b). (Note A [211] surfaces are those {211} surfaces nearest to A {111} surfaces). Since <211> directions favour dendritic growth[195] it is possible that these shells grew by a dendritic process. Surface wetted areas were always round and the electrical characteristics as exemplified by Figure 20(c) were uniformly bad for both surfaces. In view of the irregularity of the regrowth this latter observation is not surprising.



(a)



(c)

(b)

Figure 20. Junction prepared under highly nonequilibrium alloying conditions ($dT/dt > 50^{\circ}C/sec$) (a) Diagrammatic section illustrating the round junction and regrowth shell (b) A regrowth shell with triangular hole bounded by A{211} surfaces (x225) (c) The electrical characteristic at $77^{\circ}K$.

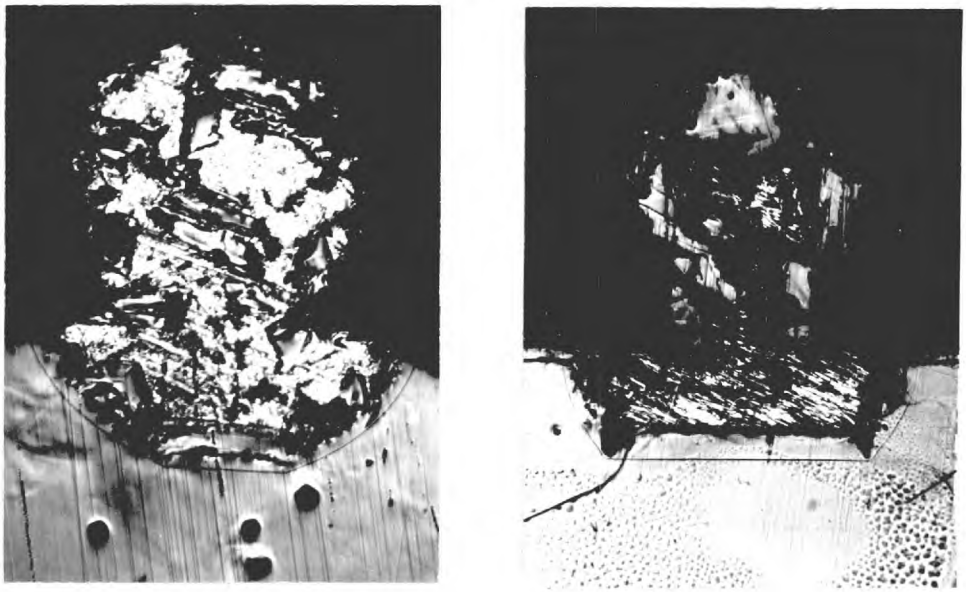
(b) Nonequilibrium alloying, $5^{o}C/sec < {}^{dT}/dt < 10^{o}C/sec$.

In this region experiments were carried out using either alloying preceeded by prewetting or alloying without prewetting. In the latter case wetting began only when the temperature reached approximately $300^{o}C$ and then only took place over a limited area. By the completion of the alloy cycle the pellet had invariably assumed a pear-like shape with the neck in contact with the crystal. The final wetted area in all cases was circular and was larger than the pellet contact area. The spreading of the wetted area was greater on B surfaces than on A surfaces. Junctions formed on A surfaces were, in general, round with a slight flattening at the bottom as shown in Figure 21(a), while those formed on B surfaces were flat with slight rounding near the edges as shown in Figure 21(b). The regrowth was poor on both surfaces but there were usually larger areas of good regrowth on the B surfaces. Electrical measurements were not made on any of these junctions but it was assumed that Minamoto's[55] observations of poor characteristics from A diodes and of good characteristics from B diodes would apply to this case.

When pellets were prewetted, junctions on both surfaces were flat at the centre and rounded at the edges as shown in Figure 22. However, in general, the wetted area and the flat area of the junction were larger on the B surface. Surface wetted areas on both surfaces tended to be round with slight flattening on three sides. The regrowth was poor on both surfaces. No electrical measurements were made on these junctions.

(c) Quasiequilibium alloying $(dT/dt) < 1^{o}C/sec$.

For experiments involving slow rates of temperature rise, all pellets were prewetted. A few diodes were alloyed at a rate of $1^{o}C/sec$.



(a) (b)

Figure 21. Sections of junctions prepared without prewetting under nonequilibrium alloying conditions ($dT/dt \sim 10^{\circ}C/sec$) (a) alloyed on the A surface (x150) (b) alloyed on the B surface (x100).

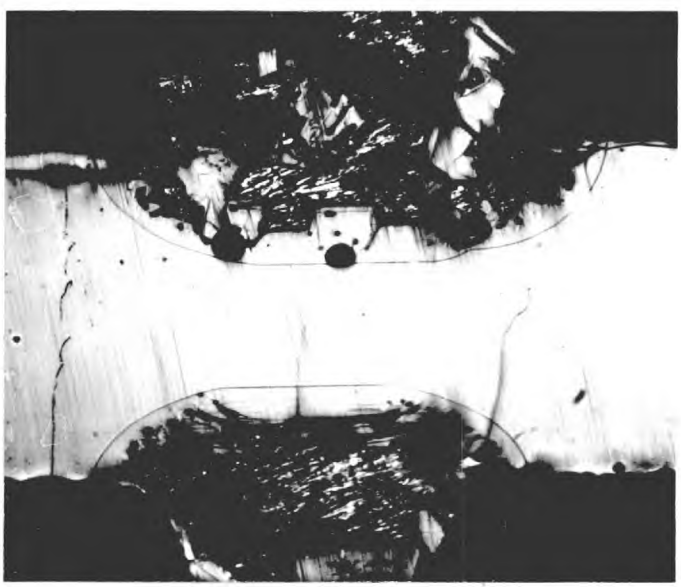


Figure 22. Section of a double junction (A uppermost) alloyed at $8^{\circ}C/sec$ after prewetting (x110).

None of these were examined electrically but sections showed flat junctions on both A and B surfaces. The spreading of the wetted area was considerably greater on B surfaces than on A and the wetted area often tended to assume an irregular hexagonal shape as shown in Fig.23. X-ray examination of several of these junctions showed that the well developed sides were associated with B {111} planes and it was thought that such areas could be used to indicate the relative ease of dissolution on facet forming planes. However, recent observations on crystal C396/u, where triangular wetted areas such as that shown in Figure 24 resulted from alloying at 5°C/min., have cast some doubts on the crystallographic significance of these latter observations. Both the shape of the wetted areas and the electrical characteristics (see Chapter 5) of junctions formed on crystal C396/u indicated it to be a special case, but the fact that some crystal defects which cannot be easily distinguished in ordinary Hall and conductivity measurements (possibly sulphur as an impurity in this case) can effect the metallurgical properties to such an extent, indicates that caution must be exercised in drawing conclusions about the relative importance of facet forming planes from observations of wetted areas.

Most of the junctions examined and all of those used for the electrical measurements described in Chapter 5 were prepared under conditions in which the rates of rise and fall of temperature during the alloying cycle were less than 5°C/min. Flat junctions were obtained on both A and B surfaces without exception. A double alloyed section is shown in Figure 25(a). Although some irregularities in the shape of the wetted areas occurred, over 50% were almost perfectly hexagonal as shown in Figure 25(b) and, with the exception of crystal C396/u

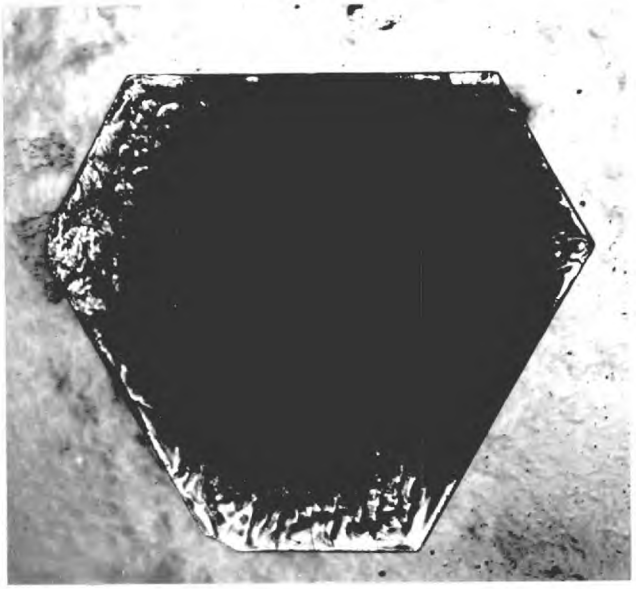


Figure 23. An irregular hexagonal wetted area observed after alloying at ~1°C/sec. The well developed sides are associated with the B side planes. (x100)

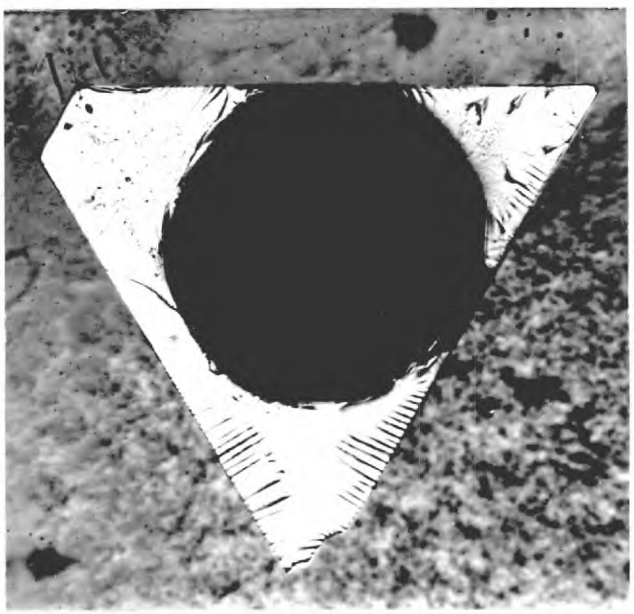
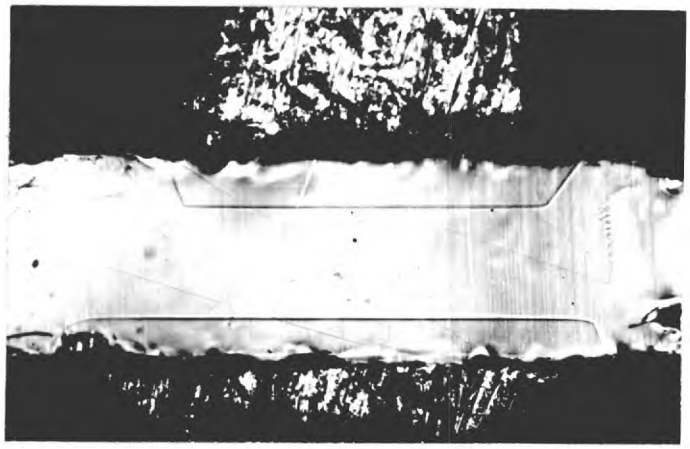
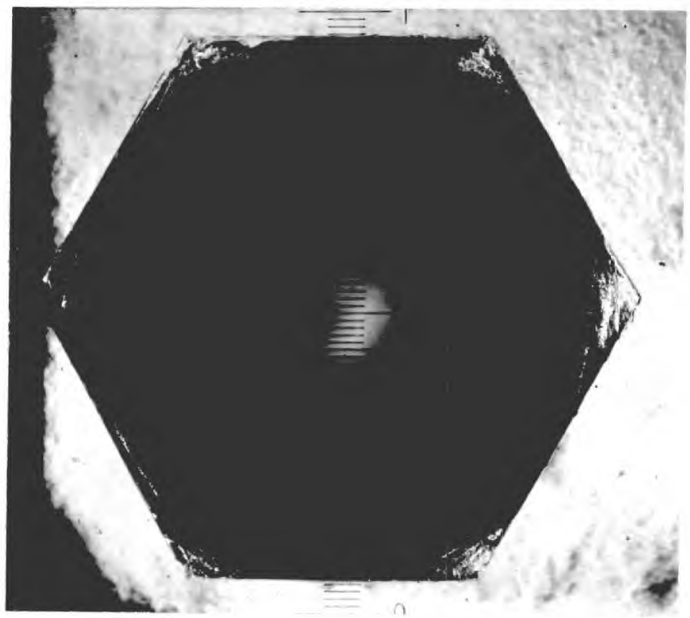


Figure 24. A triangular wetted area observed after alloying at 5°C/min on crystal C396/u . (x140)



(a)



(b)

Figure 25. Junctions prepared on {111} surfaces by prewetting and quasiequilibrium alloying ($dT/dt \sim 5^{\circ}C/min$) (a) Section of a double junction showing flat junctions on both A (upper) and B surfaces (x100) (b) A hexagonal wetted area typical of most junctions formed under the above conditions (x100).

described above, less than 10% exhibited any large departure from a regular hexagon. The regrowth (Figure 25(a)) was extensive and uniform and good electrical characteristics were obtained on both surfaces. Polar effects on the electrical characteristics will be discussed in section 5.3. From sections such as that shown in Figure 25(a) and the shape of the wetted areas, Figure 25(b), it was determined that the regrowth regions were bounded by only {111} and {100} planes as illustrated in Figure 26. The presence of {100} facets has not been observed in non-polar semiconductors such as Si and Ge. The significance of this observation will be discussed in later sections.

An elementary explanation of the above observations can be obtained by considering the processes involved in dissolution. In the simplest case, the rate of dissolution of a substance in chemical[28] or electrochemical[24] reactions is limited by one or both of two restraints. One restraint arises from the dissociation energy associated with breaking the interatomic bonds to free an atom and take it into solution. The other arises from the need for freed solute atoms to diffuse into the volume of the solvent. These same two mechanisms can determine the rate of metallic etching or alloying.

Under conditions which depart far from equilibrium, i.e. when the rate of temperature rise during alloying is very high, atoms can be freed more rapidly from any crystallographic surface than they can diffuse into the solvent. Thus the process of dissolution becomes independent of crystal orientation and the solid-solvent interface exhibits a diffusion-limited profile. When alloying is so limited the dissolved volume tends to take the shape of a diminished mirror image of the solvent volume. This effect has been observed in the junctions prepared

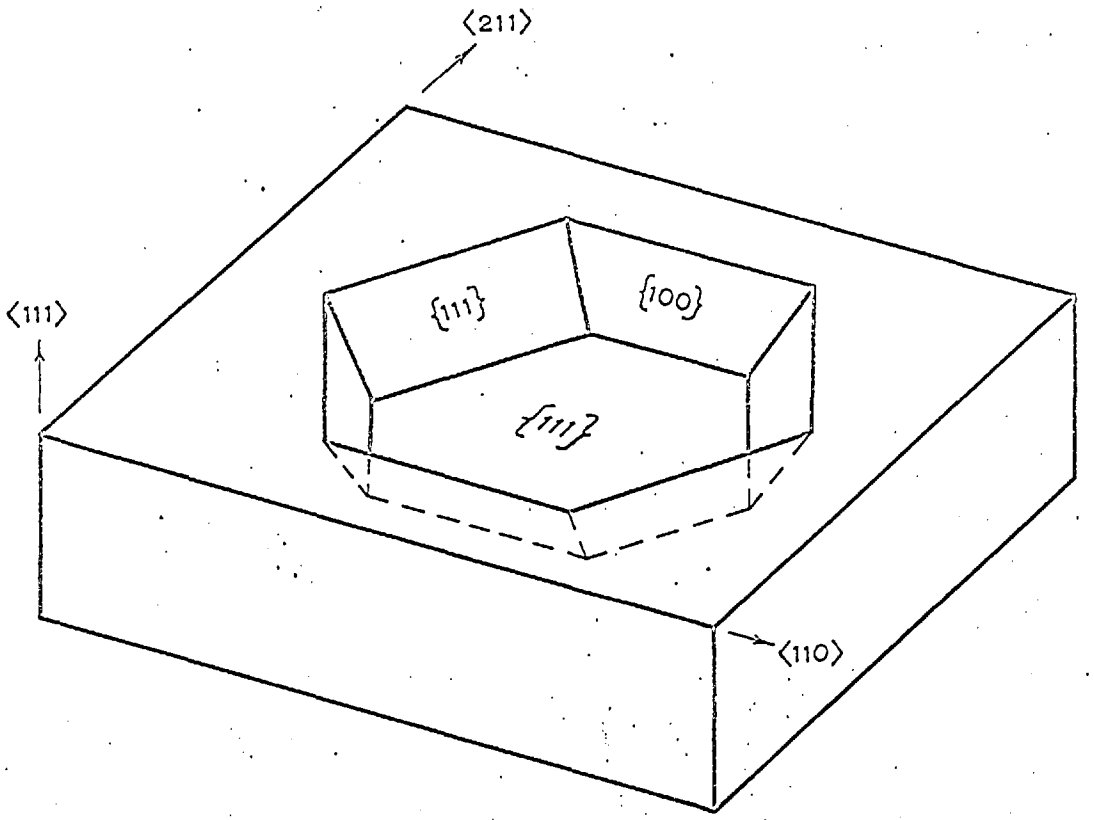


Figure 26. The shape of the dissolved volume obtained by quasi-equilibrium alloying on {111} surfaces of InSb. The {111} side facets are of opposite type to the uppermost {111} surface of the die and alternate with the {100} facets to form the sides of the dissolved volume.

under highly nonequilibrium alloying conditions and explains the junction profile shown in Figure 20(a).

When the rate at which atoms can be freed from a surface becomes important the most slowly dissolving surfaces tend to develop and to form the boundaries of the dissolved volume while faster dissolving surfaces tend to extinction. It is apparent from Figure 26 that the rate-limiting planes in InSb are $\{111\}$ and $\{100\}$. The tendency for wetted areas to spread more on B$\{111\}$ than on A $\{111\}$ surfaces gives direct evidence for a higher dissolution rate on A surfaces than on B. This evidence coupled with that of other autors[46][52] points to a considerable difference in the ease both of dissolution and of growth between A and B surfaces. It would be expected therefore that under some particular alloying conditions the dissolution process on one surface could be diffusion-limited while that on the other could be rate-limited. This effect has been observed in junctions prepared without pre-wetting under nonequilibrium alloying conditions, Figure 21(a) and (b). In this situation the contact angle between the pellet and the crystal was always greater than 90° and hence, over the contact area, the pellet appeared to be an infinite source of solvent. With no apparent shape limitations applied by the solvent volume, dissolution proceeded, limited largely by either one or other of the two restraints. In this case alloying on the B surfaces was largely rate-limited and produced flat junctions while that on the A surface was largely diffusion-limited and produced round junctions. When prewetting was carried out, however, the contact angle between the pellet and the crystal was always less than 90°. Even under conditions where rate-limitation would occur on both surfaces for an infinite solvent

source, in this situation the shape of the solvent volume caused diffusion-limited profiles near the edges where the amount of solvent was limited. The junctions in Figure 22 showed this tendency to be partly diffusion-limited and partly rate-limited under the same conditions on either A or B surfaces.

In the light of this discussion it is apparent that if the rate of temperature rise were decreased until the diffusion rate into the solvent became unimportant, (i.e. quasiequilibrium) the alloyed regions would become bounded by flat rate-limiting planes on either A or B surfaces. That this is so is demonstrated in Figures 25 and 26.

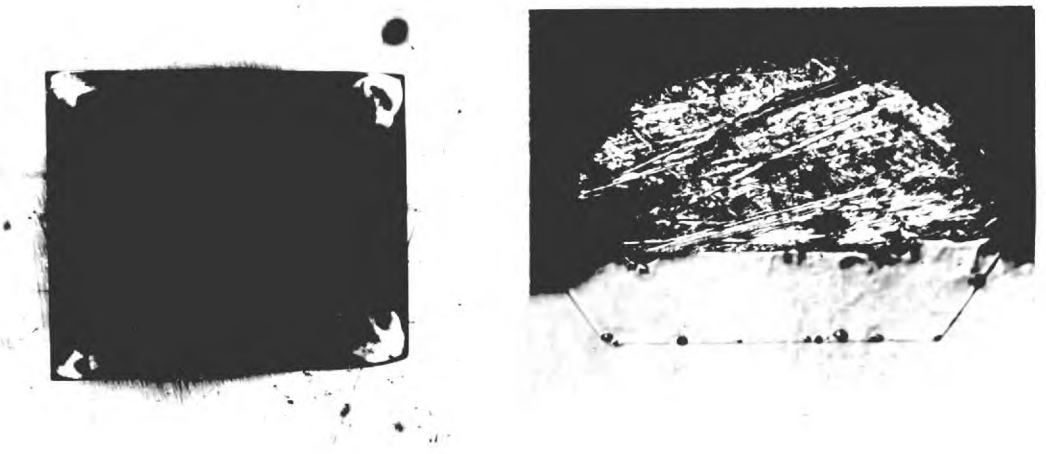
4.2. Alloying on {100} and {110} Surfaces.

4.2.1. Experiments on InSb.

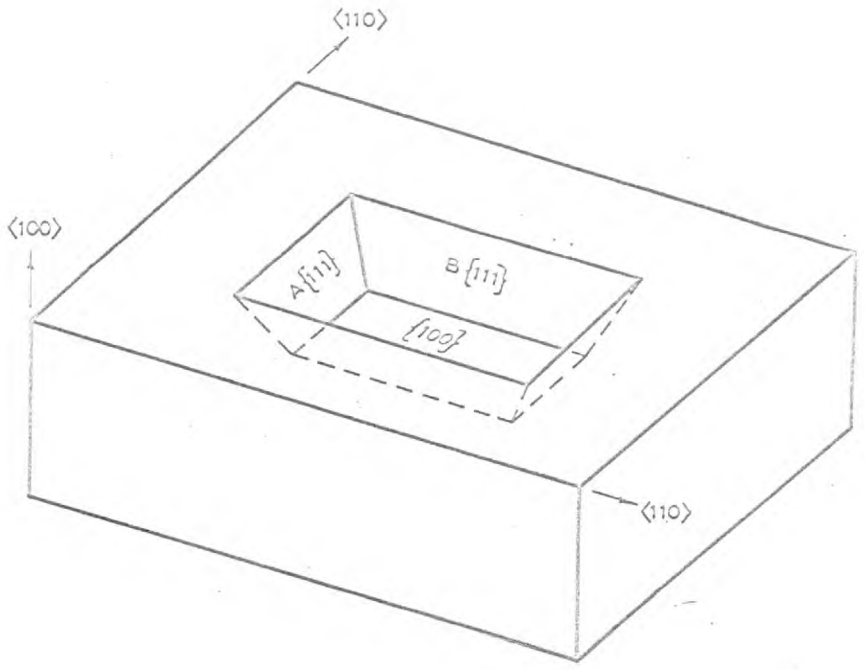
The following experiments were performed to determine if the {100} and {110} planes are rate-limiting, or facet forming, planes. This information is important for two reasons; first, the facet forming planes have a fundamental relationship to the processes of crystal growth and secondly, it would be possible to avoid polar effects in alloyed p-n junctions if good, flat junctions could be formed on non-polar surfaces such as {110} and {100} .

For these experiments {110} and {100} oriented dice were cut from InSb having a donor concentration of $10^{14}cm^{-3}$ and a dislocation density less than $5 \times 10^{3}cm^{-2}$. Pellets were radiation wetted to the dice and the junctions were alloyed to a maximum temperature of 400°C with the rate of temperature rise during the cycle always less than 5°C/min.

The surface wetted areas of junctions formed on {100} dice were invariably rectangular as shown in Figure 27(a). Sections such as



(a) (b)



(c)

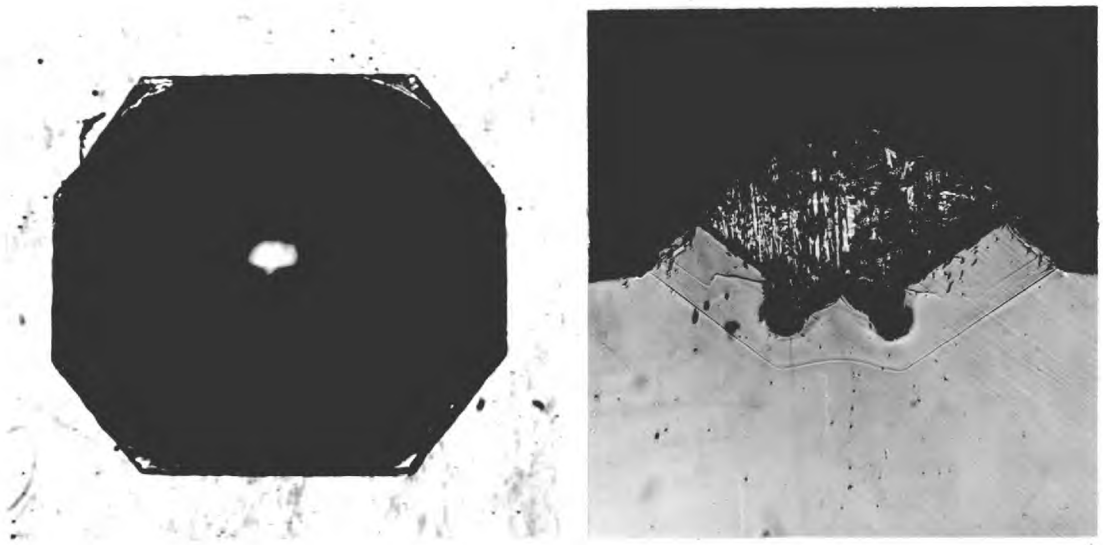
Figure 27. Junction prepared on a {100} surface of InSb by prewetting and quasiequilibrium alloying (a) Rectangular surface wetted area (x80) (b) Section showing the crystallographically flat {100} bottom and {111} sides (x80) (c) The shape of the dissolved volume, opposite sides are the same type of {111} surface and the B {111} surfaces form the long sides of the rectangle.

that of Figure 27(b) have shown that the {100} is a facet forming plane in InSb. The dissolved volumes were again defined by {111} and {100} planes only, as shown in Figure 27(c). It was also found that the long sides of the rectangular wetted areas were always associated with the B {111} surfaces. This further confirmed the conclusion drawn in section 4.1. about the relative rates of dissolution of A {111} and B {111} surfaces. It is of interest in this connection to note that as early as 1956 Millea and Tomizuka[196] reported the same rectangular morphology appearing on the {100} surfaces of InSb in experiments where the surface was observed as the temperature was raised slowly to the melting point. The straight sides of the melted areas on the surface were found to correspond to the {111} planes intersecting the surface but the rectangular shape was tentatively attributed to small temperature gradients. It would seem that until now the significance of these observations has been missed.

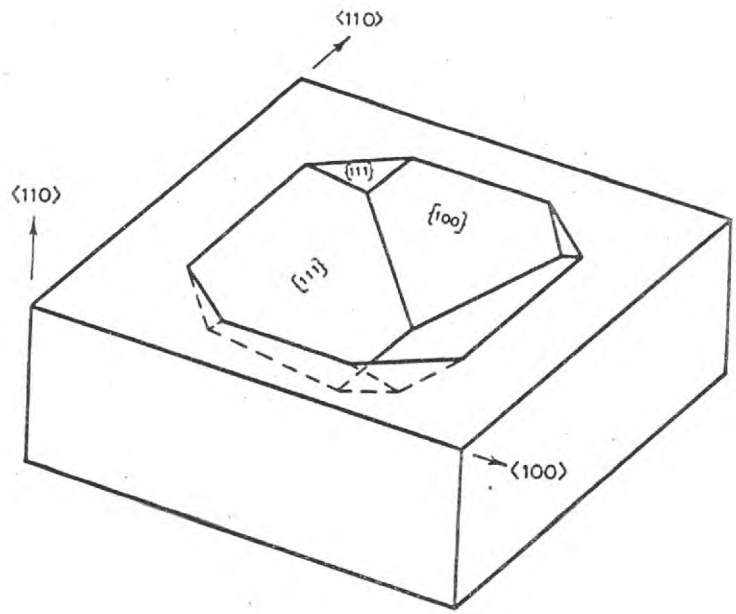
The surface wetted areas of junctions formed on {110} dice generally assumed the form of an irregular octagon as shown in Fig.28(a). Sections such as that in Figure 28(b) have shown no evidence to indicate the {110} as a rate-limiting plane. The shape of the dissolved volume was determined from the surface wetted area and the sections and was shown to be composed of {111} and {100} surfaces only, as shown in Figure 28(c).

The relative dissolution difficulty on A {111} , B {111} and {100} surfaces was estimated using an optical microscope to examine the sections and surface wetted areas. It was concluded that

B {111} > A {111} > {100} .



(a) (b)



(c)

Figure 28. Junction prepared on a {110} surface (a) Wetted area (x120) (b) Section through the well developed sides (x100) (c) The shape of the dissolved volume. Opposite sides of the trough are formed by large {111} facets of opposite type, the smaller {111} facets are opposite in type to the large facet which they touch.

4.2.2. Experiments on Germanium.

From the above experiments it was apparent that in InSb the {111} and {100} are the only facet forming planes. In elemental semiconductors such as germanium only the {111} has been reported[22][57][197][198][199] as a facet forming plane. If these observations are correct they imply that the polar nature of InSb in <111> directions has some effect on the nonpolar {100} surfaces. To make certain that other facet forming planes could not be observed in elemental semiconductors under similar conditions to those used for InSb, the following experiments were carried out on germanium.

Wafers of single crystal germanium oriented in <111>, <110> and <100> directions were cut from a crystal in the same manner used for InSb. Spherical pellets of indium were radiation wetted to the wafers and the junctions were alloyed in hydrogen. The maximum alloying temperature of the cycle was 700°C and from 500°C to 700°C the rate of temperature rise was 2°C/min. This slow rate of temperature rise gave a dissolution rate somewhat lower than that for the experiments on InSb and therefore it should have been possible to observe weaker faceting tendencies.

The observations of junctions alloyed on {111} surfaces confirmed the reports of other workers[22][55]. The wetted areas tended to be hexagonal with three sides rounded and three well defined. The rounded sides appeared in place of the {100} planes illustrated in Figure 26. When sectioned these junctions were shown to be bounded by flat {111} planes on the bottom and at the well defined sides, Figure 29(a).

Junctions alloyed on {110} surfaces were characterised by oblong wetted areas. In some cases the long sides were well defined and in all

(a)

(b)

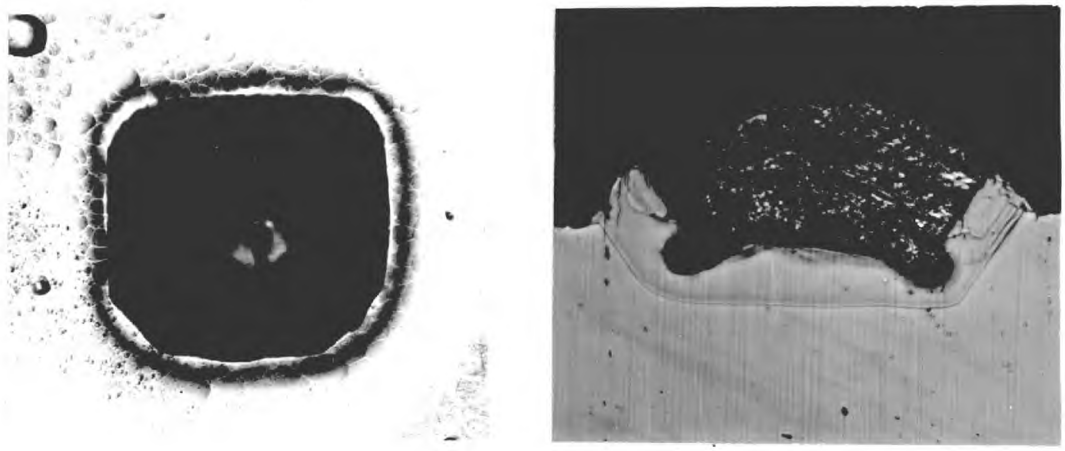
Figure 29. Sections of junctions formed on germanium (a) On the {111} surface, note the extremely flat junction (x100) (b) On the {110} surface, sectioned through the well developed {111} sides (x100).

cases they were found to be associated with the {111} planes forming a shallow angle with the {110} surface, see Figure 28(c). In a few cases the vertical {111} planes gave short well defined edges on the ends of the oblong wetted area but no evidence for the presence of {100} sides was observed. Sections such as that in Figure 29(b) revealed no facet on the {110} plane but only the well developed {111} side facets as observed in InSb, Fig. 28(b).

Junctions formed on {100} surfaces tended to have square wetted areas as shown in Figure 30(a). From the section shown in Fig.30(b) it is clear that the sides of the dissolved volume were formed by four {111} planes similar to those shown in Figure 27(c), except that A and B planes are the same in Ge and therefore the side {111} facets formed a square. The bottoms of the dissolved regions, although relatively even, were not crystallographically flat in any case. It must therefore be concluded that the {100} is not a facet forming plane in germanium.

In one instance the filament current was set very high during radiation wetting on a {100} surface. At the high pellet temperatures reached, indium was sputtered on to the hot germanium surface. These small indium pellets were subsequently removed by etching in dilute HCl. The etch pits left are shown in Figure 31 and clearly illustrate the shape of the dissolved volume on {100} surfaces.

This result is interesting because it demonstrates the advantages of the micro-alloying technique for investigating metallurgical effects. In the crystal growth or melt back experiments often used for metallurgical observations, the masses of the materials used are large and therefore small deviations



(a) (b)

Figure 30. Junction prepared on a germanium {100} surface (a) Square wetted area (x75) (b) Section showing flat {111} sides but no crystallographically flat bottom (x100).

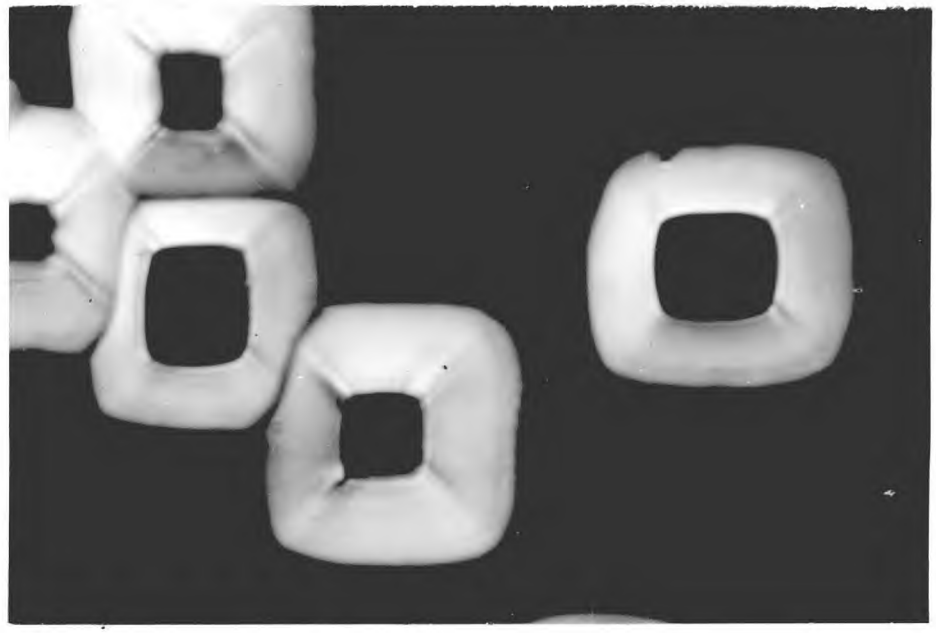


Figure 31. The pits left after etching away indium which had been sputtered onto a hot germanium {100} surface (x1500 under dark field illumination).

from equilibrium cause large rates of growth or dissolution. In decanting experiments, which have been used extensively to examine growth surfaces, it is very difficult to decant or quench quickly enough to avoid growth which occurs in conditions other than those for which observations are desired. In micro-alloying as exemplified by this sputtering experiment the amount of solvent is extremely small and therefore large departures from equilibrium (i.e. short alloying cycles involving relatively large temperature rates of change) give rise to dissolution or growth rates which are relatively low. For this reason well defined crystallographic planes appear as boundaries to the alloyed region. In the sputtering experiment effective quenching was also achieved because the rate of cooling at the end of the radiation cycle was exponentially greater than the prior heating rate. For this reason negligible regrowth occurred and the etch pits observed accurately represented the dissolved volumes. It should be noted, however, that it was not possible from the observations made on these small pits to determine the flatness of the {100} bottoms since small deviations such as those in Figure 30(b) would not appear under the illumination used.

4.3. Surface Models.

The discussion of rate-limiting planes given in section 4.1. can be used to explain all the observations in sections 4.1. and 4.2. except the difference in dissolution rates on A {111} and B {111} surfaces and the existence of the {100} as a facet forming plane in InSb while it is not in germanium. In order to

understand these observations it is necessary to know why a particular crystallographic plane is facet forming. Since facets form as a result of difficulty in nucleation during growth or in removing atoms during dissolution the nature of the atomic bonding in the surface must be considered. In the following parts of this section surface models for the {111} and {100} surfaces will be presented and in section 4.4. these models will be used to explain the metallurgical observations of this work and of the work of other authors.

4.3.1. {111} Surfaces.

In defining {111} surfaces in Figure 1 it was assumed that these surfaces are formed by triply bonded atoms and not by singly bonded atoms. This is reasonable on two accounts, firstly, in separating two parts of a crystal along a {111} surface, it would be considerably easier to separate the widely spaced indium and antimony layers than the closely spaced double layer; and secondly, a layer of singly bonded atoms would be inherently unstable. It is, therefore, generally assumed[24][27][55][200] that under ideal conditions the A surface is formed by triply bonded group III atoms while the B surface is formed by triply bonded group V atoms. It is probable in this situation that the group III atoms would be bonded to the surface by deformed sp^2 bonds and the group V atoms by deformed p^3 bonds. However, from geometrical considerations, Figure 32, the sp^2 bonds of the group III atoms would be directed at 20° from the favoured sp^3 bond direction of the crystal and thus the deformed sp^2 bonds would be highly strained in compression.

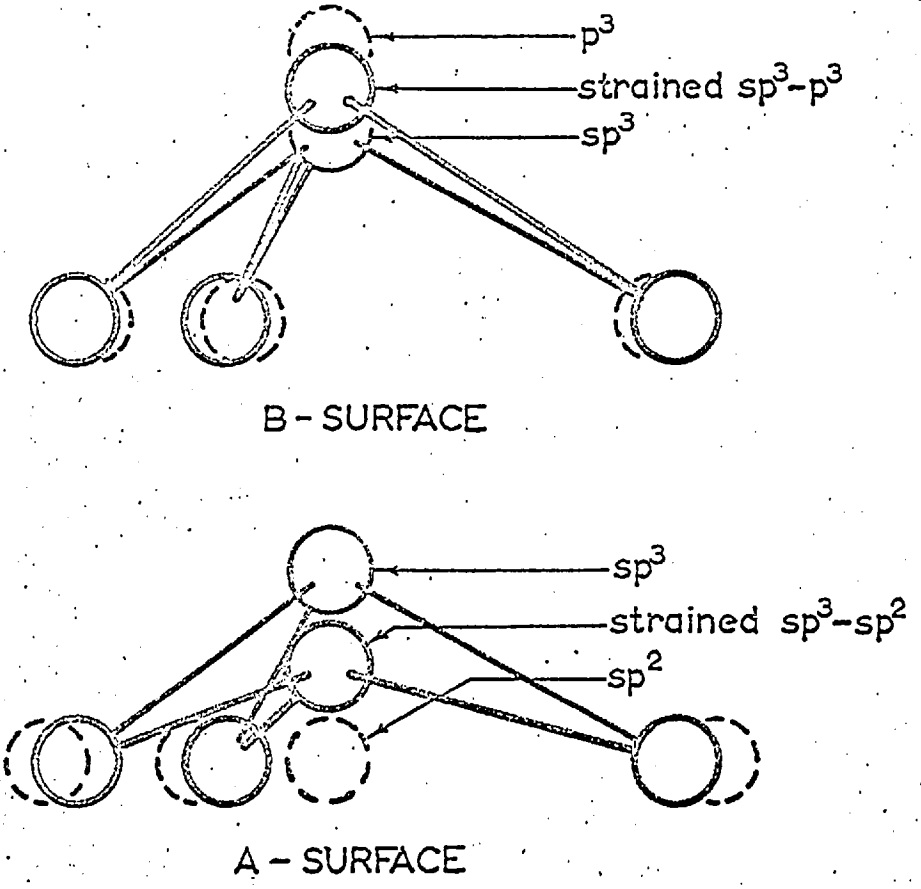


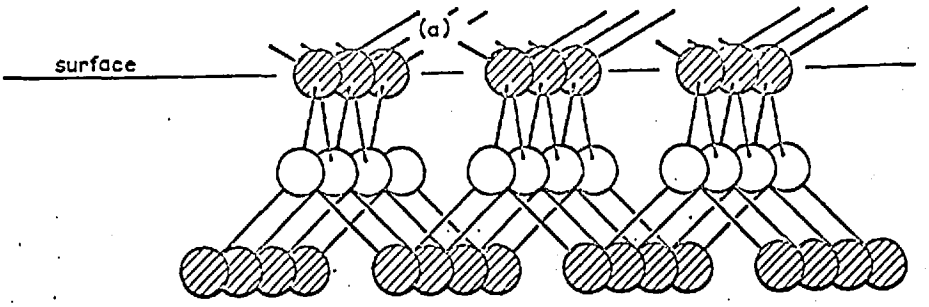
Figure 32. Models for A{111} and B{111} surfaces of III-V compounds. The position of the B type surface atoms (strained sp^3- p^3 bonding) and the A type surface atoms (strained sp^3- sp^2 bonding) are shown in bold lines. The ideal positions for p^3 and sp^2 bonding are shown in dotted lines.

Similarly the p^3 bonds of the group V atoms on the opposite face are directed at 15° from the sp^3 bond direction and thus the deformed p^3 bonds are strained in tension. The bond deformation required at a B surface is thus much smaller than that for the corresponding A surface, and p^3 bonds are well known[201] to display bond angles deviating from 90° by as much as 16°. It is apparent, therefore, that there is a considerable difference between the bonding of atoms in A surfaces and that in B surfaces.

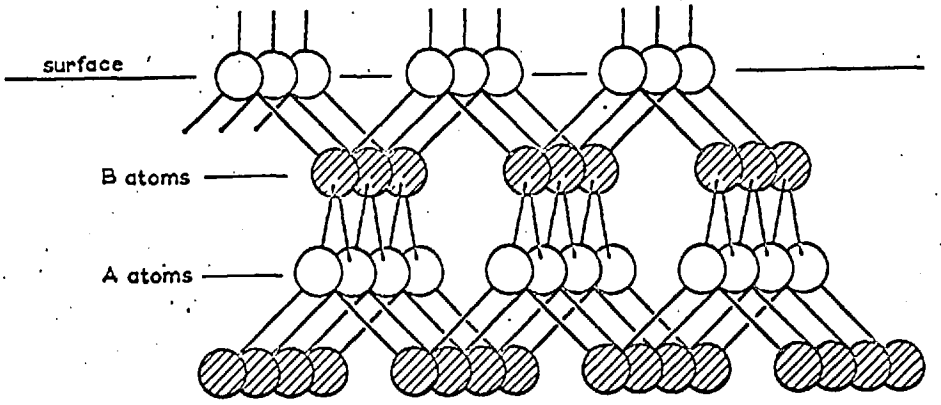
Gatos and Lavine[27] have used a similar model to interpret their observations of chemical and physical polarity effects. Their attempt to explain polar growth phenomena[47] is at variance with the present treatment and with experiment[46][52]. Holt[202] has criticised this model and suggested an electron resonance bonding approach. However, Holt's objections have since been answered by Gatos[203]. Haneman[204] has suggested a redisposition of surface atoms to minimize surface energy. The arguments that follow are not affected by such a redisposition.

4.3.2. {100} Surfaces.

It would seem reasonable on the basis of the electronic configurations of group III and group V atoms to assume that A atoms tend to form sp^2 bonds on any surface while B atoms tend to form p^3 bonds. If such bonding tendencies are considered to exist on {100} surfaces where every atom is doubly bonded to the crystal, then B atoms could be expected to assume the configuration shown in Figure 33(a). It is possible that because of the deformation towards sp^3 of the p^3 bonds that the s orbitals of the B atoms are



(a) B{100} surface



(b) A{100} surface

Figure 33. Models for {100} surfaces in III-V compounds (a) Composed of B atoms in p^3- tending to sp^3 bonding configuration (b) Composed of A atoms in the sp^2 bonding configuration so that dangling bonds are perpendicular to the surface.

also deformed towards the sp^3 configuration. In any case they easily change to sp^3 when a suitable atom moves into a bonding position. It is apparent that a {100} surface composed of B atoms provides many favourable sites for further atomic bonding (i.e. site (a) in Figure 33(a)) since all atoms bonding to the surface can form two bonds. Such a surface is similar, therefore, to the {100} surface of non-polar semiconductors.

It is possible for A atoms to bond into a {100} surface in a manner which gives only a small ($\sim 10^{\circ}$) deviation between the sp^3 crystal bonding configuration and the sp^2 bonding tendency of the A atoms. Thus A atoms bond stably into a {100} surface with relatively little strain and with single dangling bonds as shown in Figure 33(b). With such a configuration a {100} surface composed of A atoms appears similar in some respects to a {111} surface since it provides no favourable sites for further atomic bonding. That is, initially, an atom can form only a single bond to the surface layer. Now since in <100> directions the crystal canbe considered as alternate layers of A and B atoms as shown in Fig.33, it is possible that {100} surfaces are always composed of A atoms in the stable configuration illustrated in Figure 33(b). Such a surface is considerably different from the {100} surface presented by a non-polar semiconductor and the implications of such an atomic configuration on the metallurgy of {100} surfaces will be considered in the following sections.

4.4. Crystal Growth and Dissolution.

4.4.1. Introductory Considerations.

The model for {111} surfaces, although relatively simple, is capable of providing an explanation for the chemical and physical polar effects reviewed in Chapter 1. This is substantial justification for using similar principles in considering the {100} surfaces and for attempting to use the {111} model to explain the metallurgical polar effects. The way in which the model explains the chemical and physical polar effects will first be described briefly.

On the basis of the model for {111} surfaces, it is reasonable to assume that the group III surface atoms become active electron acceptors in opposition to their usual metallic role as electron donors. Thus atoms, ions, or molecules which can donate electrons could bond to the group III surface atoms, completing the sp^3 bonding structure in the surface and releasing the strain of the deformed sp^2 bonds. On the contrary, however, the group V atoms remain electron donors since their s orbitals are assumed not to be involved in bonding. Since the group III atoms have all valence electrons involved in bonding it is to be expected that the A surface would be chemically more noble than the B surface. It is also apparent, because of the two loosely bound s electrons associated with each atom on the B surface, that this surface would be much more subject to chemical oxidation. In the case of anodic oxidation a combination of the electron excess on the B surface and the weaker bonding of the group V atoms would lead to faster oxidation. The preferential etching of dislocations on the A

surface would also be expected from this model and has been treated in some detail by Gatos and Lavine.[27]

Physically, the greater strain involved in the bonding of group III surface atoms would give rise to the greater degree of disorder observed in X-ray experiments on A surfaces. The observation that mechanical damage arising from controlled abrasion is of lesser extent on A than on B surfaces[33] and that decomposition temperatures are higher on A surfaces[38] might be interpreted in terms of a 'prestressing' of the A surface bonds, where the initial compression must be overcome before the bond can be put in tension and broken. The experiments of Hanneman, Finn and Gatos[35][36] could also be interpreted in terms of such a prestressing.[205]

Crystal growth in III-V compounds has been treated at some length by Sangster.[200] However the polar nature of the crystal was not considered beyond its effect on the stoichiometry of the growing surface. To apply the above surface models to the explanation of the observed metallurgical effects requires further consideration of the processes of nucleation and crystal growth, as influenced by the different properties of the A {111}, B {111} and {100} surfaces. In so doing however it must be realized that the structure of the surface layers of atoms at a solid-liquid interface will by no means be that of the solid crystal at lower temperatures. The concentration of defects, vacancies and interstitial atoms in the surface layers will of course be very large. For simplicity these features will not be expressly considered, and the discussion will consider an ordered array. The degree of order must in fact increase very rapidly as one considers deeper

lying layers of atoms in the solid, since all III-V compounds exhibit a high degree of stoichiometry, and the self-diffusion coefficients of the individual components are very low.[83]

It is generally considered that crystal growth occurs by a step process[206][207][208]. Because a complete discussion of growth processes would not be relevant to this thesis only growth on surfaces which may be formed of steps bounded by A {111}, B {111} and {100} planes will be considered. These are surfaces in the families {hh1} and {h11} and represent two extremes in the polar surfaces not of the form {111}. Ideally it is not possible to construct surfaces in these families which are not composed of stepped {111} and {100} planes. Since atoms are densely packed in the {111} and {100} planes it is possible that these are the growth-limiting planes in any case. The experimental observations of sections 4.1. and 4.2. indicate that the only facet forming or growth-limiting surfaces in InSb are the A {111}, B {111} and {100} planes. These observations are also confirmed in crystal growth experiments[23][209] where the same planes have exhibited faceting. Dendritic growth often results in crystals which are almost entirely bounded by stepped {111} surfaces[210] and the only planes delineated by growth striae under constitutional supercooling conditions are the {111}.[211][212] It is not unreasonable then in treating growth on an idealized surface to assume that crystal growth in III-V compounds on surfaces of the form {hh1} or {h11} occurs by steps bounded by A {111}, B {111} and {100} planes. Steps as small as a {111} double layer or a {100} monolayer will be allowed. It will also be assumed that the

dissolution process is the reverse of the growth process.

4.4.2. Growth on {111} surfaces.

With the above considerations in mind, the implications of the models of A and B surfaces on growth processes can now be considered. The concentration of atoms bonding to a surface is determined by the change in free energy during bonding. The attachment of atoms by means of single bonds is therefore unfavourable since the change in free energy is a minimum. The change in free energy, ΔG, can be written formally as

$$\Delta G = \Delta H - T\Delta S$$

where ΔH is the change of enthalpy or total heat, ΔS is the change of entropy or disorder of the system and T is the absolute temperature.

In the case of triply bonded B atoms, the p^3 bonds are only slightly deformed. The triply bonded A atoms on the other hand must form highly strained sp^2 bonds. The higher strain in the bonds of A atoms arises not only because of the larger difference between sp^2 and sp^3 bonding configurations but because of the compressive nature of the strain. This is illustrated in Figure 34 which also shows how, in either case, the strain gives rise to an increase in bond energy. When an atom singly bonds the strain energy of the corresponding triply bonded atom is released and thus the change in free energy for single bonding is increased and can be expressed as

$$\Delta G = (\Delta H_{bond} + \Delta H_{strain}) - T\Delta S$$

However, because of the greater strain on the A surface,

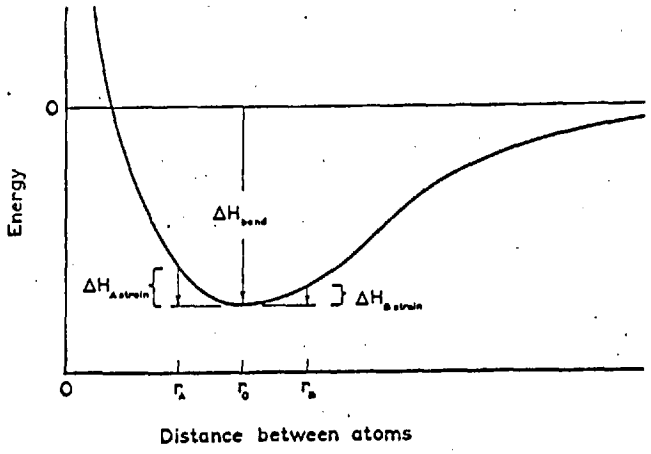


Figure 34. The energy-distance relationship between two atoms (after Joos[213]), r_o corresponds to the bonding or minimum energy position. Note that for similar departures from r_o the compressive strain energy $\Delta H_{A\ strain}$ is greater than the tensile strain energy $\Delta H_{B\ strain}$.

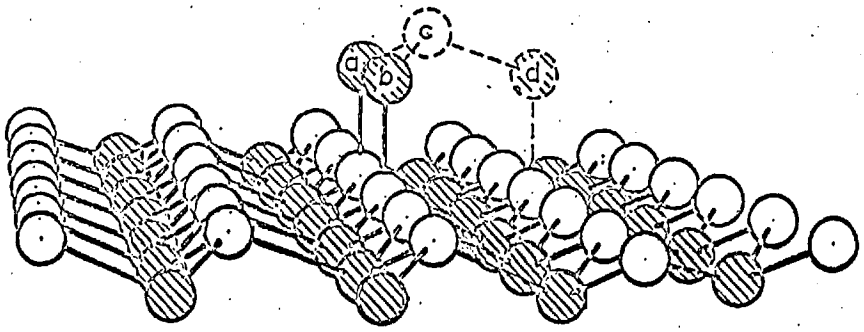


Figure 35. Formation of a nucleus on a {111} type surface. Atoms (a) and (b) must singly bond on neighbouring sites; then atoms (c) and (d) can form double bonds.

$\Delta H_{A\ strain}$ will be greater than $\Delta H_{B\ strain}$, and since a term of the form

$$\exp\ (\Delta H_{strain}/RT),$$

where R is the universal gas constant, will enter into either the mass action law, or into the relevant reaction rate equations, the corresponding stage of the growth process on the A surface will be favored.

The nucleation and propagation of steps on A $\{111\}$ and B $\{111\}$ surfaces will now be considered. Nucleation requires that at least two atoms singly bond as next nearest neighbours before growth involving double bonds can proceed, Figure 35. The probability of two atoms singly bonding as next nearest neighbours is much lower than that for the bonding of a single atom. The appropriate expression for the nucleation process will therefore contain a term of the form

$$\exp\ (\alpha\, \Delta H_{strain}/RT)$$

where α is greater than unity. The value of α rapidly decreases as the nucleus grows and the probability of an atom landing on a site which will cause further growth increases with the size of nucleus. It follows then, that for similar departures from the equilibrium liquidus temperature, the nucleation rate on an A surface will be considerably greater than that on a B surface where the strain energy is less.

On the basis of the assumptions made in section 4.4.1., the propagation of a nucleus on a $\{111\}$ surface can occur by means of A$\{111\}$ - B$\{111\}$ steps or $\{111\}$ - $\{100\}$ steps. The latter will

be considered in the next section. The propagation of A {111} - B {111} steps, Figure 36, will be considered in two stages. In the first stage at least one atom (e.g. atom (a) in Figure 36) must singly bond in the corner of the growth step. In the second stage subsequent atoms form pairs of bonds (e.g. atoms (b) and (c) in Figure 36) and the entire corner of the original step quickly becomes occupied. This second stage is favoured because atoms of both species are incorporated at the same rate and all atoms added can form two bonds simultaneously. At this stage the bonding energy may be decreased by up to $3\Delta H_{strain}$ but ΔH_{strain} is a small fraction of the total bond energy[35] and thus the net change in free energy is still almost twice as great as that involved in the first stage. The rate-limiting step in the process is therefore the formation of the first single bonds. Further growth requires the repetition of this two-stage process, thus microscopic growth proceeds causing macroscopic lateral movement of the step.

4.4.3. Growth on {100} Surfaces.

Because it is possible for growth on {111} surfaces to be propagated by {111} - {100} steps, it is necessary to examine the nature of crystal growth on {100} surfaces before proceeding to further investigate the implications of the difficulties of nucleation and of step propagation on {111} surfaces. In section 4.3.2. it was pointed out that, because an atom can form two bonds when bonding into a {100} surface composed of B atoms, the change in free energy is large and the process is favoured. Such a surface provides no difficulty to growth and atoms can bond into the

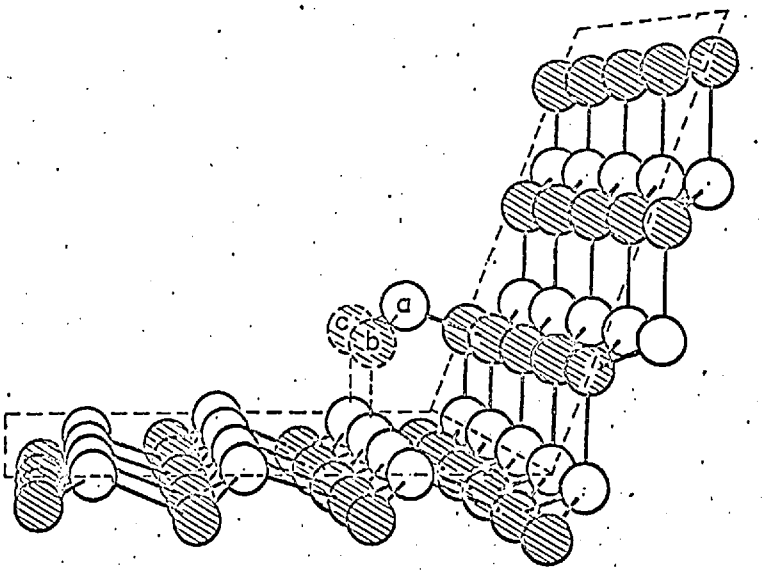


Figure 36. Growth on an A{111} - B{111} step. Atom (a), for example, must form a single bond; then the corner of the step can be completely filled by atoms which double bond (i.e. atoms (b) and (c)).

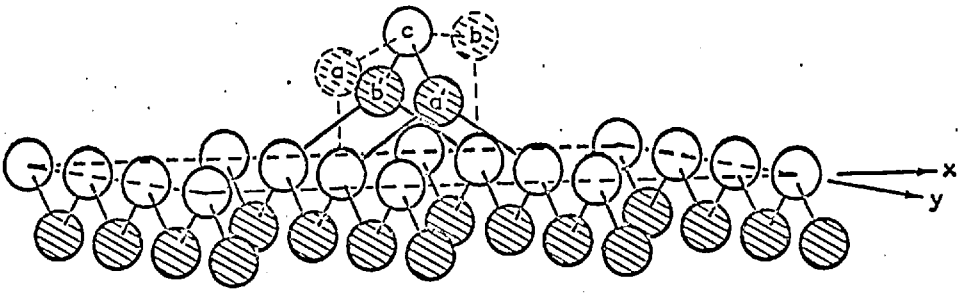


Figure 37. Formation of a nucleus on a {100} surface. Atoms (a) and (b), for example, must singly bond before atom (c) can doubly bond and the whole structure doubly bond into the (a') (b') (c) arrangement.

lattice at any point on the surface without the need to form a nucleus. Growth on the {100} surface of a non-polar semiconductor is similar and probably slightly more favourable because of the lack of any strain in the surface bonds. On a {100} surface composed of A atoms, however, atoms bonding to the surface can form only one bond and therefore, as discussed in section 4.4.2, growth is less favoured. Furthermore, a difficult nucleation stage is required on this surface, similar to that on {111} surfaces, before the relatively easier step propagation can proceed. Several possibilities for the nature of the nucleation process exist and the relative probabilities of each can be determined only if the ratio of strain energy to bond energy and the dependence of strain energy on deformation are known. However, since the establishment and propagation of any nucleus which can be considered depends on the same arguments applied in different sequences, the principle involved can be demonstrated by considering a particular nucleation process. Nucleation by what appears to be the most probable process will be considered here.

Nucleation on an A {100} surface (hereafter called {100}) requires that at least two atoms singly bond as next nearest neighbours in a direction where $|x| = |y|$ (see Figure 37) before growth involving the formation of double bonds can occur. When two atoms so bond, as for example atoms (a) and (b) in Fig. 37, a third atom, (c), can doubly bond between them. A combination of the desirable bond configuration for atom (c), which is an A-type atom, and the second order interaction of bonds lower in the crystal causes this structure to assume a configuration tending

to the desired crystal configuration. In assuming the crystal configuration indicated by atoms (a') and (b'), two bonds are formed and therefore the process is energetically favored. The structure formed by atoms (a'), (b') and (c) can be considered the smallest nucleus capable of achieving any degree of stability.

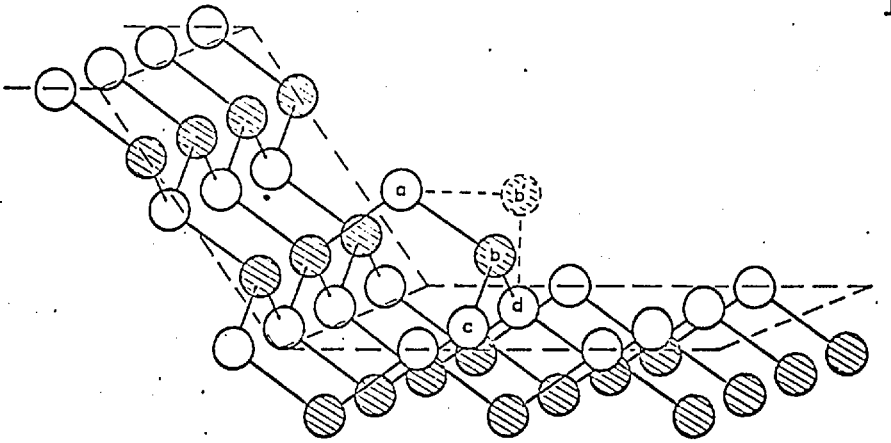
The probability of two atoms singly bonding as described above will, of course, be much less than the probability of a single atom singly bonding. In this instance it will even be less than the probability of two atoms singly bonding as next nearest neighbours on a {111} surface because the number of possible combinations is smaller. However, as already discussed, other nucleation processes could be considered (i.e. two atoms could singly bond as next nearest neighbours in the x-direction or one atom could doubly bond to two adjacent atoms in the y-direction) and if the probabilities of all these processes were combined it is possible that the overall probability for nucleation on a {100} surface would be greater than that for either an A {111} or a B {111} surface. In any case it is apparent that nucleation is required for growth on a {100} surface and that the difficulties encountered in establishing such a nucleus are similar to those encountered on {111} surfaces in both polar and non-polar semiconductors.

The subsequent propagation of the nucleus of Figure 37 occurs by means of {100} - {111} steps in the x and y directions. It is obvious that there are two different types of step: the {100} - A {111} propagating in the x-direction and the {100} - B {111} step propagating in the y-direction. These two

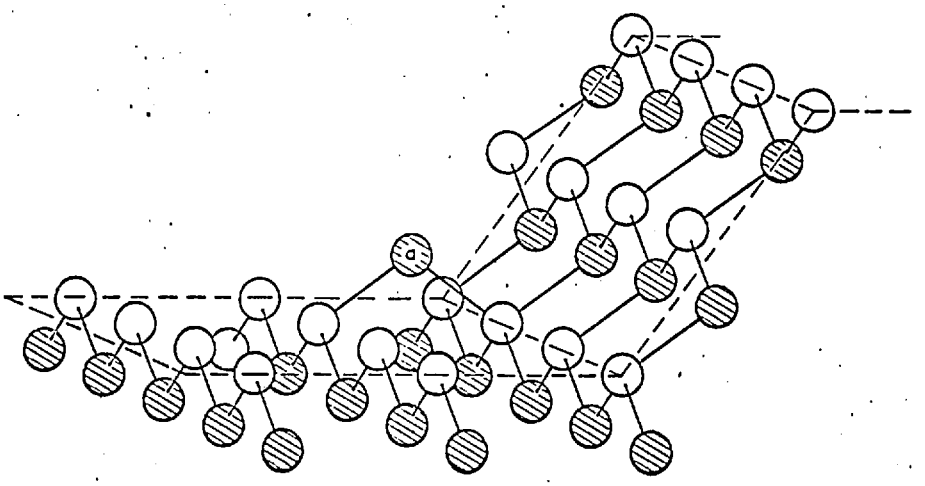
types of step will be considered separately and are illustrated in Figure 38(a) and (b).

The {100} - B {111} step is analagous to the A{111} - B{111} step in that its propagation involves two stages: a difficult stage and an easy stage. However each stage is somewhat more complex since there are several possible ways in which each can proceed. The difficult stage is to establish a configuration in the step similar to that of atoms (a) and (b) in Figure 38(a). This could result from atom (a) singly bonding followed by atom (b) either triply bonding in position (b) or doubly bonding in position (b') and then singly bonding into the (b) position. It could also result from atom (b) singly bonding in position (b') followed by atom (a) doubly bonding and then a further single bond being formed to give the (a) - (b) configuration. Once this difficult step is accomplished the atoms (c) and (d) have been forced closer to an sp^3 configuration and therefore growth along the step corner, whether by double bonding or a combination of single and double bonding, is relatively easier. These two stages repeat continuously during step propagation and again the rate of propagation is limited by the difficult stage.

Propagation of the {100} - A{111} step would seem at first sight to be easy since all atoms bonding into this step could form two bonds. However, B atoms tend to bond into the corner of this step and the angular difference between the surface bonds of the crystal and the p^3 bonds of the B atoms is 51°. The interaction of such bonds tends to be weak and therefore the tendency to form double bonds is weak as it is to a greater extent on the {100}



(a) {100} - B{111} step



(b) {100} - A {111}

Figure 38. Step propagation (a) by means of {100} - B{111} step where a stage of single bonding (i.e. atom (a) or (b')) is required and (b) by means of a {100} - A{111} step where propagation involves double bonding.

surface where the angular difference between the crystal and the B atom bonds is 90°. The first atom could become double bonded into the corner either by striking a site just when the thermal movement of the atoms causes that site to assume a favourable bonding configuration or by singly bonding into the corner of the A{111} surface and sticking sufficiently long for thermal movement to bring the corresponding {100} atom into a favourable configuration for the second bond to form. In either case the probability of the first atom (for example atom (a) in Figure 38(b)) double bonding is low but once such an atom has bonded the neighbouring atoms in the {100} surface become strained towards the sp^3 configuration making subsequent double bonding more favourable. Thus again growth on a {100} - A {111} step proceeds in two stages, the first difficult and the second relatively easier.

On the basis of the above discussion of {100} - A{111} and {100} - B {111} steps, the discussion of A{111} - B {111} steps in section 4.4.2. and the conclusion drawn in section 4.2.1. with regard to the relative ease of growth or dissolution on A{111}, B{111} and {100} surfaces, it would seem reasonable to place the ease of propagation of the above steps in the following order:

$$\{100\} - A\{111\} > \{100\} - B\{111\} > A\{111\} - B\{111\} .$$

4.4.4. Growth on {110} surfaces.

The consideration of nucleation processes on {111} and {100} surfaces has resulted in an explanation for these planes developing as facets during crystal growth and dissolution. For the sake of completeness the nature of the {110} surface will also be

considered. If the Bravais rule[214] which relates faceting tendencies to the atomic density in a plane were applied, faceting in the zinc-blende or diamond structure would be expected on {110} surfaces in preference to either the {100} or {111} surfaces. However, if the atomic structure of the {110} surface, as illustrated in Figure 39, is examined, it is apparent that the entire surface can be considered to be composed of A{111} - B{111} steps. Growth on such a surface involves the propagation of these steps as discussed in section 4.4.2. That is, the single bonding of one atom (e.g. atom (a) in Figure 39) in each step is followed by double bonding down the full length of the step. Therefore no severe nucleation difficulties exist on {110} surfaces and no {110} facets would be expected.

4.4.5. Implications on Metallurgical Phenomena.

If crystal growth in a ⟨111⟩ direction is considered it is apparent from section 4.4.2. that the nucleation of new steps is the most difficult stage of growth and therefore faceting would be expected on {111} surfaces. It is also clear that initial nucleation on the B surface would be much more difficult than that on the A surface and thus the overall growth or dissolution rate for comparable unconstrained growth or alloying conditions would be greater on the A surface. Similarly the treatment of {100} surfaces in section 4.4.3. has shown that a difficult nucleation stage is required for growth in polar, but not in non-polar semiconductors. Faceting on {100} surfaces would be expected, therefore, in the polar case only. Thus the models

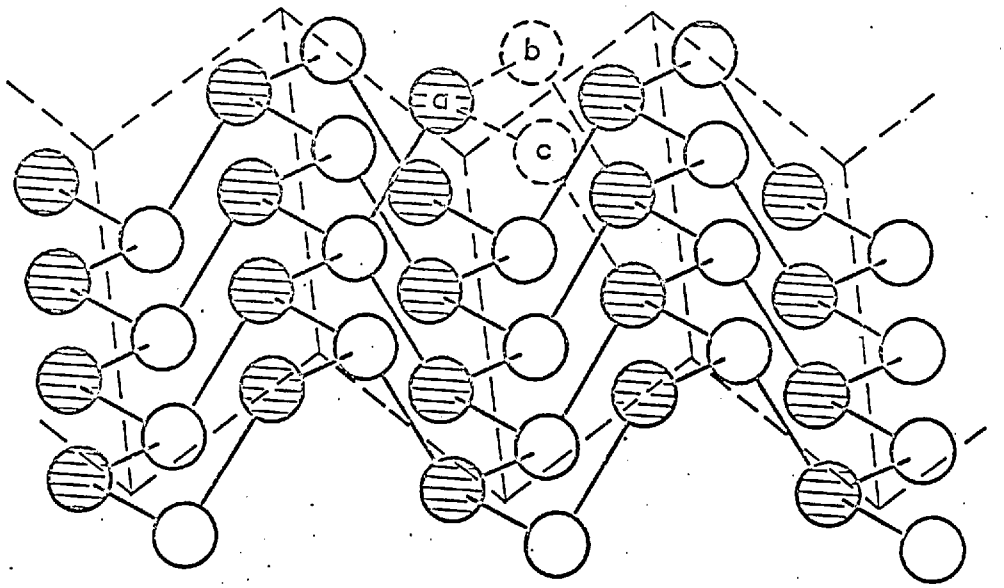


Figure 39. A {110} surface illustrating how it can be considered to be composed of A{111} - B{111} steps.

for $\{111\}$ and $\{100\}$ surfaces have explained the present experimental observations as well as those of other authors.[46][52] It would, however, be both reassuring and profitable if these models could also be used to explain or achieve a more fundamental understanding of the other polar metallurgical effects which have been reported. In the remainder of this section faceting, growth surface morphology, facet segregation coefficients, crystal growth in $\langle hh1 \rangle$ and $\langle h11 \rangle$ directions, twinning and the introduction of dislocations during crystal growth will be discussed in terms of these surface models.

Because nucleation is more difficult on B than on A surfaces the amount of supercooling required for growth on a B surface will be substantially greater than that required for similar growth on an A surface. Since less supercooling is required for step propagation following nucleation, the formation of large B facets during crystal growth and during the converse process of dissolution would be expected. Figure 40 illustrates how the larger supercooling required for nucleation on B surfaces leads to large central B facets on B oriented crystals and large B side facets on A oriented crystals. The smaller supercooling required on A surfaces leads to smaller A facets. The difference between the supercooling required for nucleation and step propagation gives rise to a pulsating type of growth which leaves the facet free to absorb impurities from the melt during each nucleation period. The combination of absorption and rapid step movement upon nucleation could easily lead to the anomalous segregation coefficient observed on $\{111\}$ type facets.[23] Since A and B

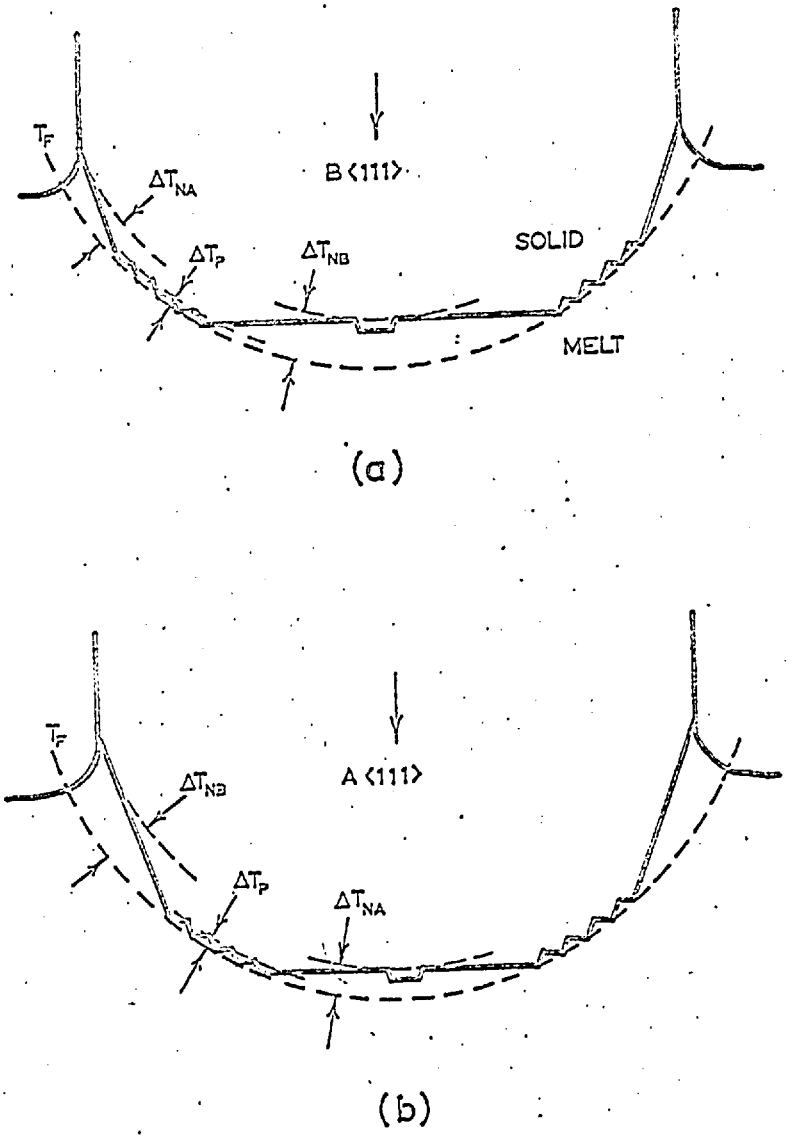


Figure 40. Growth face morphology for growth in (a) B⟨111⟩ and (b) A⟨111⟩ directions with a convex freezing isotherm T_F. The supercooling required for step propagation ΔT_p, for nucleation on an A surface ΔT_{NA} and for nucleation on a B surface ΔT_{NB} are illustrated. The difference in facet development as a result of the difference between ΔT_{NA} and ΔT_{NB} can be seen by comparing (a) and (b).

surfaces probably have different absorption properties and step propagation on the highly supercooled facets is probably very fast, it would not be surprising if a difference in facet segregation coefficient were observed on A and B facets. These conclusions are in agreement with experimental observations.[23][54][215]

The family of directions ⟨hh1⟩ can be considered to give growth surfaces composed of a roof-like structure of A {111} - B{111} steps (see Figure 36). In this case alternate roof faces are of opposite polarity type and step propagation involves a stage of relatively difficult single bonding which is polarity dependent. Because of the unequal lengths of the two types of surface in this roof structure, growth in orientations of the form ⟨hh1⟩ would be expected to show polarity dependent properties ranging from the extreme of ⟨111⟩ to the absence of such effects in the ⟨100⟩ directions.

The family of directions ⟨h11⟩ can be considered to give growth surfaces composed of a roof-like structure of {100} - {111} steps. However, growth in an A ⟨h11⟩ direction would involve {100} - A {111} steps while that in a B ⟨h11⟩ direction would involve {100} - B {111} steps. Because of the difference in the nature of propagation of these two steps, growth in ⟨h11⟩ directions would also be expected to show some polarity dependence. The relative ease of propagation of {100} - {111} steps over A {111} - B {111} steps could lead to easier crystal growth in ⟨h11⟩ directions than that in ⟨hh1⟩ directions. There would seem to be some experimental support for these considerations since it is found[200] that orientations of the form ⟨h11⟩ are

particularly favourable for crystal growth and that they do exhibit some polarity dependence.

Twinning is one of the most common causes of failure in the growth of single crystals.[3][46] Twinning occurs more frequently in A oriented crystals than in those oriented in the B direction. From the previous discussion of growth rates, it will be assumed that an A oriented crystal is characterised by a rounded convex crystal-liquid boundary with well developed B side facets while B oriented crystals have a large central B-facet and less well developed A side facets (Figure 40). Twinning, if it occurs, will be more likely on the B facets since the greater rate of step propagation consequent upon nucleation of these facets will enhance the chance of stabilizing and propagating an initially mis-oriented nucleus. Since twinning on a central facet is very rare,[3] it is probable that the formation of a mis-oriented nucleus is associated with the gas-liquid-solid boundary. The presence of insoluble impurities at this point[216] or the presence of oxygen in the ambient gas[199] both enhance the chance of twin formation. If insoluble impurities or oxidation are a source of mis-oriented nuclei, it would not be surprising if the probability of forming a mis-oriented nucleus were higher on the chemically active B surface. Therefore, if a mis-oriented nucleus can form more easily and is stabilized more readily on a B facet than on an A facet, a higher incidence of twinning in A oriented crystals where the B side facets extend to the liquid-gas boundary would be expected.

Dislocations may be introduced into growing crystals in a

variety of ways. Some of these are: (a) to relieve the strain caused by the incorporation of impurities,[206] (b) by propagation from the seed crystal[217] and (c) by the occurrence of slip to relieve thermal stresses accompanying pulling.[218][219]

The concentrations of impurity atoms, except under conditions of constitutional supercooling,[206] will seldom be large enough to have any significant effect and this discussion will be limited to growth from pure stoichiometric melts.

The propagation of dislocations from the seed, or arising during wetting of the seed has been considered by Dash[217]. This mechanism may be one of the major sources of dislocations in the low melting point semiconductors.[199]

The introduction of dislocations by thermal stress can be significant in the growth of the high melting point semiconductors, where large thermal gradients may be encountered while the pulled material is still in its plastic condition. This case will be considered since there is some evidence of differences in dislocation densities arising from the use of differently oriented seeds in GaAs.[45]

The stress patterns arising as a result of the thermal geometry of a typical pulling furnace have been discussed by Billig[218] and Van Bueren et al.[219] It is concluded that the active components of the stress pattern are of the general form shown in Figure 41, and the magnitude of these stresses decreases exponentially towards the seed end of the crystal.

In the zinc-blende and diamond structures the principal slip planes are {111} planes. The relative positions of these planes

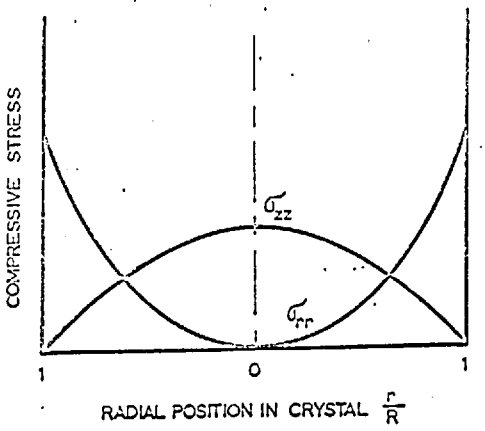


Figure 41. Thermal stress distribution in a cylindrical crystal of radius R under conditions of radial cooling. σ_{zz} is axial compressive stress σ_{rr} is radial compressive stress. (After van Bueren et al[219])

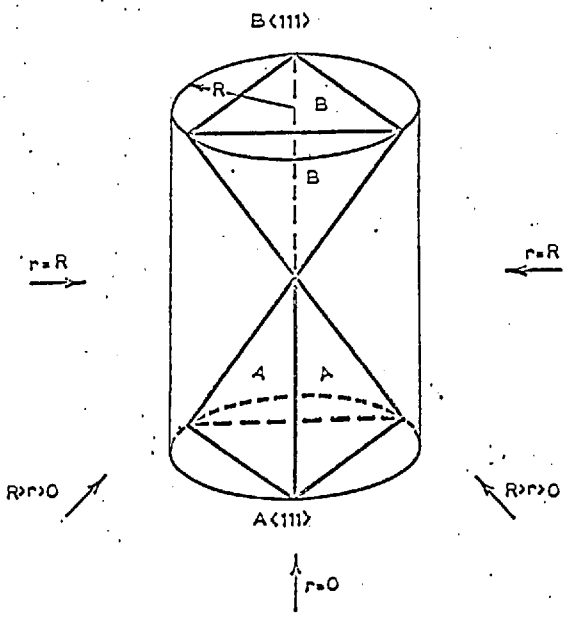


Figure 42. Relative positions of the tetrahedra of A{111} planes and of B{111} planes. The arrows indicate the resultant thermal compressive stress for various radial positions in a crystal growing in the A⟨111⟩ direction.

in a crystal growing in the A $\langle 111 \rangle$ direction is shown in Fig.42. If the outer regions of such a crystal are considered, the stress will be a pure hoop stress and slip will occur equally on both sets of inclined A $\{111\}$ and B $\{111\}$ planes. If smaller radii are considered, then from Figure 42 is obvious that the resultant stress direction will have rotated and slip will occur principally in the A planes. Finally at the centre the stress will be pure axial compression and hence slip could again occur on both sets of planes.

From the above considerations it is apparent that in the zinc-blende structures, crystals grown in the A direction will have the thermal stresses relieved largely by slip on A surfaces and vice versa. However with the stress distribution considered, slip occurs in such a way that on A planes α dislocations are introduced whilst β dislocations are introduced on B planes. It has been shown[28] that if conventional etches are employed only α dislocations intersecting A surfaces are delineated. Thus, if conventional etches were used for etch pit counts during crystal growing experiments, one would expect to observe higher densities in crystals grown in the A direction where α dislocations are favoured. Etch pit counts under such circumstances do not necessarily indicate a higher density of dislocations in A grown crystals. Nonetheless, because α and β dislocations are different, a difference in dislocation density might be expected even if both types were counted. Van Bueren[219] has pointed out that the slip patterns found in germanium crystals can be best accounted for if it is assumed that the strain introduced by the temperature distribution is only partially relieved by slip. This is readily

acceptable in view of the rapid decrease in dislocation mobility with decreasing temperature and stress. If dislocations are generated at or near the crystal surface then their motion along the glide planes will eventually bring them into the core of the crystal. Under particular conditions of growth the number of dislocations which move into the crystal will depend on their mobility and on the activation energy which determines the temperature at which they become virtually immobile. Peissker et al[43] and Willoughby and Bell[44] have given substantial evidence of differences in mobility and activation energy for α and β dislocations. The α dislocations are appreciably more mobile and have a lower activation energy than β dislocations. These differences would suggest that the core dislocation density in crystals grown in the B direction should be lower than that in crystals grown in the A direction.

Finally the observed differences in mobility and activation energy between α and β dislocations must be considered. If it is supposed that the terminating atoms in an edge dislocation find themselves in a situation similar to that of the triply bonded atoms in an A $\{111\}$ or B $\{111\}$ surface, the greater strain in the indium (sp^2) bond will give rise to a greater mobility and lower activation energy. The same conclusion is reached if rearrangement in the dislocation core to reduce the number of dangling bonds is considered. Any such rearrangement would involve greater deviations from the (sp^2) planar configuration of indium than for the p^3 arrangement of antimony and thus the β dislocations are more likely to rearrange and become less mobile.

C H A P T E R 5

Indium Antimonide p^+n Diodes.

5.1. Properties of Materials Used.

Before presenting the results of electrical measurements performed, as described in Chapter 3, on diodes fabricated by the techniques detailed in the same chapter; it is necessary to describe the properties of the InSb from which the dice used for diode fabrication were obtained. The properties of the regrowth regions of the diodes will also be assessed since it is not possible to fully interpret electrical measurements on p-n junctions unless the properties of both p and n regions are known.

5.1.1. Intrinsic Carrier Concentration.

It is clear from an examination of the theory given in Chapter 2 that the intrinsic carrier concentration is an important parameter in the electrical characteristics of a p-n junction. Low level diffusion currents (equation 2.15) are dependent on the minority carrier concentrations, which are related directly to the square of the intrinsic concentration (equation 2.11b) if the material is extrinsic. High level diffusion (eqn. 2.17 and 2.18) and recombination (equation 2.20) currents depend directly on the intrinsic carrier concentration. A knowledge of the intrinsic carrier concentration is necessary, therefore, before any detailed interpretation of electrical measurements can be made.

The intrinsic carrier concentration for a semiconductor with a single maximum in the valence band and a single minimum in the conduction band can be expressed as[15]:

$$n_i = 2\left(\frac{2\pi kT}{h^2}\right)^{3/2} (m_e m_h)^{\frac{3}{4}} \exp-\left(\frac{E_g}{2kT}\right) \quad \ldots\ldots (5.1)$$

If the energy gap varies linearly with temperature, $E_g = E_{go} - \alpha T$ where α is a constant, equation (5.1) can be written as:

$$n_i = K_i T^{3/2} \exp - \left(\frac{E_{go}}{2kT}\right) \quad \ldots\ldots (5.2)$$

where $K_i = 2\left(\frac{2\pi k}{h^2}\right)^{3/2} (m_e m_h)^{\frac{3}{4}} \exp\left(\frac{\alpha}{2k}\right)$

Measurements of intrinsic concentration in InSb are found to obey equation (5.2) and there is good agreement among workers[18][60] that, for temperatures above 150°K,

$$n_i = 5.7 \times 10^{14} T^{3/2} \exp - (0.125/kT) \quad \ldots\ldots (5.3)$$

Since the temperatures at which useful measurements can be made on most InSb diodes lie below 150°K, it is necessary to obtain values of intrinsic concentration for temperatures at which its measurement is not yet possible. It is common practice to simply extrapolate the experimental expression of equation (5.3) assuming that it applies at temperatures below 150°K. However, because of the critical dependence of electrical characteristics on the intrinsic carrier concentration, a more careful extrapolation was required for this work.

The optical measurements[59] of energy gap in InSb have shown that its variation with temperature is essentially linear down to about 150°K. Thereafter it cannot be considered linear and for temperatures above 20°K the energy gap is given quite accurately by:

$$E_g = 0.25 - (2.9 \times 10^{-4})T - \frac{0.32}{T}$$

Thus equation (5.3) cannot be used at temperatures much lower than 150°K. However, since InSb appears to have a direct energy gap, it would seem reasonable to assume that more accurate values of n_i could be obtained by substituting the measured values of the optical energy gap into equation (5.1). There is considerable justification for this assumption since the value of E_{go} from equation (5.3) agrees well with the extrapolated value obtained from optical absorption measurements and, if the most reliable values of the effective masses and the value of α obtained from optical measurements are substituted into equation (5.2), the value of K_i is in good agreement with the experimental value in equation (5.3).

For these reasons the values of n_i used in this work were calculated from equation (5.1) using the optically measured values of the energy gap.[59] The values of n_i as a function of reciprocal temperature are plotted in Figure 43. For comparison the values obtained from equation (5.3) are also shown. It can be seen that these values differ by about 30% at 77°K.

Equation (5.1) is often written[164] as:

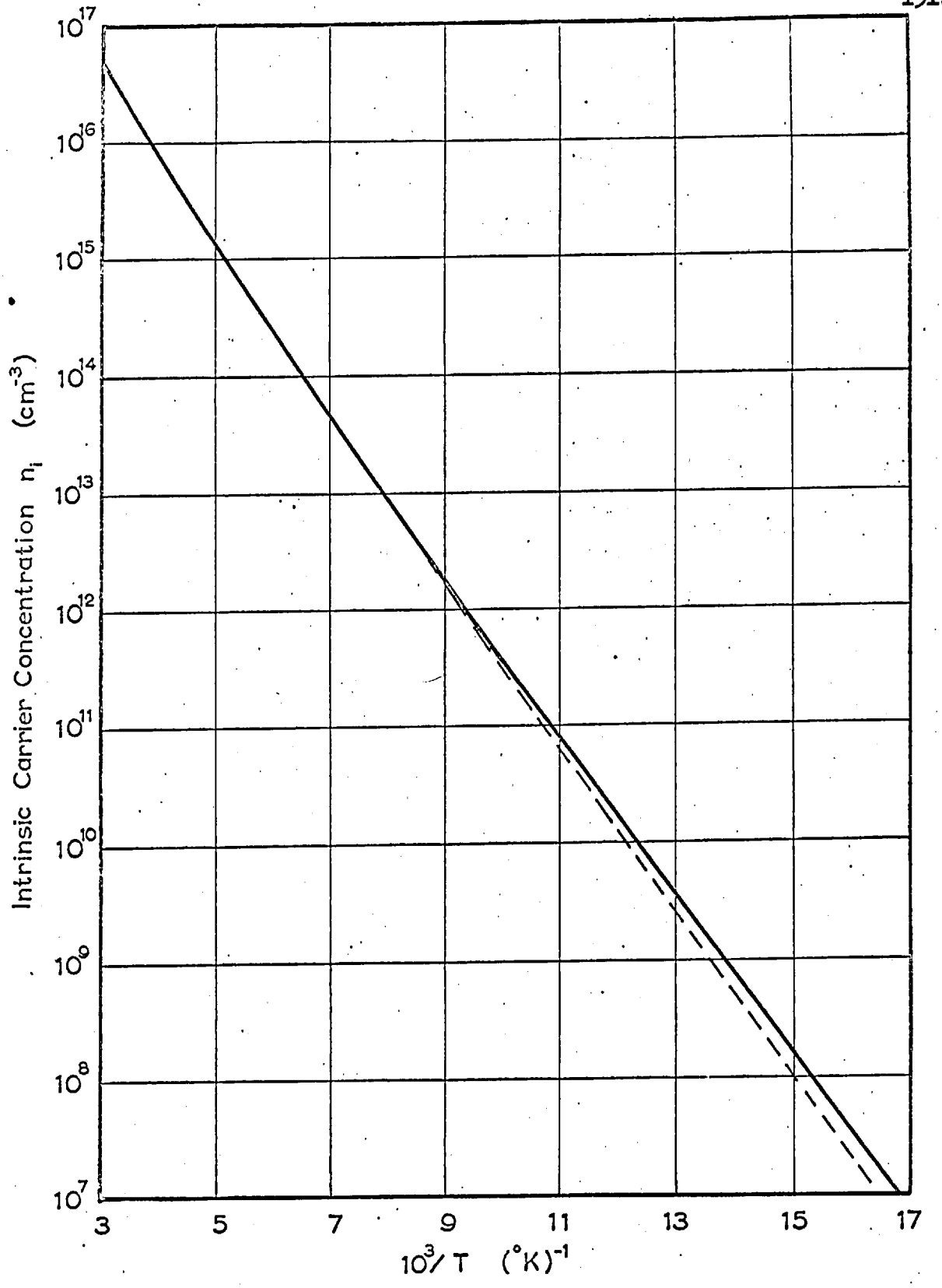


Figure 43. Intrinsic carrier concentration in InSb corrected for the non-linear temperature variation of the energy gap. The dashed curve results if a linear temperature dependence is assumed.

$$n_i = (N_c N_v)^{\frac{1}{2}} \exp(-E_g/2kT) \quad \ldots\ldots (5.4)$$

where $N_c = 2(2\pi\, m_e kT/h^2)^{3/2}$ and $N_v = 2(2\pi\, m_h kT/h^2)^{3/2}$

N_c and N_v are the effective density of states in the conduction band and valence band respectively. Since concentrations of electrons or holes which approach N_c or N_v imply the onset of degeneracy, it is useful to know the values of N_c and N_v as functions of temperature. These are plotted in Figure 44 as are the carrier concentrations required to cause the Fermi level to move to within 2kT of either band edge.

5.1.2. Properties of the Base Crystals.

As mentioned in Chapter 3, standard Hall and conductivity measurements were performed on most of the materials used in this work to determine the carrier concentrations and mobilities, and to estimate the degree of compensation. In general, these measurements were done on samples taken from the same slice as that used for dice in making diodes. Most measurements were done over the temperature range from 77°K to 300°K, although in a number of cases temperatures as high as 400°K and as low as 63°K were used. The results of these measurements are plotted in Figures 45, 46 and 47.

For a further estimation of the nature and perfection of the crystals, a slice of each was etched briefly and an etch pit count was made. From these measurements it was determined that for most crystals the dislocation density was less than $10^4 cm^{-2}$ and sometimes less than $10^3 cm^{-2}$. In crystal MCPB the dislocation density

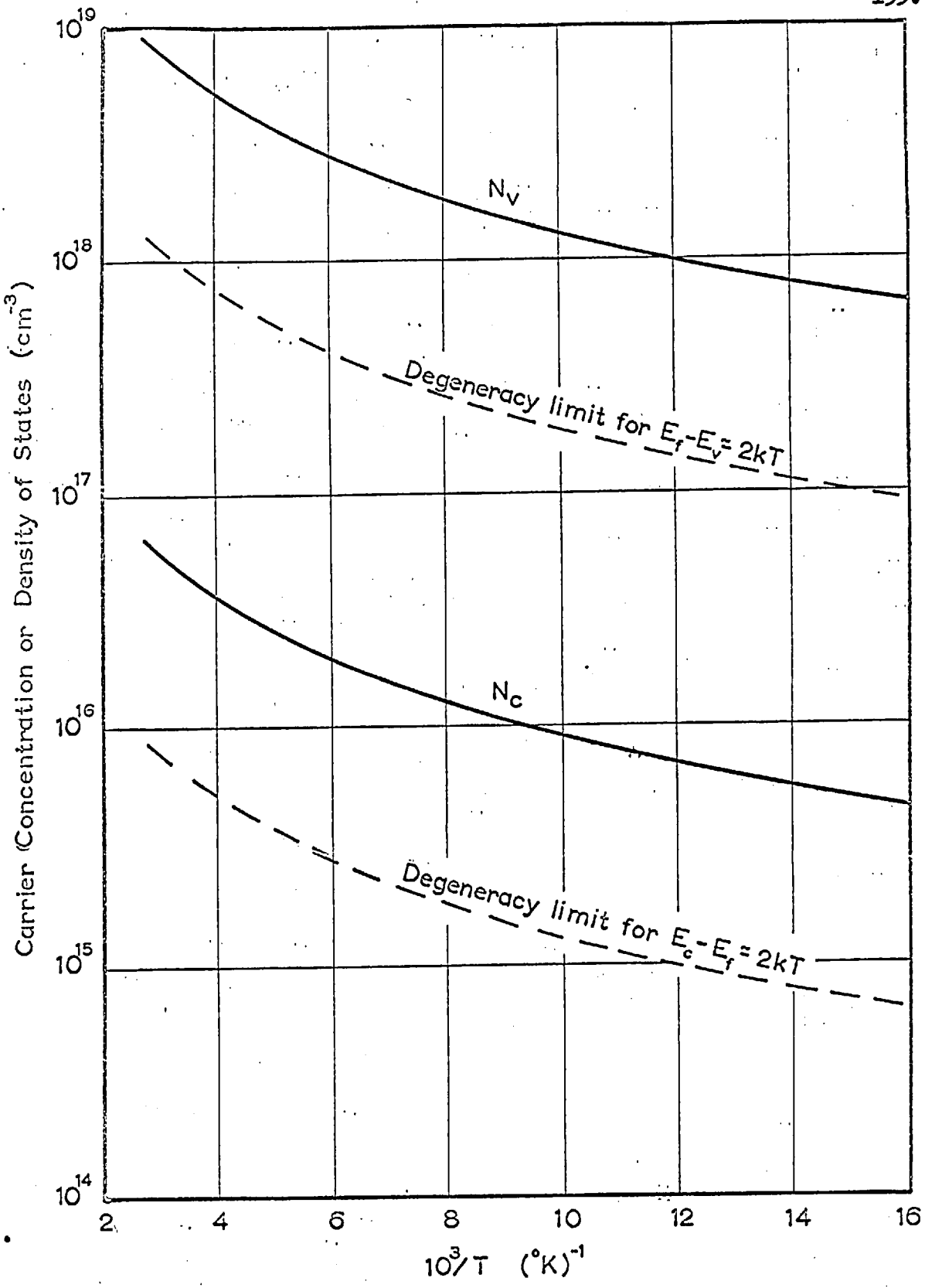


Figure 44. The effective density of states in the conduction band N_c and the valence band N_v as a function of temperature.

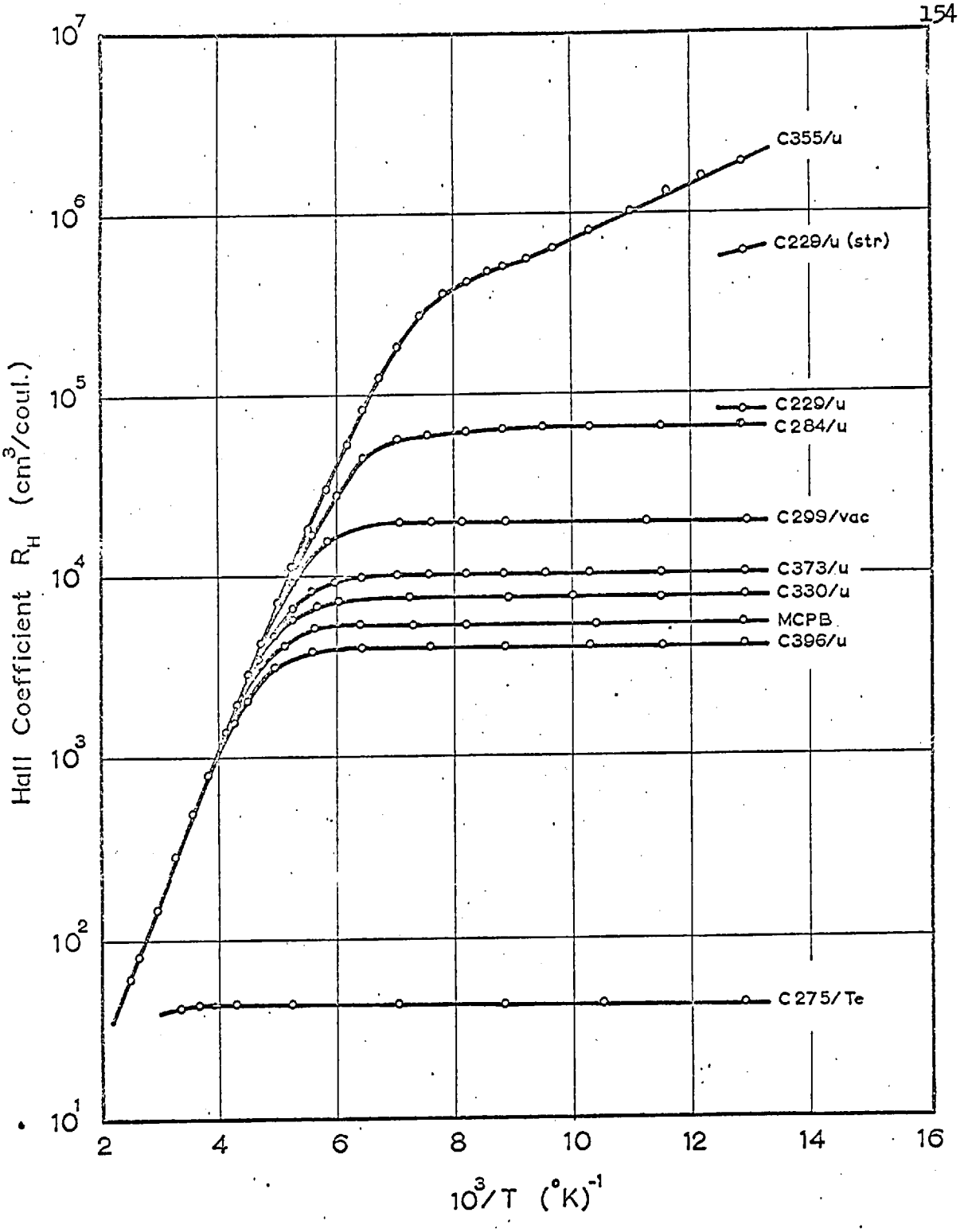


Figure 45. Measured values of Hall coefficient as a function of temperature on InSb crystals used for diode fabrication.

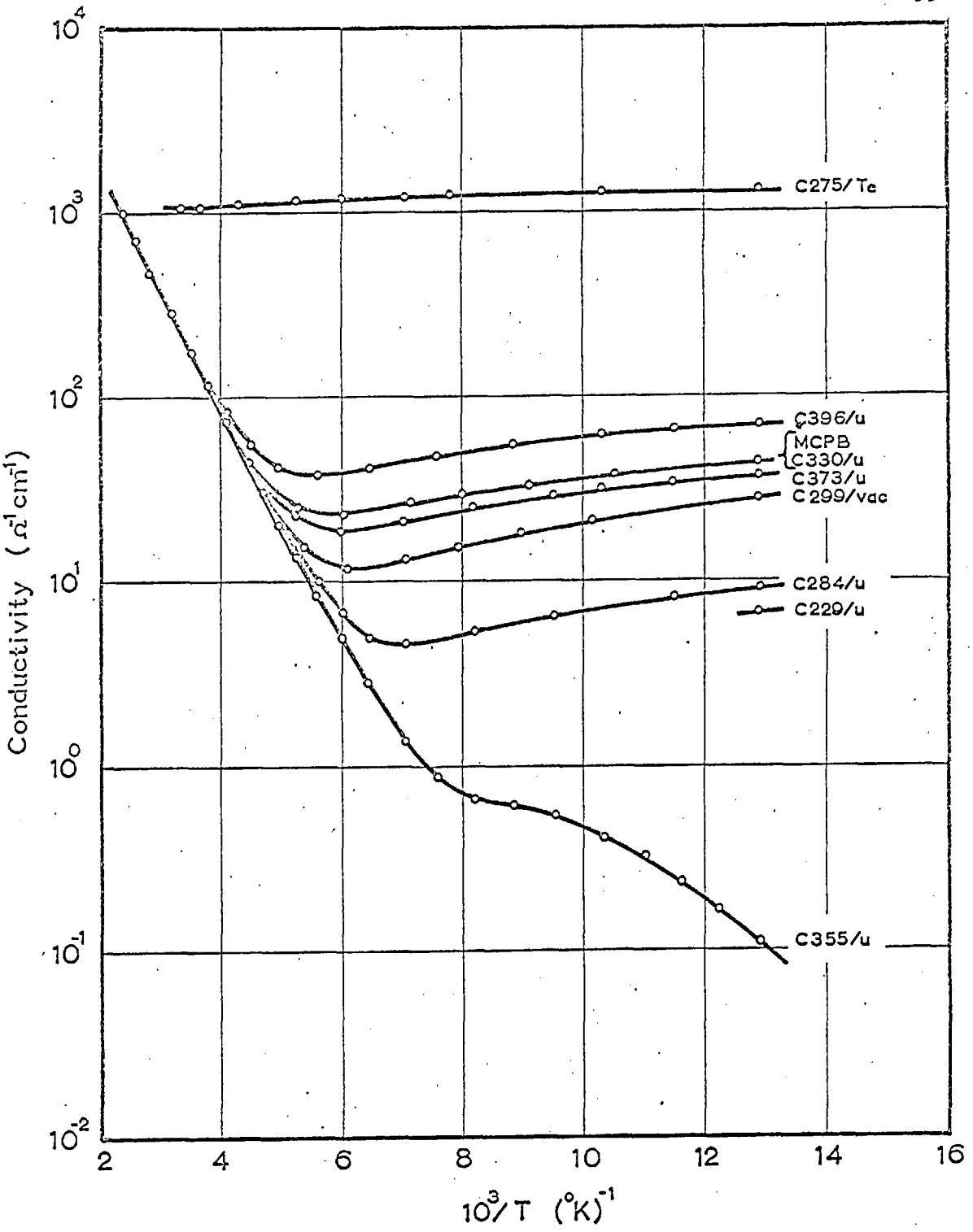


Figure 46. Measured conductivity of the InSb crystals used for diode fabrication.

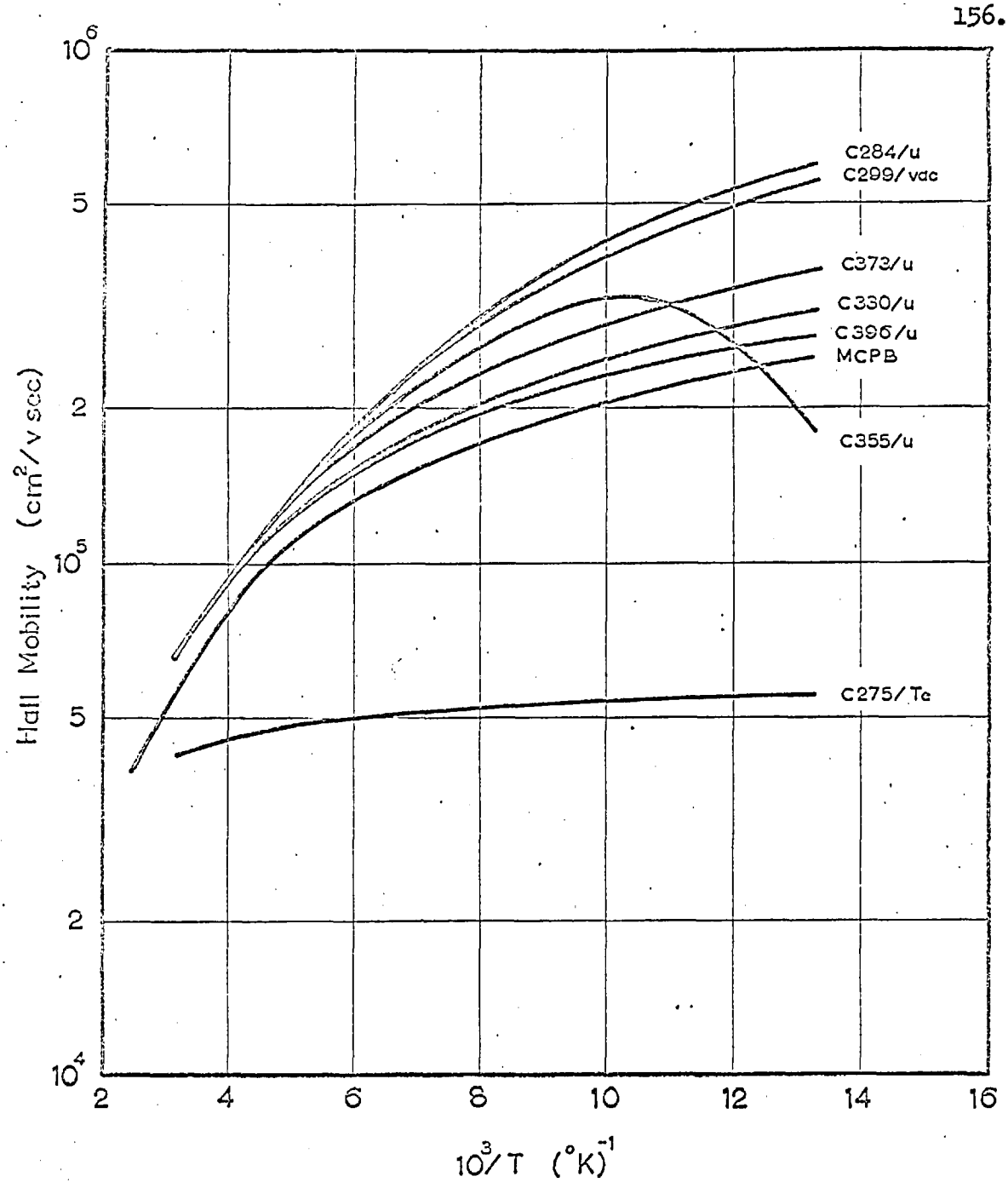


Figure 47. Hall mobility calculated from Figures 45 and 46.

was not uniform but on average it was about $10^5 cm^{-2}$. Crystal C229/u was purposely strained to introduce a high dislocation density for the experiments described in section 5.6. The dislocation density in this strained sample was quite uniform at about $3.6 \times 10^6 cm^{-2}$.

The electrical measurements and the etch pit counts indicated that most of the crystals were highly perfect single crystals with a minimum concentration of compensating impurities. This latter observation was obtained by comparing the mobilities of Fig.47 with the best mobilities reported[18] for InSb single crystals having similar impurity densities. Some exceptions were crystals MCPB and the strained samples of C229/u already discussed. Crystal C355/u is also an exception. The slope in the curve of Hall coefficient indicates the presence of an energy level either acceptor or donor, deep in the energy gap. The fact that the mobility drops off with decreasing temperature at relatively high temperatures, indicates the presence of a compensating impurity which is either present at relatively high concentration or a very effective scattering centre for electrons. Since most pure, undoped, uncompensated InSb crystals are about 10^{14} n-type[3], it is probable that the compensating impurity in this case is a low vapor pressure p-type impurity present at a concentration somewhat less than $10^{14} cm^{-3}$. Measurements[220] on p-type samples doped with copper have indicated that copper introduces deep levels in InSb and that it acts effectively to scatter electrons. The drop of electron mobility observed in InSb doped with copper at $10^{15} cm^{-3}$ compared to crystal C355/u makes copper contamination

during the growth of this crystal a plausible explanation for its electrical properties. Crystal C396/u, although giving apparently normal bulk electrical properties and low dislocation density, did not produce good diodes and the metallurgical properties were anomalous as discussed in section 4.1. An examination of the growth details of this crystal have shown that it was quite possibly contaminated with sulphur.

5.1.3. Diffusion Potential and Properties of the Regrowth.

In a highly asymmetric p-n junction it is not generally important to know anything about the heavily doped region since the electrical characteristics at most current levels are determined by the properties of the lightly doped region. However, at high current levels where the minority carrier diffusion current in the highly doped region can become significant, (see equations 2.17 and 2.18) it is necessary to know something of the properties of the highly doped region. Also the diffusion potential, which determines the space charge region width and therefore affects the capacity and the recombination currents at low forward and reverse voltages, can only be determined accurately if the impurity concentrations of both regions are known.

In these experiments pellets of indium-1% cadmium were used for alloying. If the effective value for the cadmium segregation coefficient[3] on a {111} surface is used to calculate the cadmium concentration in the regrowth region, a value of $5 \times 10^{19} cm^{-3}$ is obtained. Since the segregation coefficient is over three times smaller off a {111} surface and the solid solubility of cadmium in InSb as determined by diffusion measurements[31] is about $5 \times 10^{19} cm^{-3}$,

$5 \times 10^{19} cm^{-3}$ must be considered the upper limit of the possible impurity concentration in the regrowth. The lower limit for impurity concentration must be greater than $2 \times 10^{16} cm^{-3}$ since delineation of p-n junctions by etching, as illustrated in Chapter 4, depends on the regrowth being extrinsic at room temperature. The lower limit of possible impurity concentration in the regrowth was further narrowed by experiments performed on crystal C275/Te. This crystal was tellurium doped to $2 \times 10^{17} cm^{-3}$ n-type and was, therefore, degenerate. (see Figure 44) Junctions were formed on this material using indium 1% cadmium pellets. The electrical characteristics of these junctions (Figure 48) exhibited a very considerable amount of tunneling and thus it must be concluded that the p-type region was also degenerate. For this to be true the impurity concentration must be greater than $5 \times 10^{17} cm^{-3}$. (See Figure 44)

Values of the diffusion potential, ψ_D, at 77°K were calculated for various values of impurity concentrations in the p and n regions of a diode using Fermi statistics. Figure 49 shows the variation of ψ_D for values of n between $10^{13} cm^{-3}$ and $10^{16} cm^{-3}$, and for p = 10^{18} 5×10^{18} and $10^{19} cm^{-3}$. It is apparent that the diffusion potential depends on the impurity concentrations in the two regions and thus, if the diffusion potential ψ_D could be measured in a diode where n is known, p could be determined.

There are two methods sometimes used to measure ψ_D. One is to measure junction capacity as a function of reverse bias and find the value of ψ_D which makes equation (2.31) fit the results. Serious errors can result from this type of measurement as pointed out by Sparkes.[169] The other method requires current-voltage measurements

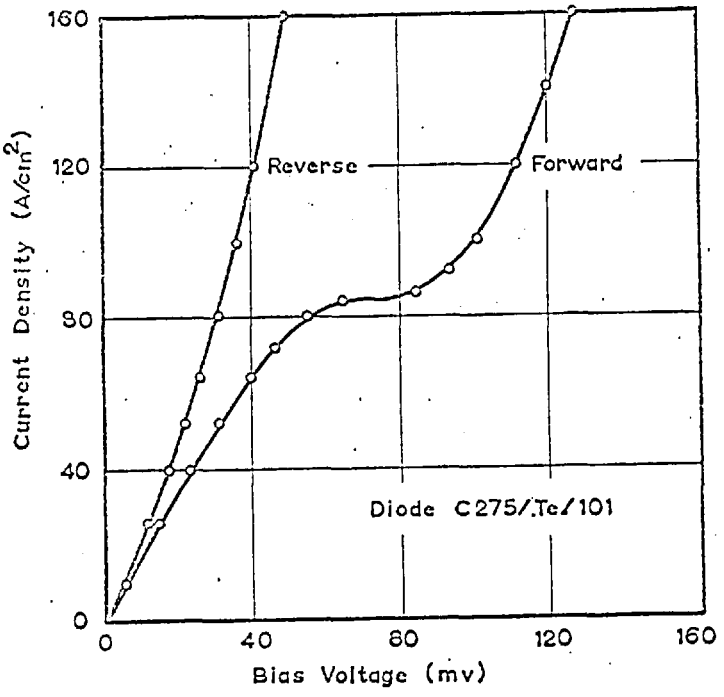


Figure 48. The tunneling characteristic at 77°K of a diode formed by alloying indium-1% cadmium to 2 x 10^{17} n-type InSb.

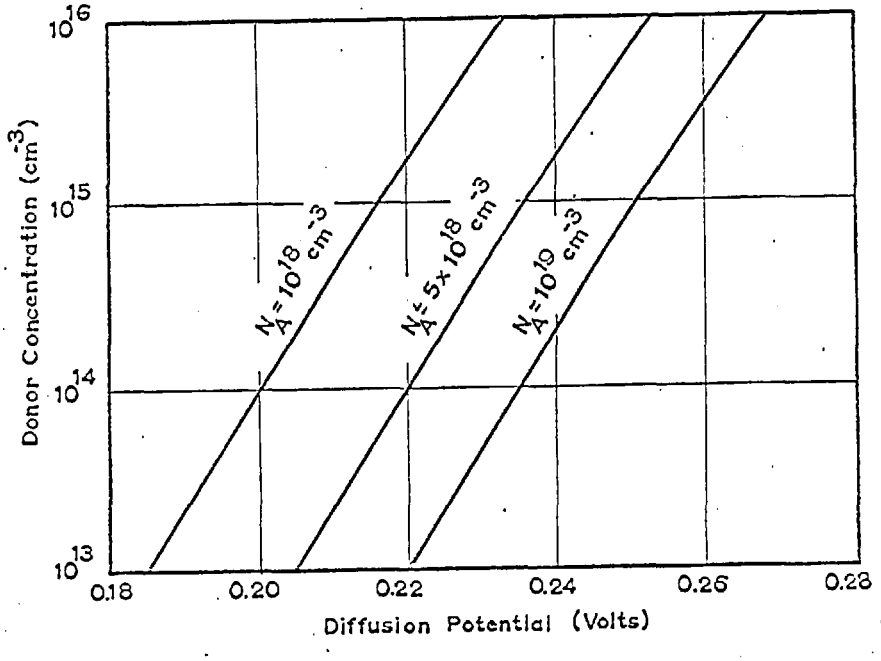


Figure 49. The variation of diffusion potential at 77°K in abrupt InSb p^+n junctions as a function of donor and acceptor concentrations.

at high currents. Its was shown in section 2.1.5. that the voltage, V_a applied to a diode is divided between the junction voltage, V, and the voltage drop due to the diode effective series resistance, R, so that $V_a = V + JRA$. Now if R is sufficiently small it will limit the diode current when $V \simeq \psi_D$ only, and, therefore the linear portion of the current-voltage characteristic at high currents can be extrapolated to zero current giving a voltage intercept approximately equal to ψ_D.

In order to obtain an estimate of the diffusion potential, measurements were made at high currents on many diodes. A typical measurement is shown in Figure 50(a) and the values of extrapolated zero current voltage for the same diode are plotted in Figure 50(b) as a function of temperature. The values of ψ_D calculated for $p = 5 \times 10^{18} cm^{-3}$ are also plotted on Figure 50(b) and it should be noted that for $p = 10^{18} cm^{-3}$ the calculated values of ψ_D would agree closely with the measured values. However, the intercept voltages obtained in the manner described above are always less than ψ_D and can easily be as much as 20% too low.(221) The impurity concentration of the p-region cannot, therefore, be less than $10^{18} cm^{-3}$ and, on the basis of the previous discussion and of measurements such as those of Figure 50, an impurity concentration of about $5 \times 10^{18} cm^{-3}$ would appear to be the most probable value. This value is in reasonable agreement with the value of $10^{18} cm^{-3}$ estimated by Chynoweth et al(145) as the impurity concentration in regrowth regions formed using indium 0.1% cadmium.

The diode regrowth regions were also examined for perfection or any evidence of strain. Sections such as those in Figure 25 have shown extensive regrowth and no gross imperfections have been revealed.

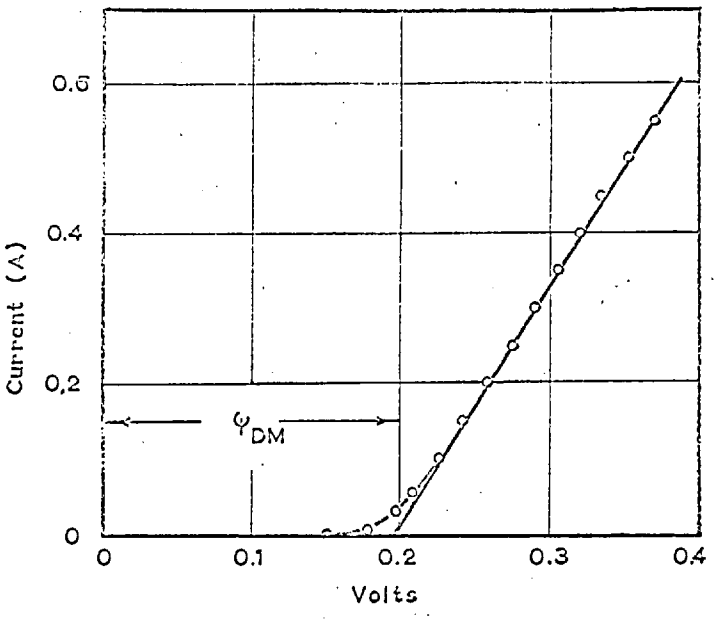


Figure 50 (a) The high current characteristic of an InSb diode showing the value of diffusion potential ψ_{DM} obtained from this measurement.

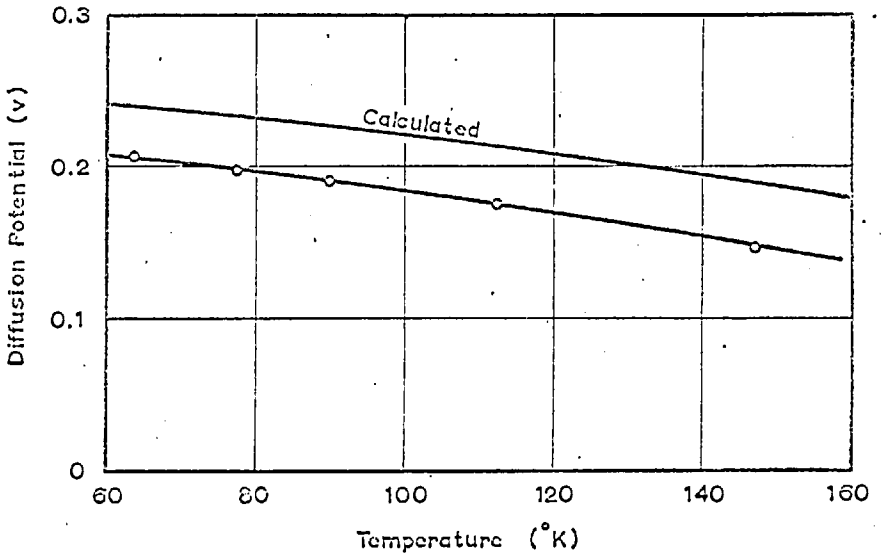


Figure 50 (b) The values of diffusion potential obtained from high current characteristics of an InSb diode compared to values calculated for that diode assuming $N_A = 5 \times 10^{18}$ cm^{-3}.

However, with the relatively large concentration of cadmium incorporated into the InSb lattice, it would not be surprising if strain and hence dislocations were introduced into the regrowth regions. If this were the case it would be possible to observe etch pits in the regrowth regions of diodes formed on A surfaces. In only a few cases, such as that shown in Figure 51, have any dislocations obviously related to the regrowth been observed. In these cases the dislocations may have been introduced for other reasons than the presence of cadmium. It must be concluded therefore that the regrowth regions were, in general, good quality, single crystal InSb having an acceptor concentration of about $5 \times 10^{18} cm^{-3}$.

5.2. Abrupt Junction and Space Charge Region Approximations.

The results to be presented in the following sections cannot be analysed using the theories presented in Chapter 2 unless the assumptions not justified in that treatment are shown to be valid. These are assumptions (4), (5) and (6) of Chapter 2 which state that the p to n transition is abrupt, that the boundaries between the space charge region and the neutral regions are well defined and that the quasi-Fermi potentials can be considered constant through the space charge region.

A junction may be considered abrupt if the transition from p-type to n-type occurs entirely within the space charge region. Since this must remain the case in forward bias, the transition width, W_t, must be several times shorter than the zero-bias width of the space charge region. If the space charge region is assumed to be well defined its width is given by equation (2.19). Values of the zero-bias width calculated from equation (2.19) for InSb junctions, where $N_A = 5 \times 10^{18} cm^{-3}$, are

Figure 51. Dislocation etch pits in a junction regrowth region. The large pit was etched up before alloying (x1000).

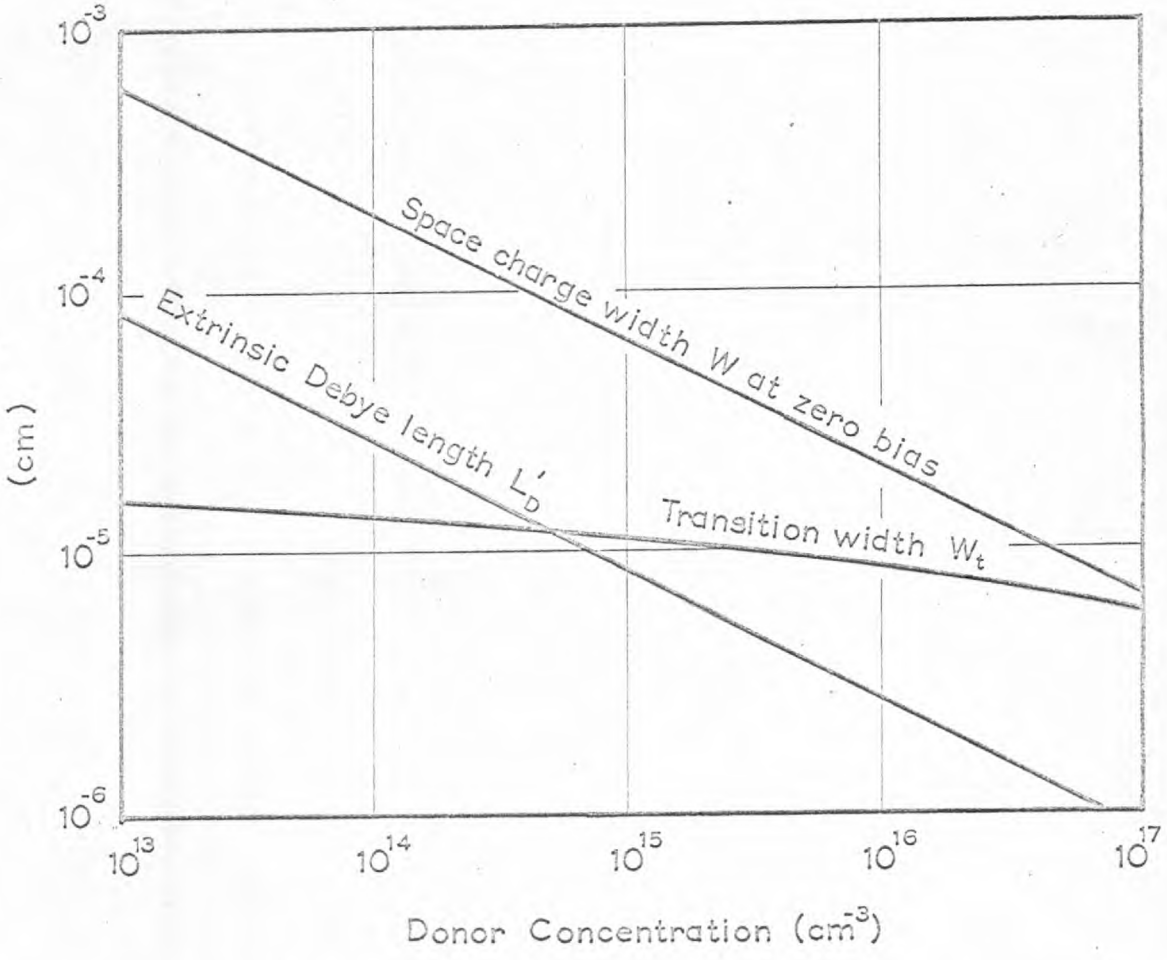


Figure 52. InSb p^+n junction ($N_A = 5 \times 10^{18} cm^{-3}$) zero bias space charge width compared to the extrinsic Debye length and the transition width from p to n.

plotted in Figure 52 as a function of carrier concentration in the n-region.

In an alloyed junction the p to n transition would be abrupt by any criterion if it were not for diffusion. To examine the nature of the transition, therefore, the diffusion profile of the alloying impurity (in this case cadmium) must be determined. For this purpose the following assumptions were made: the interfacial cadmium concentration in the solid is given by the solid solubility (i.e. $5 \times 10^{19} cm^{-3}$), the diffusion coefficient is that given by Orth and Watt[95], the diffusion gives a complimentary error function impurity distribution and the diffusion cycle is equivalent to 10 minutes at $550^{\circ}K$. Using these assumptions the transition width, defined as the distance in which p changes from $10^{18} cm^{-3}$ to one tenth of the donor concentration in the n-region, was calculated. The results are plotted in Figure 52 as a function of donor concentration. It is immediately clear that junctions formed on pure materials can be considered abrupt. However, since tunneling occurs only in very narrow junctions, the fact that tunneling was observed in diodes formed on crystal C275/Te, indicates that the junctions on impure materials are also abrupt. The reason for this is, as pointed out in Chapter 1, that diffusion in InSb does not give rise to a complementary error function profile. It gives a steeper profile and therefore the above calculations overestimate the transition width. Junctions formed on InSb by the alloying process described in Chapter 3 can thus be considered abrupt p-n junctions. This conclusion is further justified by the observations that diffused diodes behave as abrupt junctions[99] and that tunneling can be observed in InSb diodes formed by diffusion.[222]

The assumption (5), that in a p-n junction two neutral regions are separated by a depleted space charge region, neglects the sections which are neither neutral nor fully depleted. To establish the validity of this assumption it is necessary, therefore, to estimate the importance of these transition regions and to determine if the departures from neutrality in the outer regions are small.

Perturbation methods have been used to solve Poisson's equation for small departures from neutrality at the edges of the space charge region (170)(171). These solutions show that the potential, and therefore the space charge, varies exponentially with a characteristic length, L'_D, called the extrinsic Debye length.

$$L'_D = \left[\frac{kT}{q^2 \left((N_A - N_D)^2 + 4n_i^2 \right)^{\frac{1}{2}}} \right]^{\frac{1}{2}} \qquad \ldots . \quad (5.5)$$

Numerical treatments[170] have shown that the change from complete neutrality to complete depletion occurs effectively in five or six extrinsic Debye lengths. It is also generally accepted that, if the width of the space charge region as given by equation (2.19) is greater than five or six extrinsic Debye lengths, the abrupt space charge region approximation is valid. Values of L'_D for n-type InSb are plotted as a function of impurity density in Figure 52. Comparison with the corresponding widths of the space charge region shows that the assumption of an abrupt depletion layer is valid in InSb at least for all values of reverse bias. It is possible to visualize forward bias conditions where the assumption would no longer be correct because of the decreased width of the space charge region. However, for diffusion currents in forward bias this does not affect the accuracy of the theory since the

abrupt approximation was required only to establish a "zero" in the co-ordinates of the junction. The theory for recombination currents involves the width of the space charge region. However, the effect of the space charge width is very small in comparison to the other voltage dependent terms and recombination currents are only significant at low forward bias so the assumption is also justified in this case. In general, measurements of junction capacity are made in reverse bias only so the simple theory for junction capacity, which depends directly on the abrupt space charge assumption, is also acceptable.

Sah et al[158] have examined the conditions of neutrality in the regions bordering the space charge region. They have shown that in conditions of high level injection Poisson's equation in these regions becomes:

$$\frac{d^2\psi}{dx^2} = \frac{q}{\varepsilon}(p-n+N_D-N_A) \simeq \frac{kT}{2qL_o^2}$$

where L_o is the diffusion length of injected carriers. In n-type InSb L_o is approximately 10^{-3}cm. Thus, at 77°K, $\frac{d^2\psi}{dx^2} = 7 \times 10^3 V/cm^2$ and $(p-n+N_D) = 6 \times 10^{10} cm^{-3}$.

Since in general $N_D > 10^{13} cm^{-3}$, the departure from neutrality is small in comparison to complete depletion and the neutrality assumption is acceptable.

Finally assumption (6), that in forward bias the quasi-Fermi levels in the space charge region can be considered constant, must be examined. It has been shown[171] that the variation in quasi-Fermi level can be expressed as, for example:

$$\Delta\psi_p = \frac{q}{kT} \ln \left(1 - \frac{kT}{q\Delta\psi} \cdot \frac{W}{L_o}\right)$$

where $\Delta\psi_p$ is, in this case, the change in the quasi-Fermi level for holes and $\Delta\psi$ is the change in electrostatic potential in the junction. In InSb at 77°K, even at high forward bias, $\Delta\psi > \frac{kT}{q}$ and, even in an extreme case of zero bias for a junction on say $10^{13}cm^{-3}$ material, W is at least fifty times less than L_o. Thus the change of quasi-Fermi level in the space charge region will at most be a small fraction of a millivolt and the assumption is valid.

It is worth pointing out in connection with this brief treatment of these assumptions that InSb junctions at 77°K are in many respects analagous to silicon junctions at room temperature and the above assumptions have been shown to be valid or acceptable[158][171] for most p-n junctions in silicon. In germanium where some of the assumptions have doubtful validity, theories like those presented in Chapter 2 are, in general, applied without reference to these assumptions.

5.3. Direct Current Characteristics.

5.3.1. Forward Conduction.

Diodes were made on all the materials described in section 5.1.2. These diodes were prepared, etched and the electrical characteristics were measured as described in Chapter 3. The forward bias currents at 77°K of diodes formed on InSb having donor concentrations ranging from $3.3 \times 10^{12}cm^{-3}$ to $1.6 \times 10^{15}cm^{-3}$ are shown in Figure 53. For the sake of clarity, the measured values shown are only a fraction of those obtained from one diode for each base impurity concentration. The

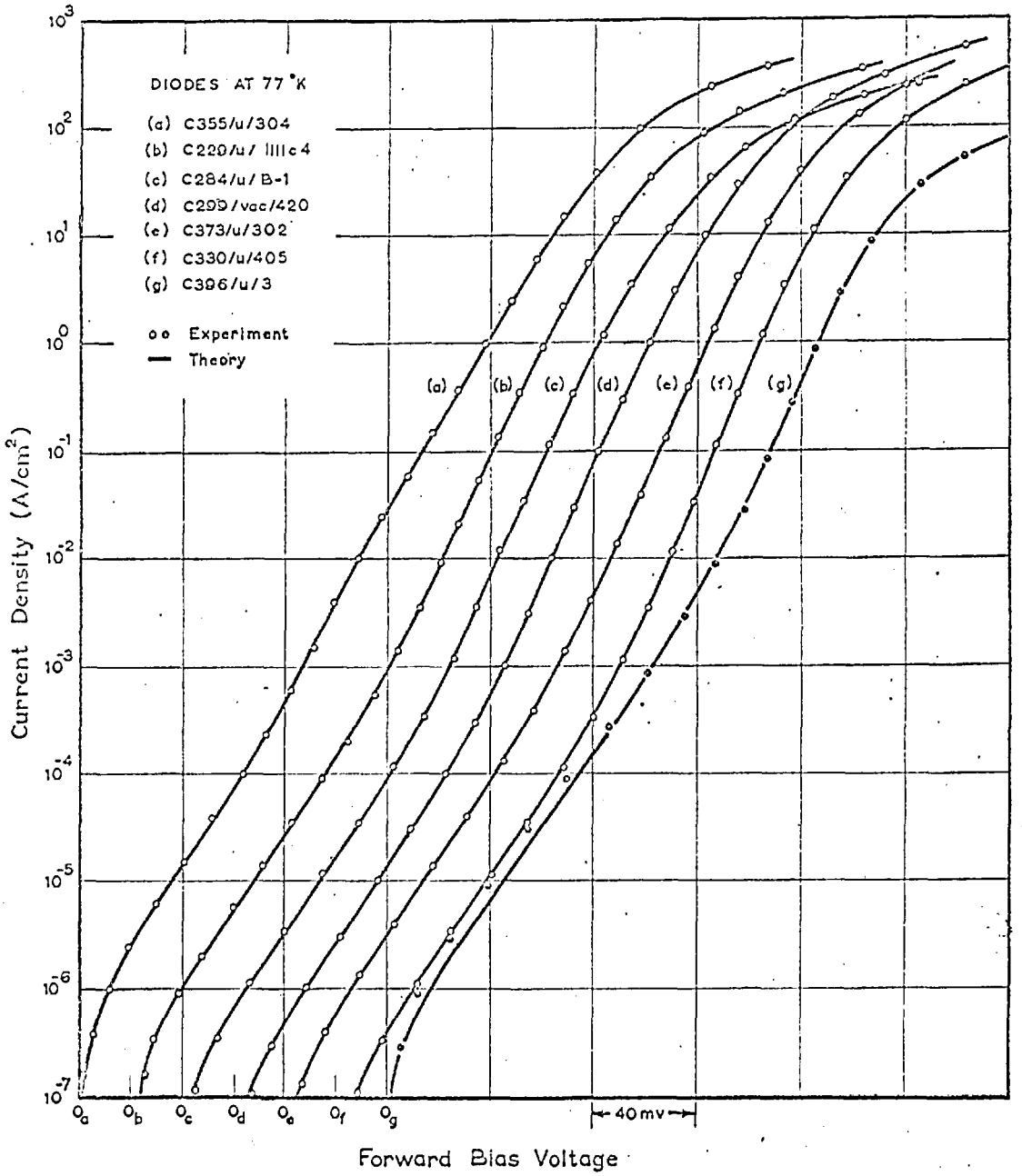


Figure 53. The forward current-voltage characteristics at 77°K of diodes formed on InSb crystals having donor concentrations ranging from (a) $3.3 \times 10^{12} cm^{-3}$ to (g) $1.6 \times 10^{15} cm^{-3}$. The solid curves are the theoretical values given by equations (2.22) and (2.23).

theory of Chapter 2 was applied by fitting equations(2.22) and (2.23) to the measurements. To do this values of the coefficients J_{Ln}, J_{rgo}, J_{Lp} and R were chosen to give a good fit to the combined results of measurements on not less than three diodes on each material. A value of J_{Hp} was not required in fitting results on InSb since J_{Hp} can be related to J_{Lp} by n_i in a material where the electron mobility and lifetime are very different from those of the holes. The solid lines in Figure 53 are the values given by equations(2.22) and (2.23) when fitted in the manner described above. Current measurements, converted to current densities by the method described in Chapter 3, on as many as five different diodes formed on the same material have generally fallen within the line widths of the curves on Figure 53, except in the high current regions where the series resistance (Eqn.2.23) becomes the dominant factor. It is immediately clear that the form of the theory given in Chapter 2, and in particular the treatment of medium level conditions, accurately describes the observations on InSb diodes. Such close agreement between the form predicted by theory and that observed by experiment is most encouraging. However, the full validity of the theory is not established until it can be shown that the values of the coefficients arbitrarily chosen to give a good fit to the experimental results also agree with the values predicted by theory.

The coefficients J_{Ln} and J_{Lp} are both determined by the minority carrier concentrations, mobilities and lifetimes in the appropriate regions (see equation 2.23). The minority carrier

concentration and mobility cannot be determined directly and therefore some uncertainty must be present in any estimate of the theoretical values of J_{Ln} and J_{Lp}. To estimate J_{Lp}, the saturation value of hole diffusion current, values of lifetime obtained from the transient measurements (see section 5.4.2.) and values of diffusivity obtained from comparable purity p-type material were used. The minority carrier concentration was calculated using the intrinsic concentration given in section 5.1.1. Assuming $N_D = 3 \times 10^{15} cm^{-3}$, the value of J_{Lp} obtained was 5×10^{-11} A/cm^2. This value is in good agreement with the observed value. Further, assuming the concentration dependence of J_{Lp} to be determined by the change in minority carrier concentration only (i.e. lifetime and diffusivity assumed concentration independent), the dashed curve on Figure 54(a) was obtained. In view of the uncertainties described above and the inaccurate assumption of concentration independent diffusivity and lifetime, the agreement with the values of J_{Lp} required to fit the measurements is considered extremely good.

Now in a p-n junction J_{Ln}, the saturation value of electron diffusion current, can be observed only at high currents in diodes formed on very pure material. The value required to give a good fit on the purest material was 4×10^{-12}A/cm^2. If N_A is assumed to be $5 \times 10^{18} cm^{-3}$ as suggested in section 5.1.3., then the value of electron mobility at 77°K will be about 3×10^4 cm^2/V sec.[18] A reasonable value of lifetime in this heavily doped region is 5×10^{-12} sec. Using these values and the value of n_i given in section 5.1.1., a value of $J_{Ln} = 3 \times 10^{-12}$A/cm^2 was calculated.

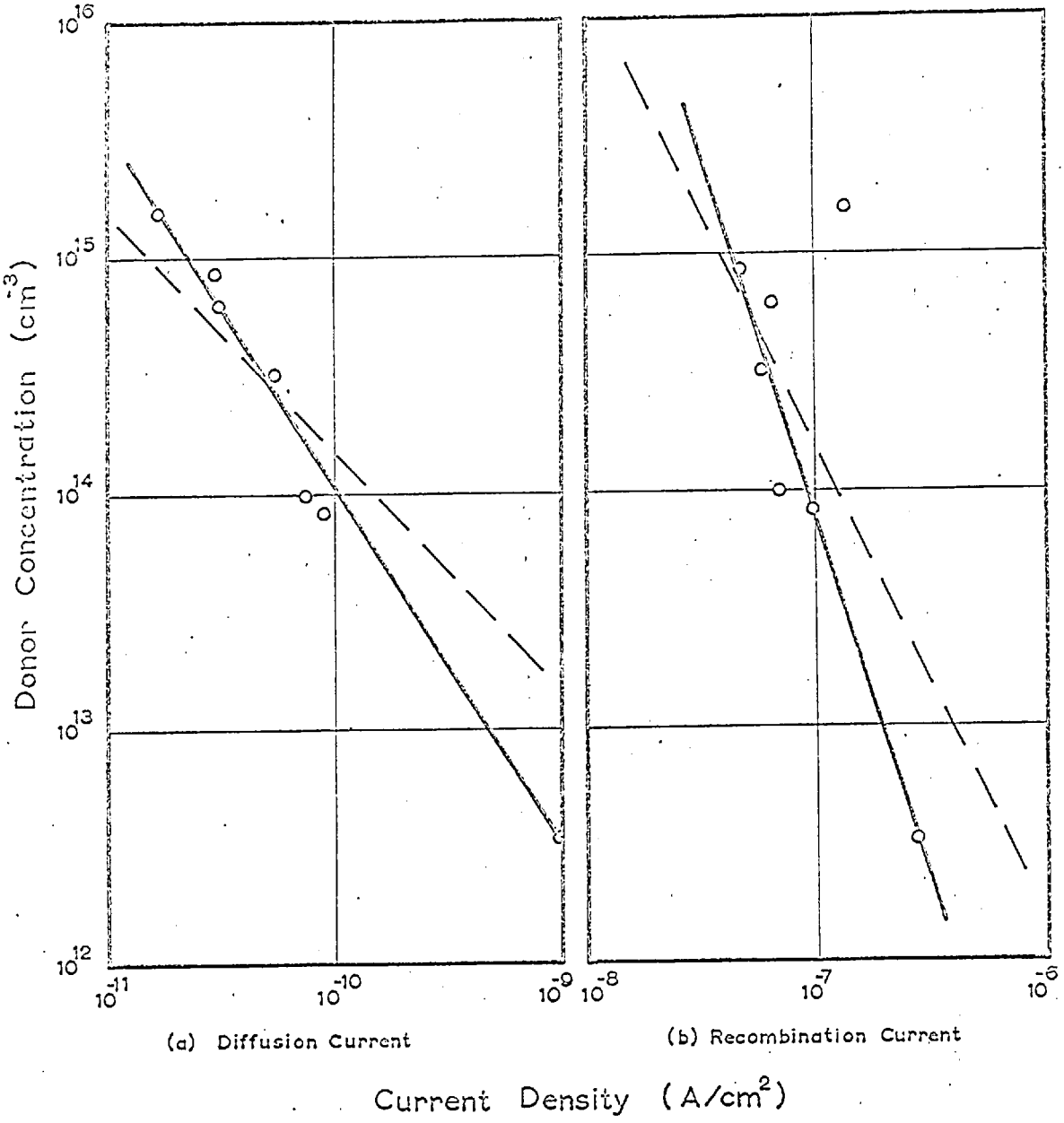


Figure 54. The values of (a) diffusion coefficient J_{Lp} and (b) recombination current coefficient J_{rgo} required to give the theoretical curves in Figure 53. The dashed lines are theoretical values for J_{Lp} and J_{rgo} assuming lifetimes and mobilities to be concentration independent.

Thus reasonable values for the bulk properties of the regrowth region give good agreement between the theoretical and observed values of J_{Ln}. Since J_{Ln} is determined by the properties of minority carriers in the regrowth region, it should be a constant for all the diodes in these experiments. The above value of J_{Ln} does give a good fit to the results on all the diodes formed on less pure material but since in these cases the effect of J_{Ln} is diminished, this cannot be considered a critical test of the accuracy of the value used.

A value of J_{rgo}, the coefficient for recombination current, was also calculated for a diode having a donor concentration of $3 \times 10^{14} cm^{-3}$. Using the value of n_i from section 5.1.1., hole lifetime $\tau_p = 5 \times 10^{-7}$ sec. (see section 5.4.2.) and electron lifetime[5] $\tau_n = 10^{-9}$ sec., the value of $J_{rgo} = 7 \times 10^{-8} A/cm^2$. This value is in close agreement with the measured value. If it is assumed that both hole and electron lifetimes are independent of impurity concentration, then J_{rgo} depends only on the square root of the donor concentration and its dependence on impurity concentration is given by the dashed line in Figure 54(b). It is apparent that, except for C396/u ($N_D = 1.6 \times 10^{15} cm^{-3}$), the values of J_{rgo} required to fit the measurements are also in good agreement with theory. If account is taken of the fact that carrier lifetimes increase as the impurity concentration decreases (see section 5.4.2) the slope of the dashed line in Figure 54(b) would be steeper and the agreement between theory and experiment would be even better.

It has already been pointed out that C396/u was a special case and although diodes on this material gave reproducible characteristics at high currents, the low current characteristics varied from diode to diode. The curve shown in Figure 53, from which the value of J_{rgo} in Figure 54(b) was obtained, was the best characteristic obtained. It should be noted that the theoretical fit to this measurement is not good in spite of the high value of J_{rgo} and therefore doubt must be placed on the origin of the currents at low bias in diodes formed on C396/u.

Finally calculations of R, the effective series resistance, were made in several instances using the known resistance of the cryostat and the conductivity and geometry of the diode wafer. These values were always in good agreement with those required to fit the measurements.

The forward characteristics of diodes were also measured as a function of temperature both to obtain information on the temperature behaviour of InSb diodes and to obtain a further check on the applicability of the theory of Chapter 2. An example of such a family of measurements is shown in Figure 55. Again for clairity, only a fraction of the measured points is indicated. Equations (2.22) and (2.23) were fitted to these results as described above and the solid lines indicate the values given by the theory. Since it has already been established that the values of the coefficients chosen for fitting at 77°K agree with theory, it is necessary in this case only to show that their temperature dependence agrees with that predicted by theory. The theoretical temperature dependence of J_{Lp} is determined almost entirely by that of the minority carrier concentration which is, from equation (2.11),

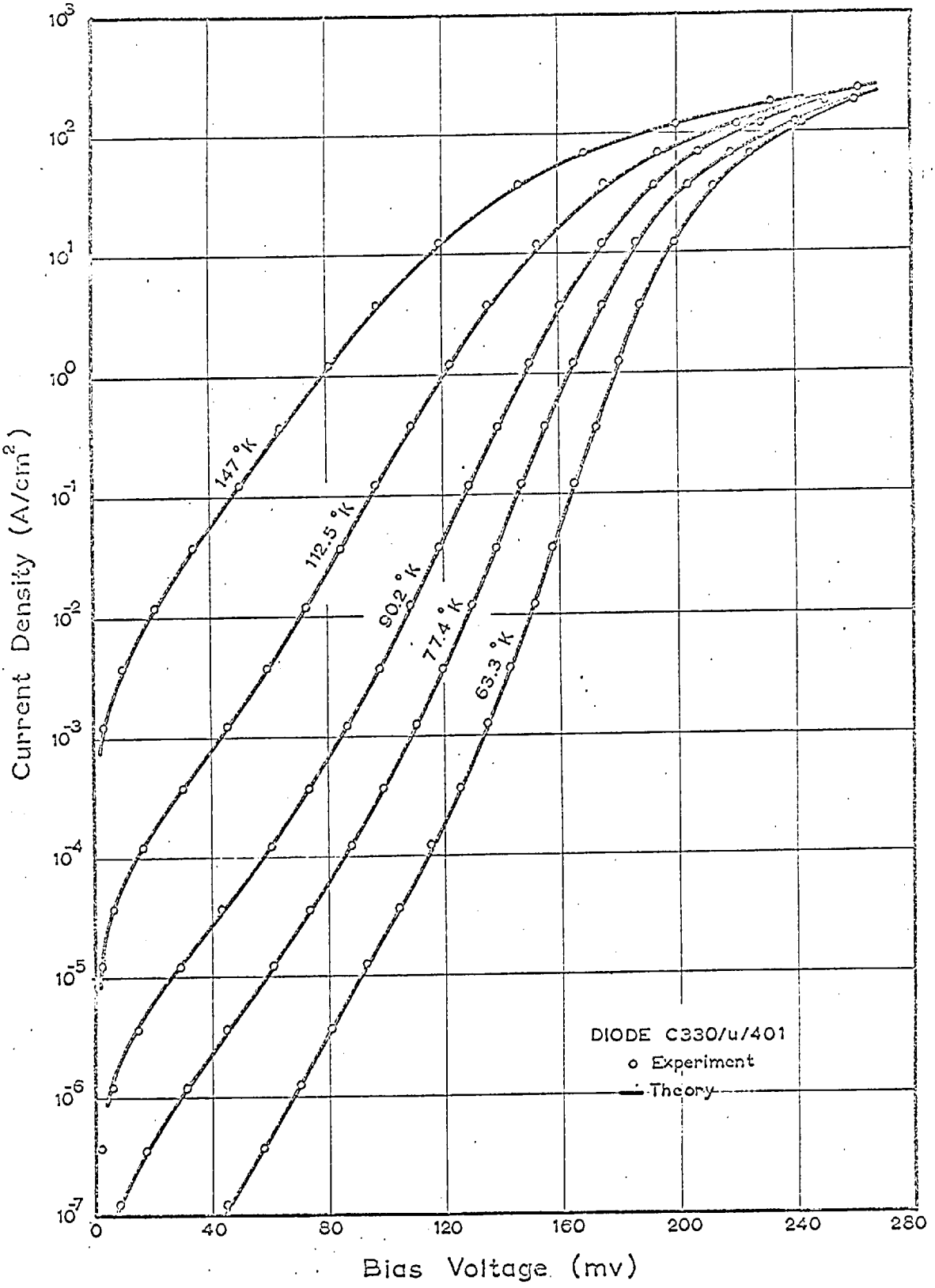


Figure 55. Typical forward current-voltage characteristics for an InSb p^+n diode as a function of temperature. Solid curves are given by equations (2.22) and (2.23).

that of n_i^2. Figure 56 shows the square root of the measured values of J_{Lp} plotted with n_i as a function of reciprocal temperature. The theoretical temperature dependence of J_{rgo} is determined almost entirely by that of the intrinsic concentration, n_i. Figure 56 also shows a comparison between the temperature dependence of the value of J_{rgo} required to fit the measurements and that of n_i. In both these cases the agreement of temperature dependence and therefore the agreement between theory and experiment is exceptionally good.

Experiments were also carried out to determine the effects of crystal orientation on the electrical properties of diodes. For this purpose junctions were prepared on the facet forming {100}, A {111} and B {111} surfaces. Although it was reported earlier (223) that a residual difference was always observed between the electrical characteristics of diodes formed on A{111} and B{111} surfaces, it has since been shown that this can be eliminated by careful chemical treatment. Diodes formed on {100} surfaces have yielded identical characteristics to those obtained from diodes formed on A {111} or B {111} surfaces of the same material. It should be pointed out, however, that failures to obtain the theoretical characteristics due to metallurgical defects were much more common on {100} and A {111} surfaces than on B {111} surfaces. In view of the discussion of Chapter 4 this is not surprising. Also chemical processing of diodes formed on B {111} surfaces was generally much more successful than that on A {111} surfaces. (see section 5.5.1.).

The forward current characteristics of InSb p-n junctions appear, therefore, to be accurately described both in their voltage and temperature dependence by the theory of Chapter 2. A recombination component having an $\exp(\frac{qV}{2kT})$ voltage dependence

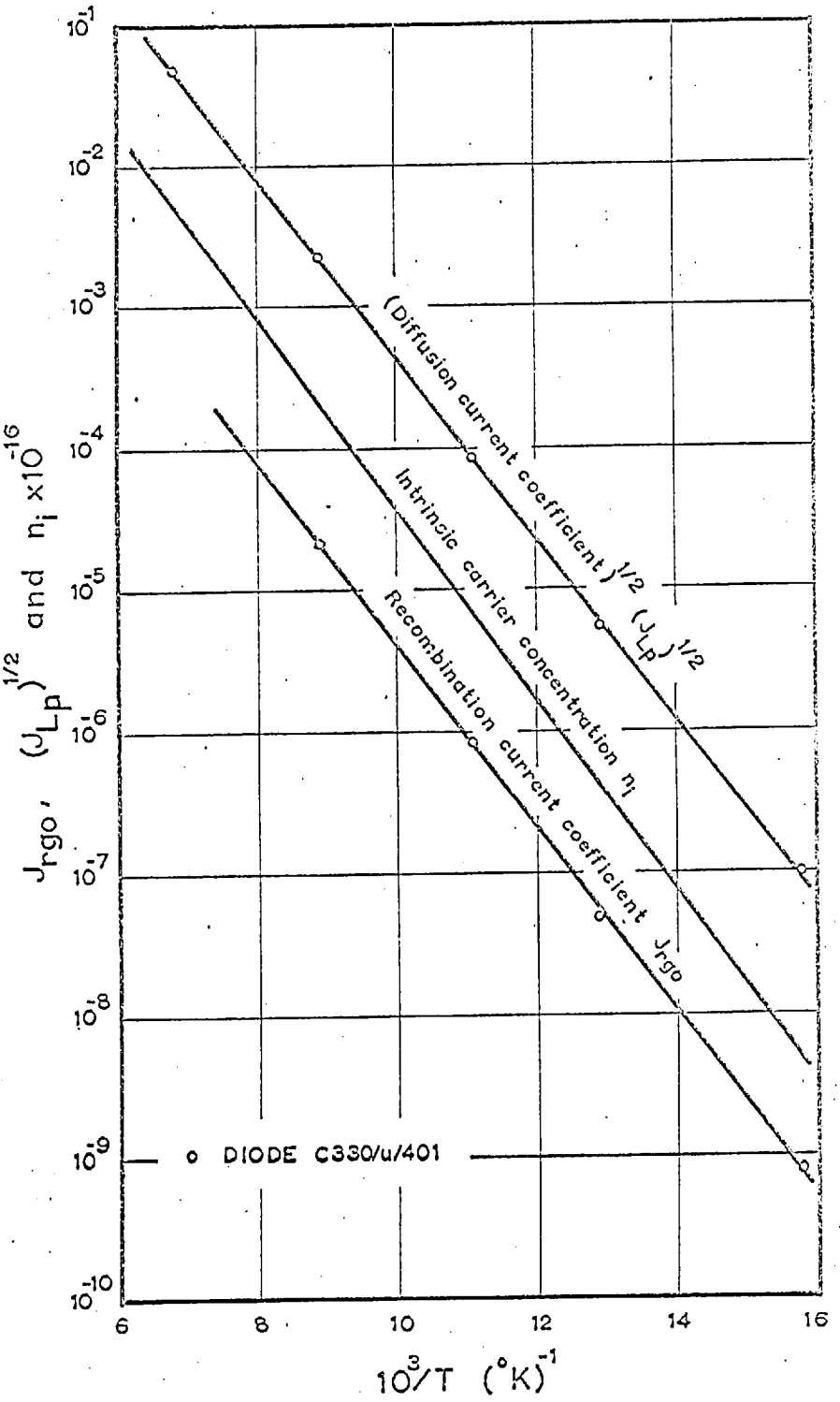


Figure 56. The temperature variation of the values of J_{Lp} and J_{rgo} required for the theoretical fits in Figure 55 compared to the temperature variation of n_i.

in the forward characteristics, indicates the presence of a trapping level near the centre of the energy gap. It is not possible to determine unequivocally from these measurements whether such a level is associated with the junction surface or is a bulk property. However, measurements on diodes where the ratio of junction area to junction periphery was varied by a factor of five showed no change in the current density characteristics. The agreement with theory also indicates the recombination component to be a bulk effect as do the reverse current measurements described in the following section.

5.3.2. Reverse Conduction and Breakdown.

The measurements of reverse characteristics of InSb p-n junctions were not as reproducible as the measurements of forward characteristics. This was ascribed to surface effects (see section 5.5.) and in particular to the influence of an anomalous hardening effect. Immediately after etching every diode exhibited a steep breakdown at relatively low voltages. If the current was increased this breakdown would move to higher voltages until it reached a limiting value. This process was called hardening. With little or no hardening the reverse currents at low bias were very small and generally very reproducible. These currents were used for measurements described later in this section. After hardening, the reverse currents at low bias were generally increased by less than a factor of five and varied from diode to diode and from etch to etch (for more detail see section 5.5.2) Figure 57 shows complete reverse characteristics for diodes on each of six different materials. The curves of Figure 57 are somewhat better

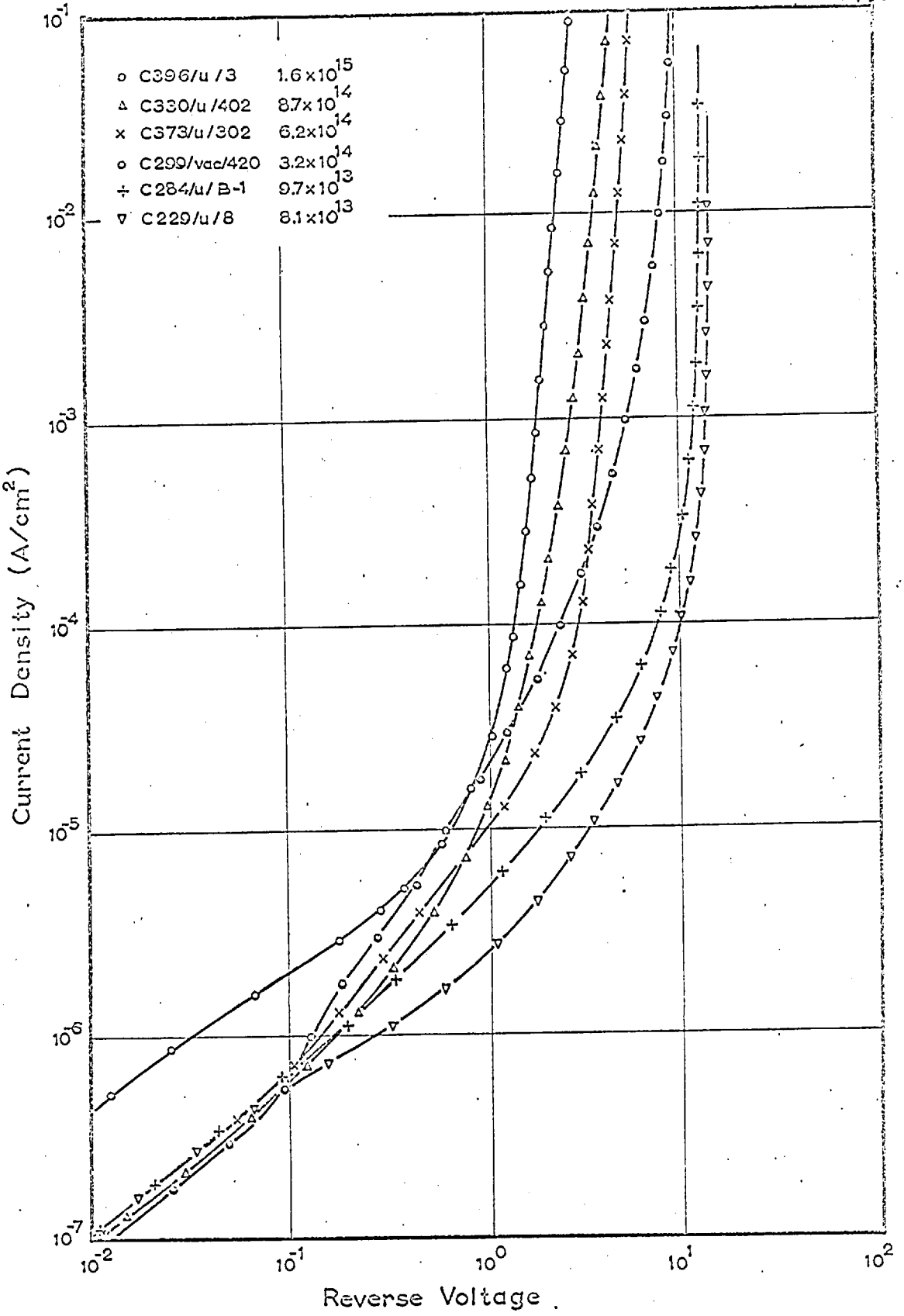


Figure 57. Reverse current-voltage characteristics of InSb p^+n junctions.

than typical in that a higher percentage of diodes often exhibited a change in the slope of the current curve at medium voltages, similar to that of the diode on C299/vac. The diodes on C596/u exhibited high currents at low reverse bias in a similar way to their behaviour at low forward bias. These reverse currents were always at least an order of magnitude greater than those expected from diodes on such material. Except for the diodes on C596/u, all diodes, after hardening, gave approximately the same values of current density at low voltages and this current quite closely obeyed a $V^{0.75}$ law. Neither of these observations correspond to any bulk theory and they are taken, therefore, to indicate the presence of a reproducible surface effect. Since the reverse currents, at least for low voltages, appeared to be dominated by surface effects, no attempt has been made to place any interpretation on these results. It should be noted however that the reverse currents observed after hardening, exemplified in Figure 57, were in general at least an order of magnitude lower than those of the best diodes previously reported in the literature.(159)

Although the hardened reverse currents at low bias were variable and seemed to have no dependence on the diode material, the high current breakdown region was quite reproducible and the breakdown voltage showed a dependence on the bulk properties. The observed breakdown voltages, defined at a reverse current density of 10^{-1} A/cm^2, are plotted as a function of donor concentration in Figure 58. For comparison the calculated values of breakdown voltage determined by assuming that the breakdown depends only on the critical field, E_i = 180 V/cm(71) are also plotted in

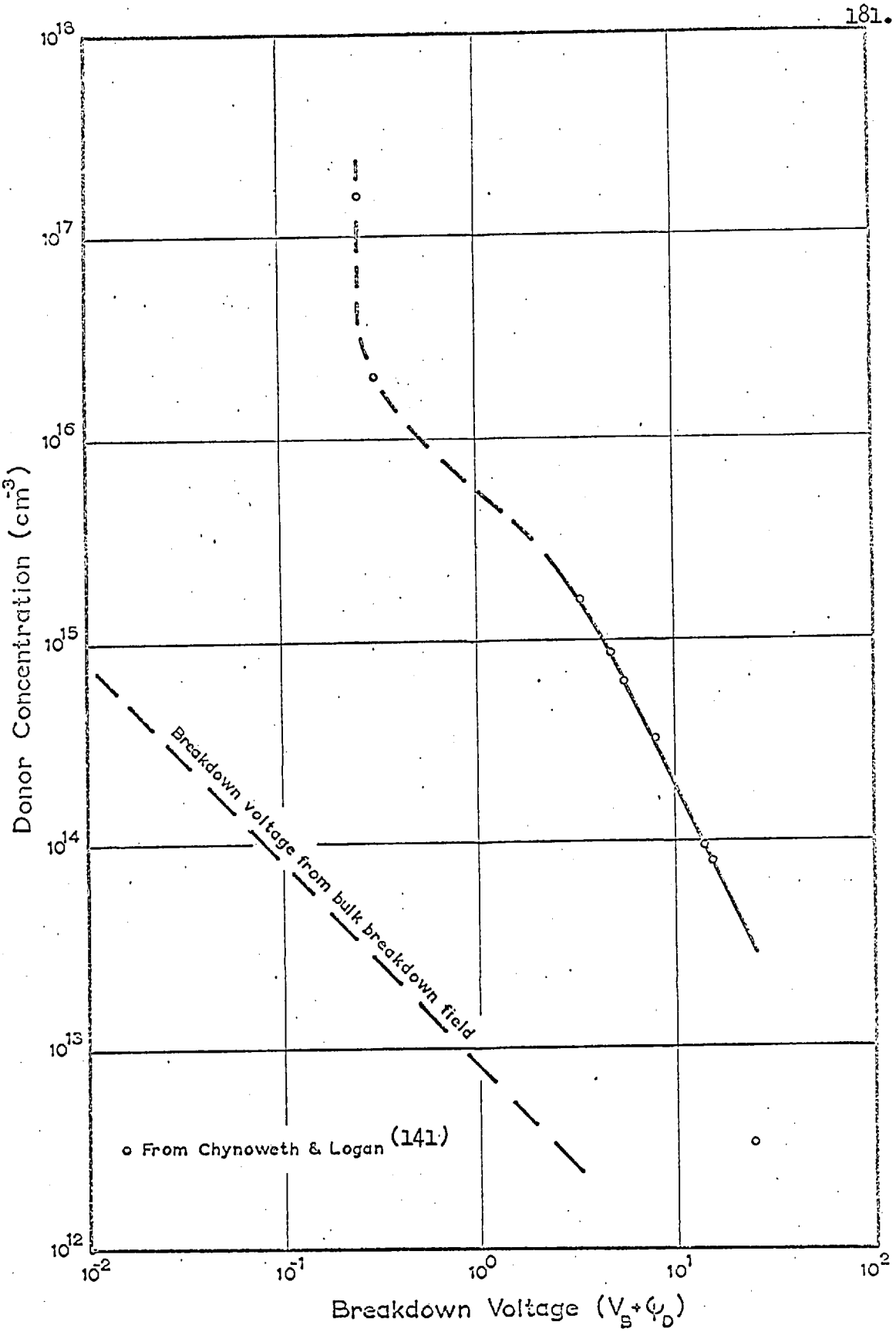


Figure 58. Breakdown voltage versus donor concentration for InSb diodes.

Figure 58. In spite of the excessive breakdown voltages observed, there is some doubt as to whether these are due to surface breakdown or a bulk effect. Hardening, as discussed in section 5.5.2, indicates that a surface effect can give a low value of breakdown voltage. When the limit of hardening has been reached there is no way of distinguishing the final breakdown from the initial low values and therefore there is no simple way of determining if the limiting breakdown is a bulk or a surface effect. Furthermore diodes broken down irrecoverably at 77°K have in some instances been permanently damaged and in other instances recovered after further etching. Such observations have only served to increase the uncertainty about the origin of the breakdown. One observation, however, strongly supports the argument for the limiting voltages being true bulk breakdown voltages. Widely differing chemical treatments, which caused low voltage currents to vary by up to four orders of magnitude, did not appreciably change the limiting breakdown voltage.

The value of breakdown for diodes formed on C355/u ($N_D = 3.3 \times 10^{12} cm^{-3}$) must be considered less reliable than the other values in Figure 58. The hardening of these diodes above 18 volts was often difficult and it was not easy to determine when or if the limiting voltage was reached. The effect of hardening on the low voltage currents was greater on these diodes than on others and the high voltage currents were very noisy. Since there is evidence for deep levels partly deionized at 77°K in this material, it is also possible that the breakdown could be enhanced by carrier multiplication through these levels. It has been shown

by Rose[224] that breakdown voltages are dependent on the trap density as well as the uncompensated impurity density. In closely compensated materials diode breakdown voltages are determined by $(N_D-N_A+N_t)$, where N_t is the density of traps, rather than by (N_D-N_A) as assumed in Figure 58. Thus, for materials where (N_D-N_A) is small compared to N_t, the axis of Figure 58 is not appropriate. It is apparent that, for C355/u, a trap concentration of about $2.5 \times 10^{13} cm^{-3}$ would bring the measured breakdown into close agreement with the other measurements. Measurements of Hall effect and photoconductivity[4][5] have shown that the density of traps in most high quality InSb is close to thisvalue and therefore this can be considered a highly probable explanation for the low breakdown voltages observed in diodes formed on C355/u.

Values of breakdown voltage for highly doped materials were estimated from the measurements on diodes formed on C275/Te and the measurements reported by Chynoweth and Logan.[141] These are also plotted in Figure 58 and show that, unlike silicon junctions (182), narrow junctions in InSb exhibit a region where the field dependence of the breakdown is of increasing importance (i.e. $r \simeq 1$ in equation(2.30)). This indicates that Zener breakdown in InSb junctions becomes significant when the concentration of the lightly doped region exceeds $5 \times 10^{15} cm^{-3}$.

It is clear that, if the observations in Figure 58 are not values of bulk breakdown, bulk breakdown must occur at voltages greater than those shown and these values are much higher than expected. It is difficult to dispute the very considerable amount

of work on breakdown in bulk specimens of InSb reviewed in Chapter 1. The field at which drift velocity saturation is observed[75] corresponds well with the optical phonon energy (0.023 ev at 300°K) obtained from infra red measurements.[225] It is interesting that significant multiplication, which requires an ionization energy of about 1.5 E_g, occurs at fields only slightly in excess of those required for drift velocity saturation. This implies that the carrier distribution at these fields includes a significant number of very energetic carriers. In any case the fields required for bulk breakdown are well established and therefore the discrepancy between them and those required for breakdown in p-n junctions must be due either to the difference in the effective length over which the field acts or to the difference in carrier concentrations in the high field regions. The carrier concentrations in the high field region of a p-n junction are very low and therefore it is possible that the number of carriers which could gain sufficient energy to make ionizing collisions would be negligible at the fields required to give significant multiplication where the population of carriers is large. However, the arguments which can be advanced in terms of the short length of the high field region in a p-n junction are more tenable.

The currents flowing in a reverse biased InSb junction arise almost entirely from generation in the space charge region, as will be demonstrated later in this section. Thus holes and electrons traverse the space charge region in approximately equal numbers and on average each electron or hole moves through only one half of the space charge region width. Since bulk breakdown occurs at higher fields in p-type[76] than in n-type[75] InSb, it is clear that for similar fields the

multiplication due to holes is much less than that for electrons and therefore only multiplication due to electrons will be considered in this discussion.

If electrons assume the saturated drift velocity ($v_d \simeq 2 \times 10^7$ cm/sec) (75) in the space charge region at breakdown, it is possible for them to pass through the space charge regions of the diodes described above in times ranging from $\sim 10^{-12}$ sec for highly doped diodes to $\sim 10^{-9}$ sec for the lightest doped diodes. Now values of the relaxation time for acoustic lattice scattering of electrons in the same materials range from $\sim 10^{-12}$ sec to $\sim 10^{-11}$ sec and the relaxation times for optical phonon scattering are not generally greatly different. Therefore, the electron transit times through the high field regions are comparable to the relaxation times. Since the electrons, which attain sufficient energy in an electric field to cause impact ionization, are certain on average to make a considerable number of collisons involving optical phonons (in silicon[226] the estimated average number is 20); the time required for an electron to gain sufficient energy in fields high enough to cause bulk breakdown would be greater than the transit time through the space charge region of a p-n junction. Thus no significant multiplication would occur in p-n junctions at the fields required for multiplication in relatively long bulk samples. This is equivalent to saying that the relaxation time of the energy distribution of electrons, for changes from the low field Boltzmann distribution to the high field distribution which gives rise to impact ionization in bulk samples, is much greater than the transit time of electrons through the high field region of a p-n junction. The observations of Prior,[227] who reported the times required to establish multiplication in InSb of the order of 5×10^{-8} sec,

would seem to substantiate this argument. Avalanche breakdown in p-n junctions would be expected therefore at higher fields than those required for breakdown in bulk material, since the ionization time decreases with increasing field.(227)

It has already been noted that the space charge width is greater and therefore the transit time for electrons is greater in junctions formed on pure materials. Thus the electrons have more time in which to make ionizing collisons and as a consequence the fields required for breakdown in lightly doped junctions will be lower than those required in highly doped junctions. In fact it is reasonable to expect the field required for breakdown to approach the value observed in bulk samples as the width of the space charge region becomes large enough to give a transit time comparable to the life time of electrons. The expected decrease in breakdown field with decreasing donor concentration in the n-regions of InSb p^+n diodes has been observed, as shown in Figure 59. The space charge widths at breakdown calculated from equation (2.19) and the maximum electric field in the junctions at breakdown are plotted in Figure 59 as a function of donor density.

In some respects the breakdown behaviour of p-n junctions in silicon and germanium is similar to that in InSb. Except for junctions on very pure material, the junction breakdown voltages are slightly higher than expected from bulk breakdown measurements. They also show a dependence on the width of the space charge region similar to InSb junctions (i.e. $r < 1$ in equation 2.30). However, the saturation drift velocities are considerably lower in Si and Ge and therefore the space charge transit times are longer. Also, for similar impurity density, the width of the space charge regions at breakdown are

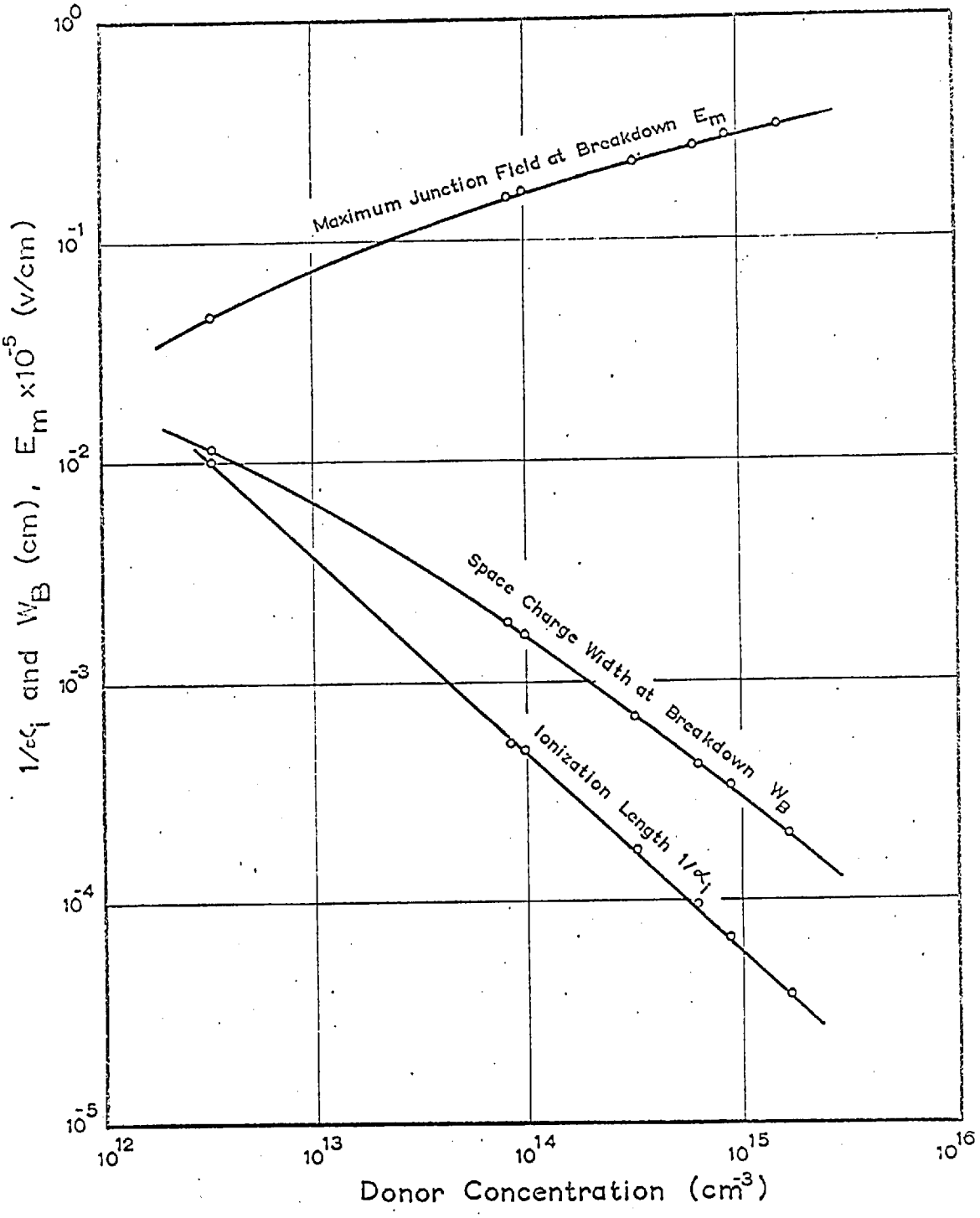


Figure 59. The values of the maximum field, junction width and ionization length in InSb p^+n junctions at breakdown plotted as a function of donor concentration.

greater in Si and Ge junctions than in InSb and therefore the transit time is even further increased. Because the space charge transit times are longer in Si and Ge and relaxation times are in general shorter, it would be expected on the basis of the previous arguments that the breakdown voltages would be in closer agreement with the values predicted from bulk measurements.

By considering the relationship between the space charge region transit time and the relaxation time for ionization, it has been established that the breakdown voltage in InSb p-n junctions would be greater than that predicted by bulk breakdown measurements and that the departure from predicted values would be greater than that for Si or Ge p-n junctions. However, these qualitative arguments do not establish that the large departures from the expected breakdown voltages in InSb junctions can be justified. Without some knowledge of the ionization coefficients as a function of electric field, an accurate analysis cannot be made. The theory of avalanche multiplication[228] is very complex and it has been found difficult[226] to simplify it sufficiently to apply it to measurements on p-n junctions. However, if it is assumed that the ionization coefficient (the average number of ionizing collisions made per centimeter of travel) is the same for electrons and holes, the ionization coefficient α_i can be shown[229] to obey the following relation:

$$\alpha_i = -\frac{4}{W_1^3}\frac{dW_1}{dE_B} \qquad \ldots\ldots (5.6.)$$

where $W_1 = \left(\frac{2\epsilon}{qN_D}\right)^{\frac{1}{2}}$ for a p^+n junction and E_B is the maximum field in

the junction at breakdown. Since the value of E_B varies with N_D as shown in Figure 59, α_i can be determined as a function of field if the value of E_B is measured on diodes having a range of donor concentrations N_D in the lighlty doped region. Using the results shown in Figure 59, α_i has been calculated for InSb as a function of the electric field from 4×10^3 V/cm to 3.3×10^4 V/cm. The values of α_i are plotted in Figure 60.

Using the values of α_i in Figure 60, it is possible to make a quantitative estimate of the average time required for a carrier to make an ionizing collison in a specified field and therefore to check if the qualitative arguments presented earlier do explain the high breakdown voltages observed in InSb diodes. In fact, it is not necessary to calculate the ionization time if it is noted that the average distance a carrier must travel before making an ionizing collison is $1/\alpha_i$. It is clear that breakdown is unlikely in a p-n junction if $1/\alpha_i$ is greater than the width of the space charge region. The values of $1/\alpha_i$ at breakdown are plotted in Figure 59 and it is clear that, in general, the width of the space charge region is about five times the ionization length. For a large multiplication factor, as implied by breakdown conditions, this ratio seems eminently reasonable.

On the basis of these calculations the values of breakdown voltage observed in InSb diodes would appear to be acceptable. However, the ionization coefficients were not obtained independently of the breakdown observations and therefore it could be argued that the above calculations are only an indication of self consistency. To obtain an independent check on the acceptability of those breakdown voltages, a more quantitative comparison between InSb junctions and those formed in Si and Ge was made.

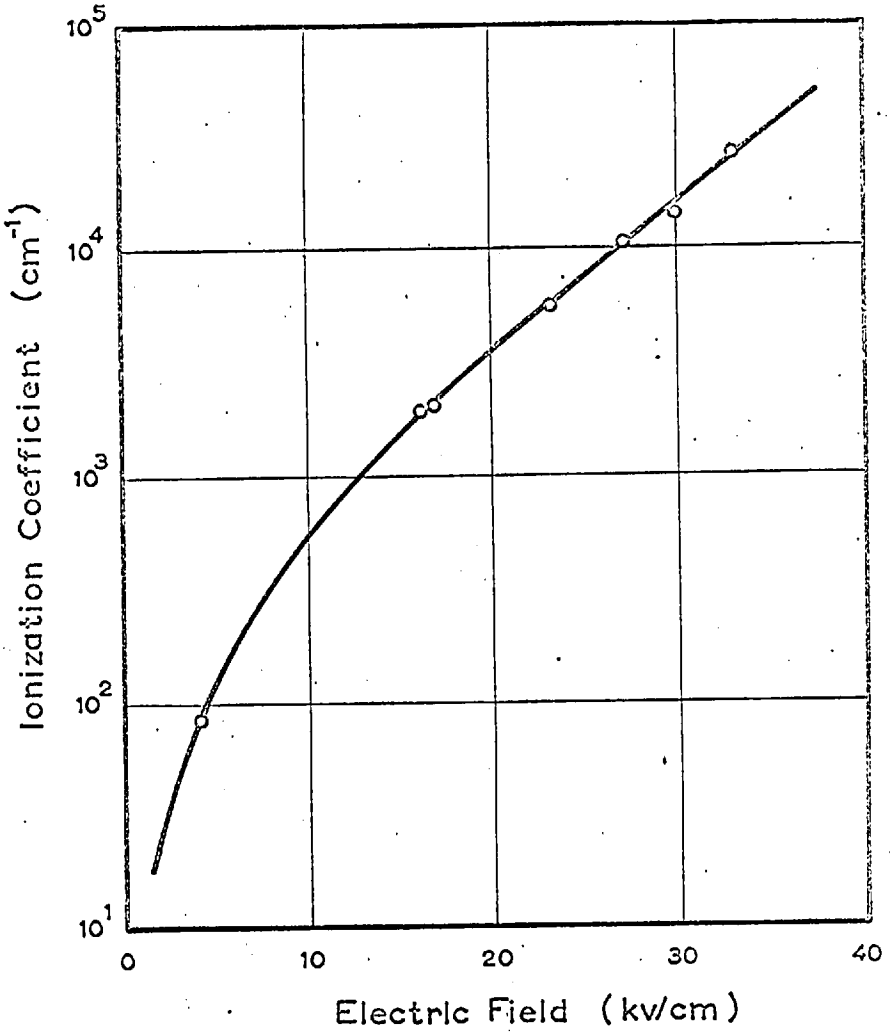


Figure 60. The ionization coefficient or the average number of ionizing collisions per centimeter of travel in InSb as a function of electric field.

The effects of different saturation drift velocity and different space charge widths in InSb, Ge and Si diodes can be normalised by comparing diodes having the same space charge transit times at breakdown. Since the energy required for impact ionization is generally more than ten times the optical phonon energy and is, in the simple case of equal hole and electron effective masses, approximately 1.5 E_g; the breakdown voltages observed in diodes having the same space charge region transit times would be expected to have values related by ratios similar to the ratios of the energy gaps of the semiconductors.

The values of space charge transit time at breakdown in silicon junctions were calculated using the breakdown voltages reported by Miller[182] and a saturation drift velocity calculated from the optical phonon energy.[226] The values are plotted as a function of breakdown voltage in Figure 61. Similar calculations were done for germanium diodes using the breakdown voltages reported by Miller[78] and the saturated drift velocity reported by Gunn.[77] These are also plotted in Figure 61. The space charge transit times for InSb diodes were calculated using the present breakdown measurements and the saturated drift velocity reported by Steele and Glicksman.[75] These values are also plotted with those of Si and Ge in Figure 61. It is apparent that for any value of space charge transit time the values of breakdown voltage are in ratios reasonably close to those of the energy gaps. This, coupled with the previous discussion and calculations, suggests that the breakdown voltages observed in InSb diodes are the right order of magnitude and not in disagreement with the bulk measurements as inferred by Lee and Kaminsky.[79]

Earlier in this section it was mentioned that, if hardening of the

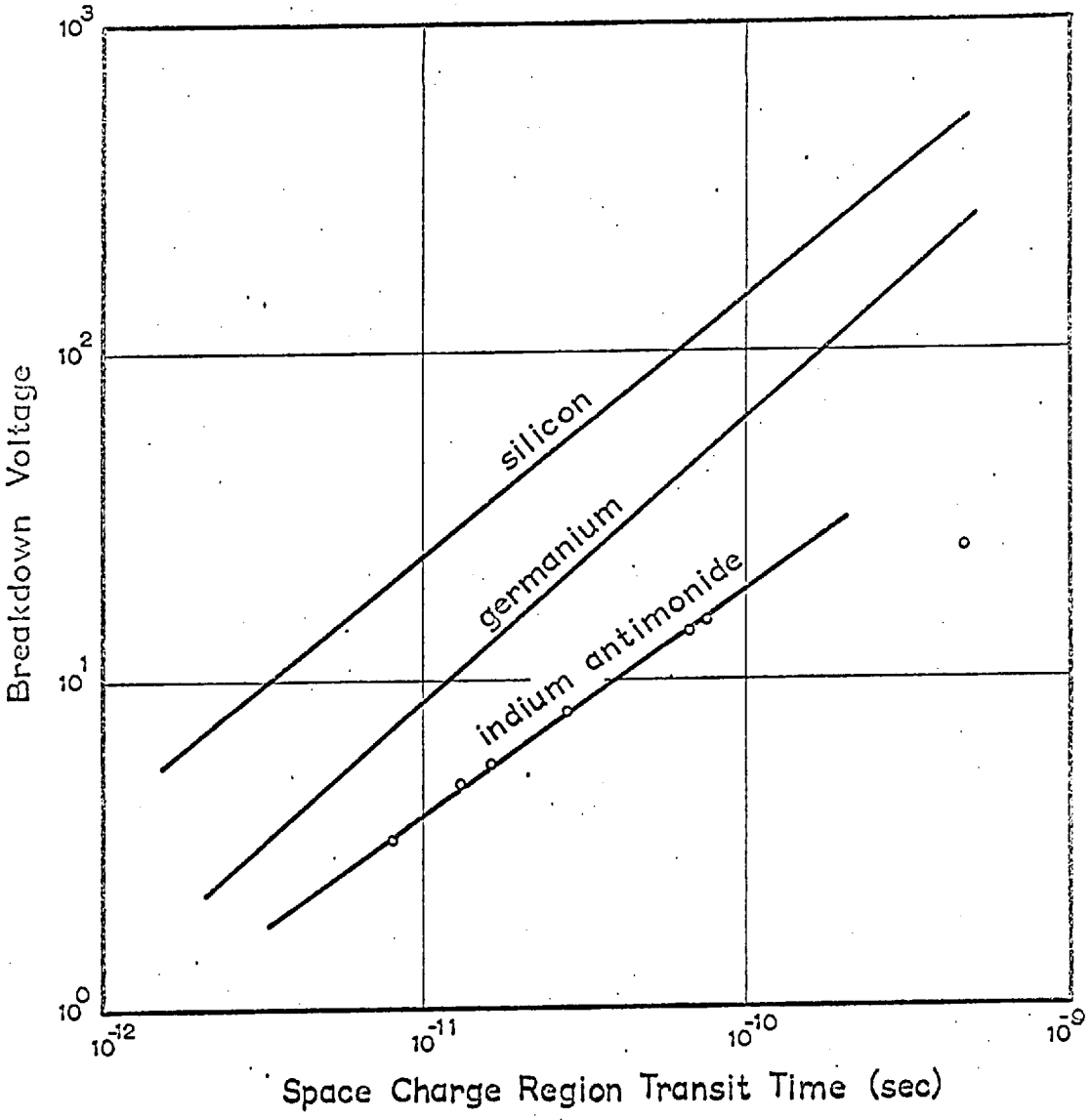


Figure 61. Breakdown voltages for p-n junctions in silicon, germanium and indium antimonide plotted as a function of space charge transit time at breakdown. The ratios between breakdown voltages for any value of transit time are approximately equal to the energy gap ratios.

diodes was avoided, the currents at low reverse bias were generally quite reproducible. These currents were measured at a constant reverse voltage of 100 mv (greater than 4kT at the highest temperatures used) as a function of temperature. The currents at low temperatures obeyed equation (2.26) and when values of current density divided by temperature to the three halves power ($J/T^{3/2}$) were plotted against reciprocal temperature, as shown in Figure 62, a straight line was obtained. This agreement with the theory of section 2.1.6. indicates that the current arises from space charge generation due to a trapping level near the center of the energy gap. From the slope of the curves in Figure 62 a value of the activation energy, $E_A = 0.120 \pm 0.004$ ev, was obtained. Using this value and an iteration procedure to solve equation (2.27) two possible values for the $0^{\circ}K$ trap energy were obtained, $E_{to} = 0.133$ ev or 0.120 ev. If the trap energy is assumed to have a temperature dependence proportional to its position in the energy gap, the values of trap energy for temperatures up to $300^{\circ}K$ are indicated in Figure 63. Since it is most improbable that the trap energy would be temperature independent, the linear dependence applied in this case would seem to be a reasonable first approximation.

Recombination measurements[4][5] on bulk samples over the temperature range from $80^{\circ}K$ to $200^{\circ}K$ have been interpreted in terms of two trapping levels assuming the levels to be temperature independent. The deepest of these two levels was located at 0.11 ev with respect to the valence band. If the average value of the upper possible level between $80^{\circ}K$ and $200^{\circ}K$ is taken from Figure 63, a value of 0.108 ev is obtained for the trapping level. These two values are in excellent agreement and, providing it can be shown that the currents of Figure 62

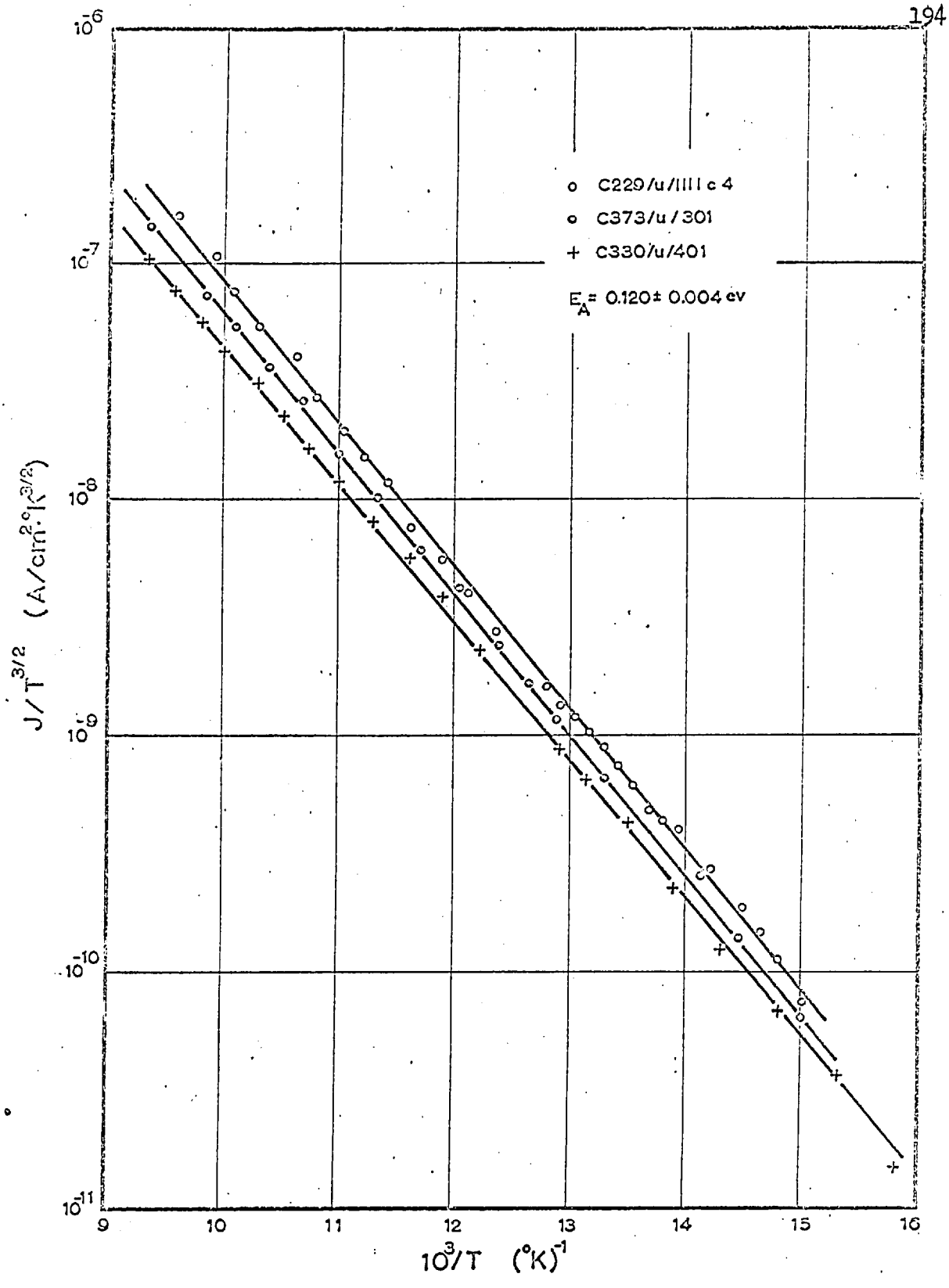


Figure 62. Reverse current below 110°K plotted as a function of temperature to obtain the activation energy associated with the trapping level.

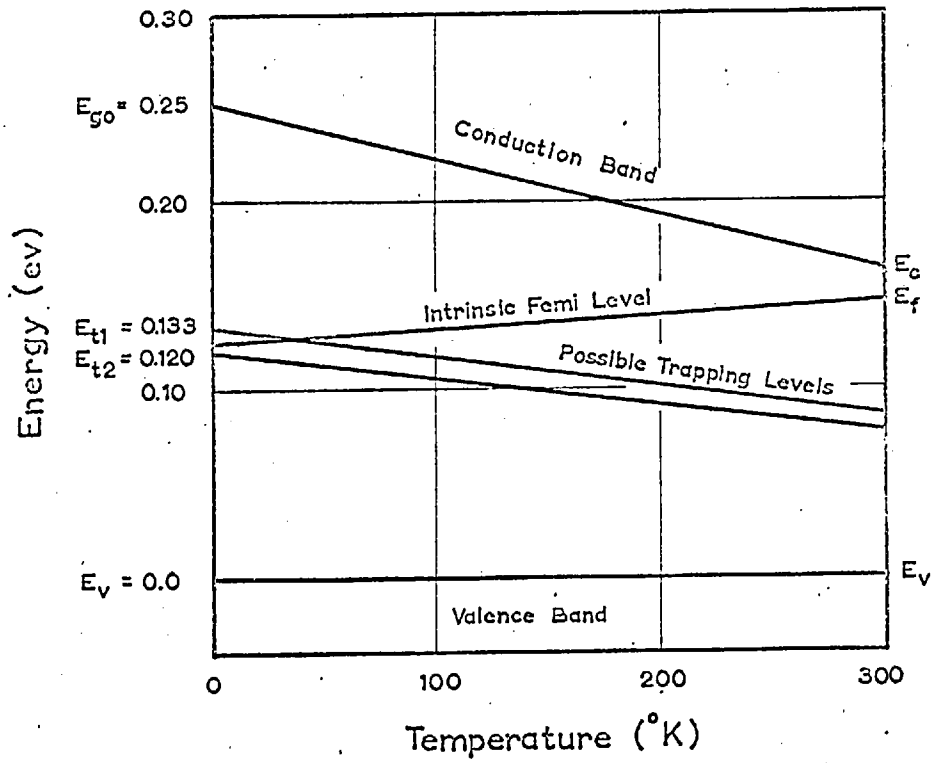


Figure 63. The variation with temperature of the possible energy levels of the traps assuming a temperature dependence proportional to that of the energy gap.

are due to recombination in the bulk rather than at the surface, this agreement would indicate that the trap energy in InSb is 0.133 ev with a linear temperature dependence of -1.55×10^{-4} ev/°K.

To establish the possibility of the observed reverse currents being a bulk effect, values of carrier lifetime obtained from pulse measurements and measurements on similar bulk samples[4][5] were substituted into equation (2.24). The theoretical values of recombination current calculated in this way were in close agreement with the observed values. For example; using equation (2.19) to calculate the space charge width, values of n_i from section 5.1.1. and values of $\tau_p = 4 \times 10^{-7}$ sec and $\tau_n = 4 \times 10^{-9}$ sec; J_{rg} for C373/u was calculated to be $7.5 \times 10^{-7} A/cm^2$ while the measured value was $8 \times 10^{-7} A/cm^2$. The parameters used to calculate values of reverse current also gave values of forward current in good agreement with those observed at low bias. Furthermore, the variation of J_{rg} as a function of impurity density, indicated by the different curves in Figure 62, is in reasonable agreement with that expected from equation (2.24). Such agreement between theory and experiment in, not only the absolute values of the redombination current, but also in the temperature and impurity density dependence, strongly suggests that the reverse currents observed are due to generation in the bulk and not at the surface. These measurements would appear therefore to support the interpretation of recombination measurements on bulk samples and to indicate the trapping level to be temperature dependent. In the analysis of recombination measurements workers have generally assumed, for the sake of simplicity and without justification, that the trapping level was temperature independent.

For reverse voltages greater than 4kT at the highest value of

temperature, the reverse current is given by equation (2.25). If it is assumed that, at the maximum temperature, the p^+ region of a p^+n diode remains extrinsic, equation (2.25) can be written as:

$$J_{rev} = K_i \frac{n_i^2}{N_A} + K_2 \frac{n_i^2}{(N_D + \sqrt{N_D^2 + 4n_i^2})} + K_3 \frac{n_i}{\cosh\left(\frac{E_t - E_i}{kT} + \ln\sqrt{\frac{\tau_p}{\tau_n}}\right)} \quad \ldots\ldots (5.7)$$

where $K_i = q\sqrt{\frac{D_{np}}{\tau_{np}}}$, $K_2 = 2q\sqrt{\frac{D_p}{\tau_p}}$ and $K_3 = \frac{qW}{2\sqrt{\tau_n \tau_p}}$

In this case τ_n, τ_p and D_p are lifetimes and diffusivity in the n region while D_{np} and τ_{np} apply to electrons in the p^+ region. The coefficients K_1, K_2 and K_3 are assumed independent of temperature.

Using the value of E_t obtained earlier, values of n_i from section 5.1.1, $N_A = 5 \times 10^{18} cm^{-3}$ and values of N_D from the measurements recorded in section 5.1.2; values of K_1, K_2 and K_3 were chosen to fit the measurements of reverse current as a function of temperature. An example of the close agreement obtained between the measured values and those given by equation (5.7) is shown in Figure 64. The values of K_1, K_2 and K_3 required to fit the measurements were compared to values calculated using the same lifetimes and diffusivities as those used to calculate theoretical values of J_{Ln}, J_{Lp} and J_{rgo} for the forward characteristics of the same diodes. Agreement between the experimental and theoretical values of K_3 were good as already described. The experimental values of K_1 and K_2 were generally about

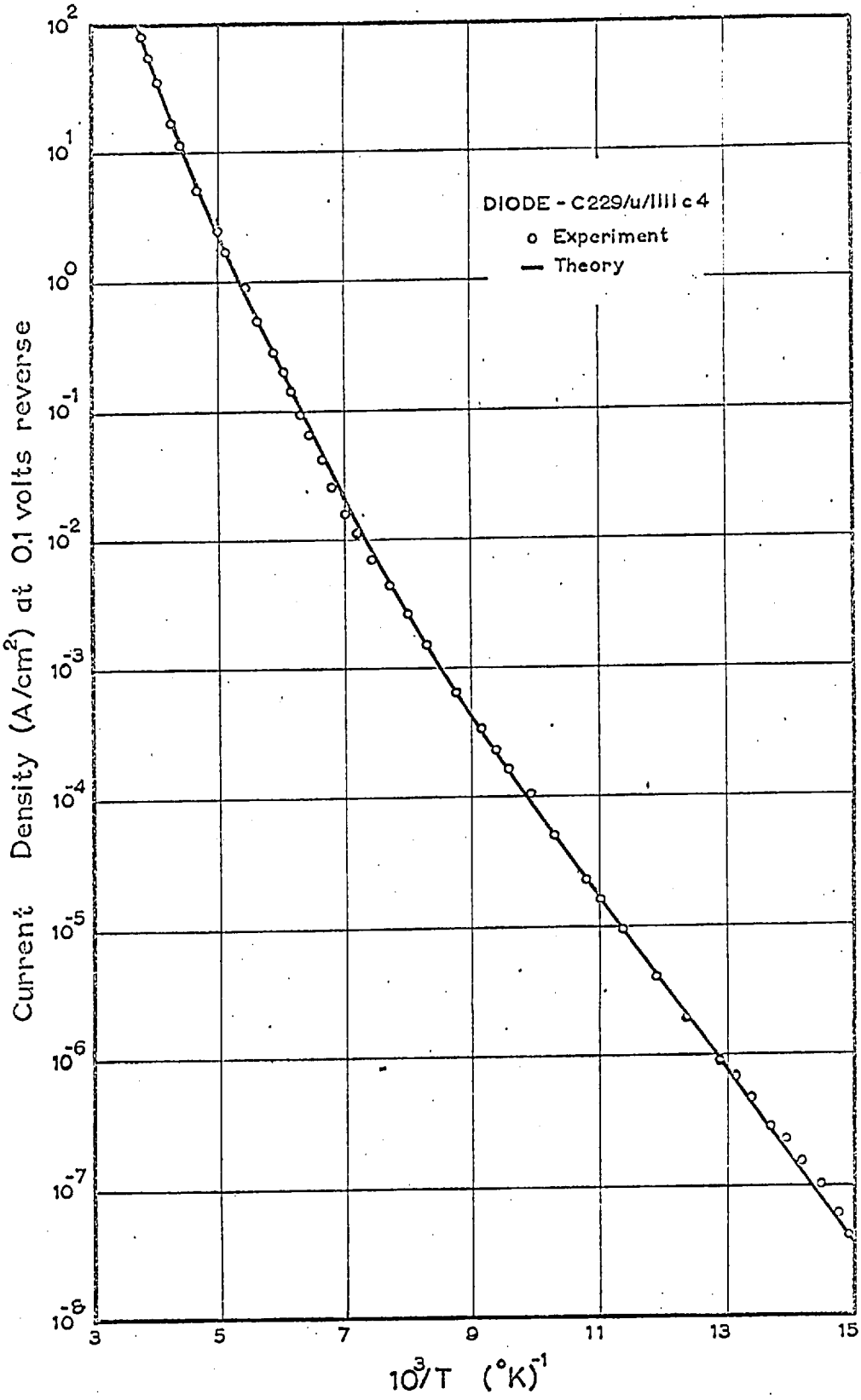


Figure 64. The agreement between measurements of reverse current as a function of temperature and that theoretically determined from equation (5.7

a factor of two higher than the calculated values. This is not surprising in view of the fact that lifetimes and diffusivities have been assumed temperature independent. The terms involving K_1 and K_2 only become appreciable at temperatures above about 150°K. In this temperature range lifetimes become rapidly varying functions of temperature due to Auger recombination and diffusivities also vary quite rapidly due to the increasing effect of acoustic lattice scattering. The degree of agreement between theory and experiment under such conditions is due to the tendency for the ratio of lifetime to diffusivity to remain relatively constant. However, lifetime does decrease more rapidly than diffusivity and therefore the increases in K_1 and K_2 can be considered to be in agreement with theory.

The close agreement between the theoretical temperature dependence of reverse current as stated in equation (5.7) and the values measured at temperatures ranging from 60°K to 300°K is further evidence to support the view that these currents arise from the bulk behaviour of the diodes. As a point of interest, it also provides some evidence for the accuracy of the corrected values of n_i given in section 5.1.1. It is perhaps pertinent to note that this treatment of reverse currents as a function of temperature shows the complete futility of the interpretation placed on earlier measurements[151] of reverse currents in InSb grown p-n junctions.

5.4. Transient Measurements.

5.4.1. Capacity measurements.

In general capacity measurements were made only on diodes which were indicated by the current-voltage characteristics to be properly

etched. Although the current-voltage characteristics of such junctions were found to agree within 10% (see section 5.3.1.), the capacity measurements varied, not only between diodes but from etch to etch, by up to 50%. The variation was even higher on the very pure materials where hardening affected the capacity (see section 5.5.2.) However, within these limits the diode capacities appeared to obey the theory of Chapter 2. Typical results are shown for diodes formed on seven different materials in Figure 65. The capacity given by equation (2.31) is also shown in Figure 65 for $N_A = 5 \times 10^{18} cm^{-3}$ and $N_D = 10^{13}$, 10^{14}, 10^{15} and $10^{16} cm^{-3}$. It can be seen that the voltage dependence of the measured capacities indicates the junctions to be abrupt. This is further confirmation of the conclusion drawn in section 5.2 concening the abrupt nature of these diodes. It is also apparent that the agreement between theory and experiment is reasonably good. Capacity effects arising from surface phenomena will be described in section 5.5.

As mentioned in section 5.1.3., capacity measurements can be used to estimate the diffusion potential, ψ_D. It is apparent from Eqn.(2.31) that on a plot of C^{-2} versus applied voltage, V, the intercept on the voltage axis will be ψ_D. InSb diode capacity measurements plotted in this way have always given valves of ψ_D about forty millivolts lower than the value of diffusion potential obtained from calculations and from the DC measurements of section 5.1.3. However, this is not surprising since Sparkes[169] more careful analysis of junction capacity has shown that, although equation (2.31) accurately describes the capacity, ψ_D obtained from the equation is not the diffusion potential.

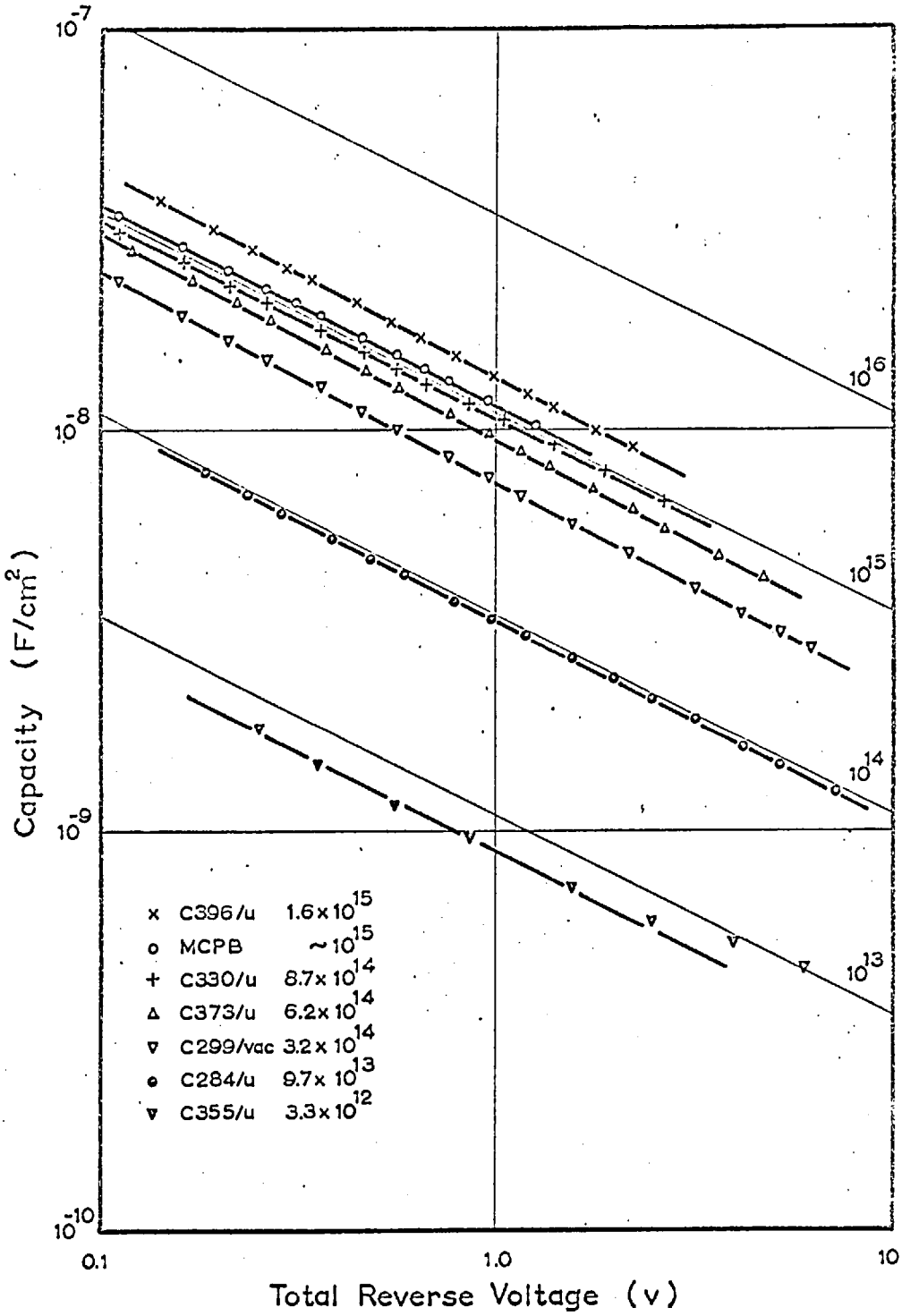


Figure 65. Values of unit area junction capacity plotted as functions of reverse voltage and of donor concentration.

5.4.2. Pulsed Lifetime Measurements.

In general hole lifetimes in InSb are about two orders of magnitude greater than electron lifetimes.[4][5] In the case of p^+n junctions the lifetime of electrons in the degenerate p^+ region has been shown (see section 5.3) to be reasonably less than 10^{-11} seconds. Therefore only the effects of holes injected into the n-region can be observed in pulse measurements on p^+n junctions.

Because of the difficulty to avoid conditions of medium or high level injections with the pulse system used in these experiments, only the reverse recovery or constant current phase of the reverse pulse response was used. Reverse recovery times were measured as described in section 3.4.4. on InSb diodes as a function of temperature, injection level, doping density, diode area and surface treatment. The reverse recovery times were converted to effective lifetimes using equation (2.33). Measurements from 63°K to 147°K have indicated that the hole lifetime increases slightly with increasing temperature. However, since the total change was generally about 10%, it was considered of doubtful significance. It was concluded, therefore, that the hole lifetime in this temperature range was essentially constant. In fact, at high levels of injection no temperature dependence was observed. Providing the procedure for chemical treatement outlined in section 3.3 was followed, the lifetime was also independent of diode area and not sensitive to the variations in surface treatment which gave rise to changes in the low current DC characteristics and in the capacity. This indicated that bulk lifetimes were observed and since the current densities were high (generally above $1 A/cm^2$) and the junction dimensions were large compared to a diffusion length, lifetimes largely independent of surface recombination would be

expected. Figure 66 shows typical values of hole lifetime measured as a function of injection level and of donor concentration. No attempts have been made to interpret these measurements quantitatively in terms of a recombination model but a number of observations can be made concerning them. Lifetimes given by radiative, Auger and trapping recombination processes are independent of injection levels for low levels of injection but become dependent on injection level for medium or high level injection. It would seem that the observed variation of lifetime in high level injection could not be accounted for by radiative recombination since lifetimes determined by this process decrease linearly with increasing injection level.[230] However, an attempt to observe recombination radiation as a function of injection level would be an interesting and informative experiment. Auger recomination would appear to be an unlikely process at these low temperatures[230] and lifetime at high level injection is inversely proportional to the square of the injection level for Auger recombination. The measurements give a similar variation with injection level to that reported for other diodes in high level injection[176] and to that predicted by the theory for a single trapping level.[231] If the curves of Figure 66 are extrapolated linearly to zero injection level the values of lifetime obtained are somewhat lower but, in general, in good agreement with those obtained from photo-conductivity experiments[4][5] on bulk specimens of similar InSb. The measurements could, therefore, be taken to indicate that a trapping recombination process continues to dominate at these high injection levels but, in the absence of a more quantitative analysis of the measurements, their real value lies in the justification that they provide for the assumptions made concerning lifetimes and surface effects

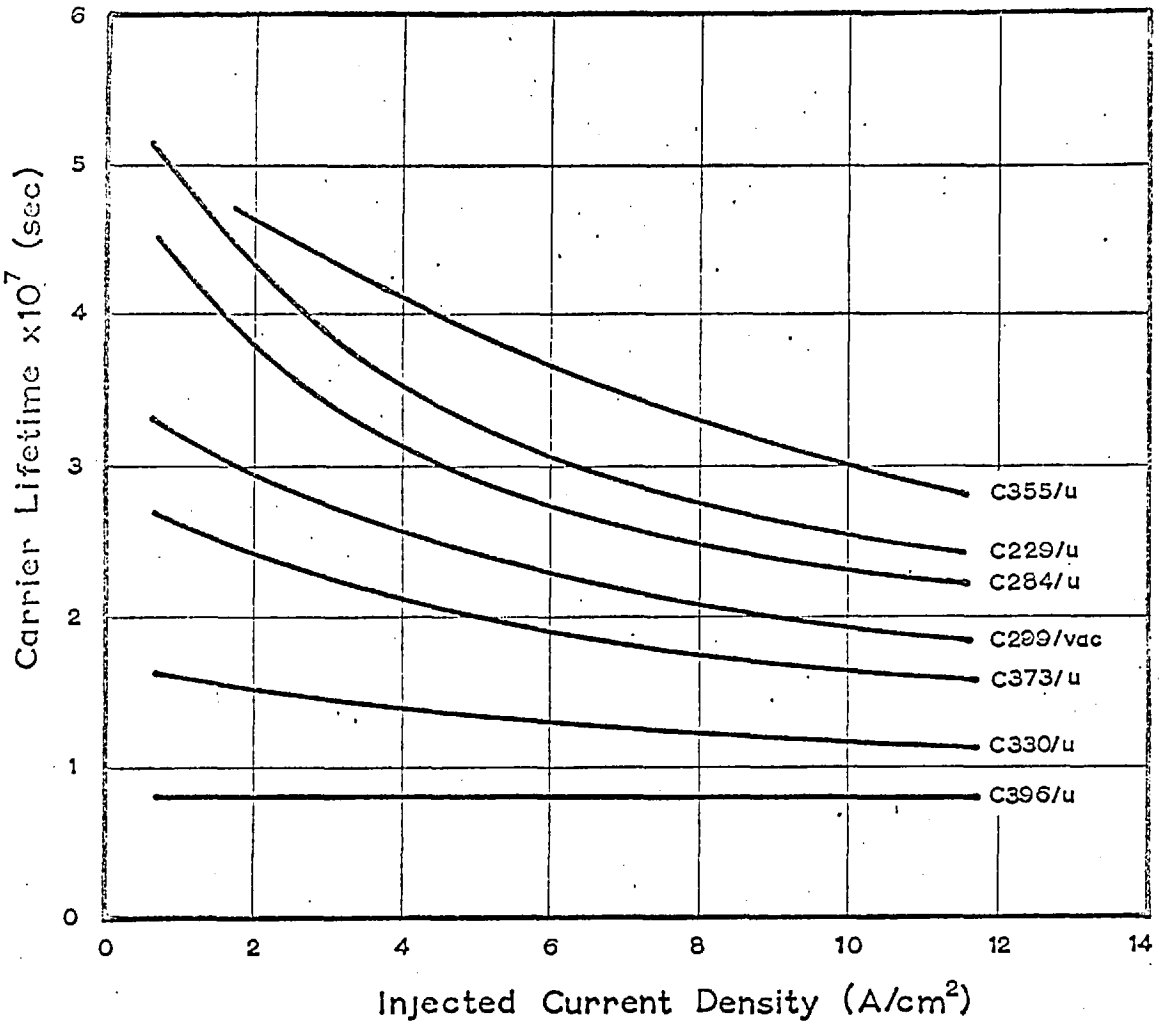


Figure 66. Typical values of carrier lifetime measured using the reverse recovery phase of the pulse response of InSb p^+n diodes.

in the interpretation of the other measurements reported in section 5.3. Since these lifetime measurements are characteristic of carriers within a diffusion length of the metallurgical junction, they also indicate that the alloying process used for diode fabrication does not significantly affect the carrier lifetime in the region of the junction.

5.4.3. Plasma Effects.

It was observed early in these experiments that when InSb diodes were driven hard into reverse breakdown a negative resistance characteristic was obtained. In this region the diode switched very rapidly from a high voltage, low current state to a low voltage, high current state. A typical characteristic observed by means of the 50c/s curve tracer described in section 3.4.2. is shown in Figure 67. In the low voltage state extremely high currents could be passed for short periods of time without causing any damage to the diode. It was noted that, when a high series resistance was used in the curve tracer, high frequency oscillations occurred in the negative resistance region. It was also found that, when a diode was DC biased into this region through a large series resistance and the 50Ω input of the sampling oscilloscope, pulses were generated by the diode. These pulses had a rise time which could not be measured on the sampling system (i.e. less than 4×10^{-10} sec) and their length was determined by the length of coaxial line connected across the diode. The pulse amplitude was one half the voltage difference between the high and low voltage states. Figure 68 shows a typical pulse generated in this way. It was clear that the diode was acting as an extemely fast switch which, in conjunction with a length of coaxial cable, formed a simple coaxial line pulse generator. It was

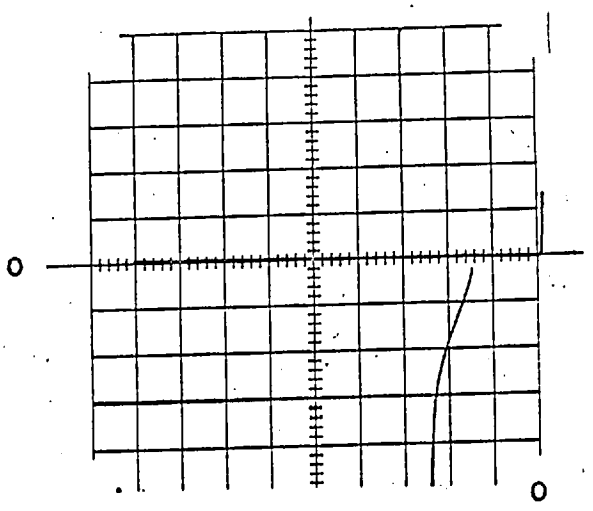


Figure 67. A typical high current reverse breakdown of an InSb p^+n diode. (1 v/div horizontal and 5 mA/div vertical).

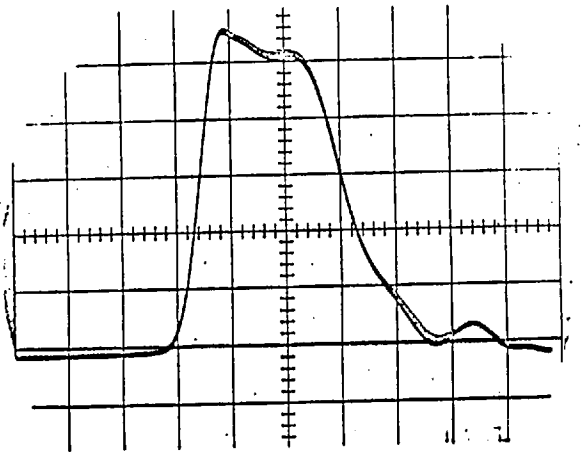


Figure 68. A pulse generated by biasing an InSb p^+n diode into the high current reverse breakdown region. (1 v/div or 20 mA/div vertical and 10^{-9} sec/div horizontal)

further observed that the pulse repetition rate increased rapidly with increasing reverse current. Judging from the sampling rate of the oscilloscope, pulse repetition rates in excess of 100kc/s were easily obtained When the diode was placed in a magnetic field the voltage of the low voltage state was increased and hence the amplitude of the pulses generated was decreased.

All of these observations indicate that in high reverse bias some type of plasma is initiated with its ensuing negative resistance characteristic. However, the nature of the plasma is not at all clear. When a diode has been damaged by passing high currents in the low voltage state for excessive times, it has been observed that in some instances the diode will recover after etching. This indicates that the breakdown can occur in the vicinity of the junction surface or in the interior of junction. If this is not a surface effect then it would appear to be a plasma having characteristics different to any previously reported in p-n junctions.

The reverse currents of silicon diodes biased close to breakdown are known[229] to arise from the superposition of small current pulses. These microplasmas each conduct a current of less than 100 μA for periods less than 2μ sec. It has been shown[232] that the microplasmas emit a continuous spectrum of photons having energies as high as 3.5 ev. The risetime of the pulses was estimated to be about 10^{-11} sec.[233] The only characteristic common to these pulses and those observed in InSb diodes is the extremely short rise time. Currents in the pulses observed in InSb diodes are three orders of magnitude higher and the pulses

occur when the diode is biased fully into breakdown, not at the onset of breakdown as in the case of microplasmas in silicon junctions.

A secondary breakdown is observed in silicon junctions[234] which is similar in many respects to these observations in InSb junctions. However, the negative resistance region gives a switching rise time of about 60μ sec. These mesoplasmas[234] are thought to be due to the formation of a plasma which actually melts a small volume of semiconductor in the junction and its time dependence is therefore associated with rather slow thermal processes. In view of the switching speeds observed in InSb this explanation for mesoplasmas in silicon would not appear applicable to the observations in InSb.

In Chapter 1 other plasma phenomena which have been observed in InSb were reviewed. The only one which would seem to have any similarity to this situation is the injected plasma used in madistor devices.[74] These plasmas also give a negative resistance characteristic and the voltage in the high current state is influenced by a magnetic field in the same way as that reported above for diodes in the high current reversed biased state. However, these are large volume plasmas which result from large injected carrier densities and again the times required to establish these plasmas are of the order of 1 μ sec. It would seem therefore that none of the plasma phenomena which are observed, and to some extent understood, either in InSb or in silicon p-n junctions are similar to that observed in reverse biased InSb p^+n junctions.

Although the nature of this phenomena is not understood, it is of interest to note that it could form the basis of a very simple fast pulse generator. It would be composed of a diode, a length of coaxial line, a source of DC current and a small electromagnet. The rise time of the pulse generator would be determined largely by the design of the associated circuitry, the pulse length would be determined by the length of the coaxial line, the pulse repetition rate would be determined by the reverse bias applied to the diode, and, although the amount of possible variation has not yet been established, it is possible that amplitude control could be achieved by variations in magnetic field. In view of the complexity and cost of available pulse generators which have comparable specifications (none with similar pulse repetition rates have comparable rise times), the simplicity of this system is almost incredible.

5.5. Surface Effects.

5.5.1. Influence of Etching and Ambients.

In Chapter 3 the care required in chemical treatment and subsequent handling of diodes was emphasized. In this section some measurements will be presented to demonstrate the effects of etching and handling on the diode characteristics. Figure 69 shows the variations in the **forward** current of a diode which resulted from variations, both intentional and unintentional, in the chemical treatment. Etch (1) consisted of CP4A only and the characteristic is typical of diodes treated in this type of etch.

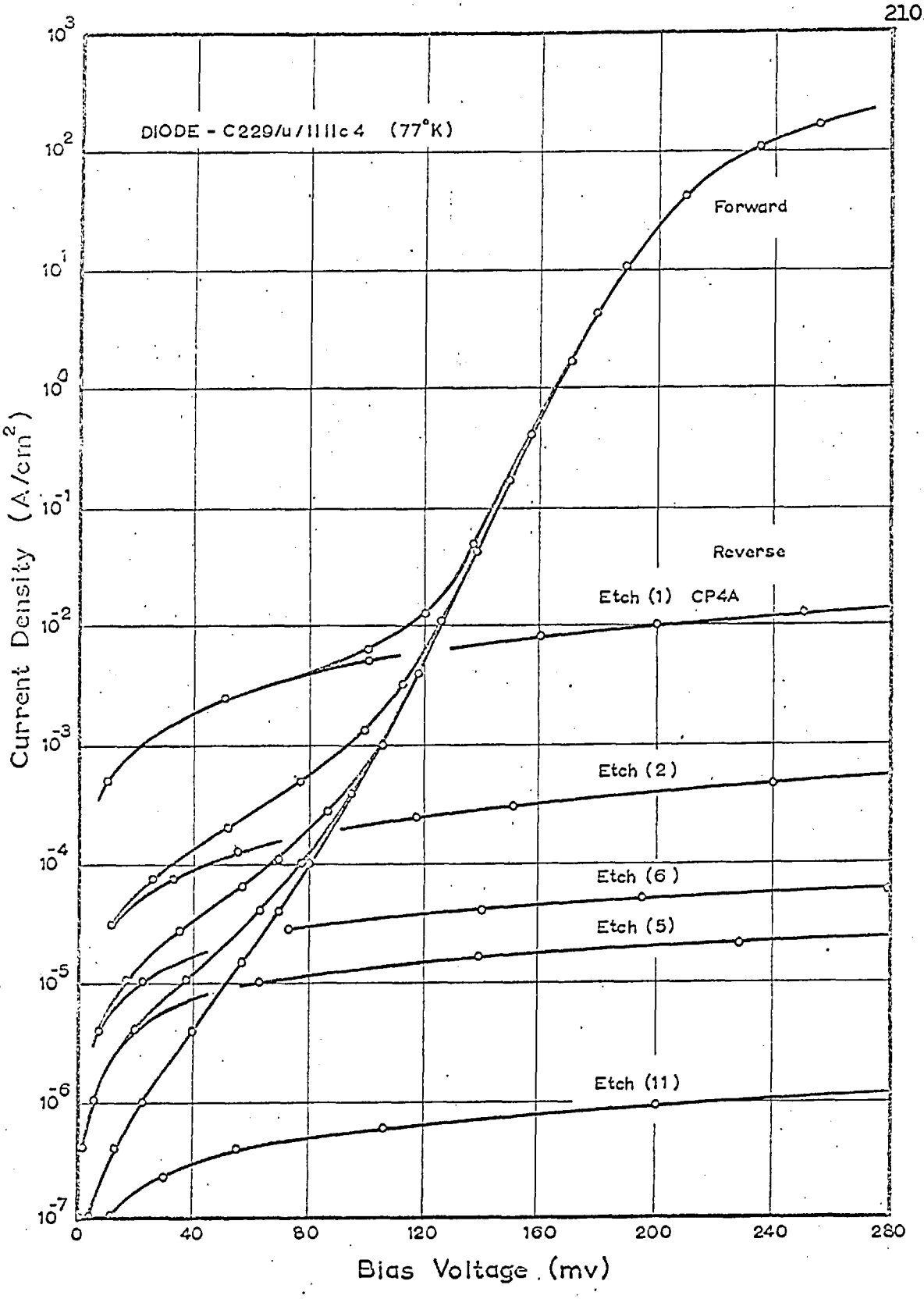


Figure 69. Variations in forward bias current which resulted from unintentional (except for etch (1)) variations in surface treatment.

The other characteristics shown were measured following the etching procedure described in Chapter 3. Although care was taken to avoid any contamination during etching, washing or drying; on occasion, this kind of variability was observed. The area change involved between etch (1) and etch (11), in the example of Figure 69, was less than a factor of three and the characteristic after etch (10) was similar to that for etch (2). It is apparent, therefore, that such variations are due to surface contamination and indicate the extreme sensitivity of diode characteristics to the nature of the surface.

This sort of behaviour is further exemplified in the measurements of reverse current as a function of temperature shown in Figure 70. The measurements in this case apply to more than one diode but it is clear that CP4A etches give rise to surface states having a low activation energy and that similar surface states can result from contamination during the chemical treatment outlined in Chapter 3. It is interesting to note that in some cases no evidence for the presence of surface states appeared until the temperature dropped below 70°K.

An observation which proved important to the techniques of handling, following chemical treatment, was that in some cases the diode surface states were photo-sensitive. Figure 71 shows the reverse characteristic of a diode treated chemically in the normal manner. After allowing it to stand at room temperature for several hours it was recooled. The reduction in reverse current indicated a change in the surface states. To determine if this was due to slow photo-sensitive surface states the cryostat was

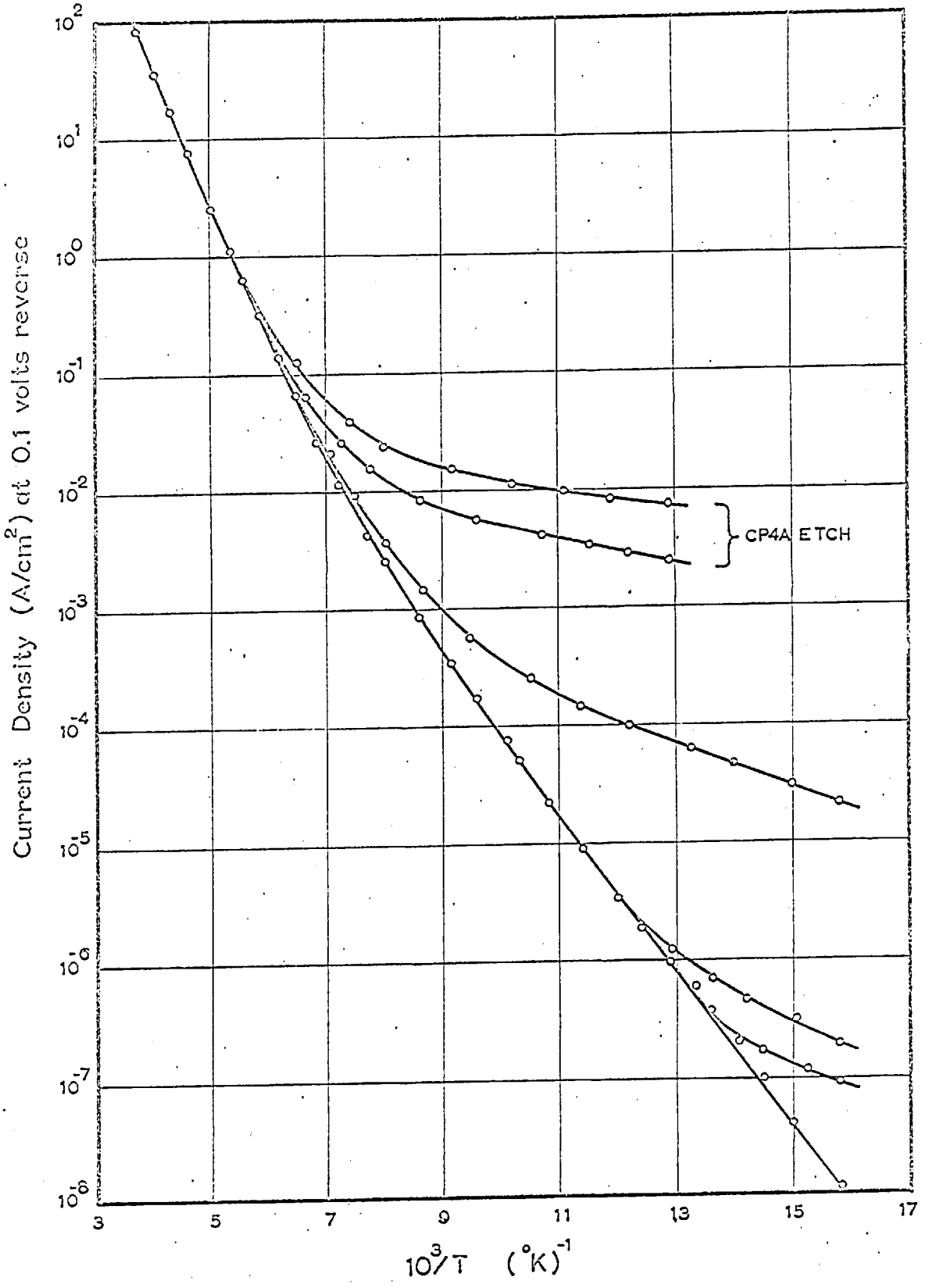


Figure 70. Variations in reverse current resulting from changes in the nature of the junction surface.

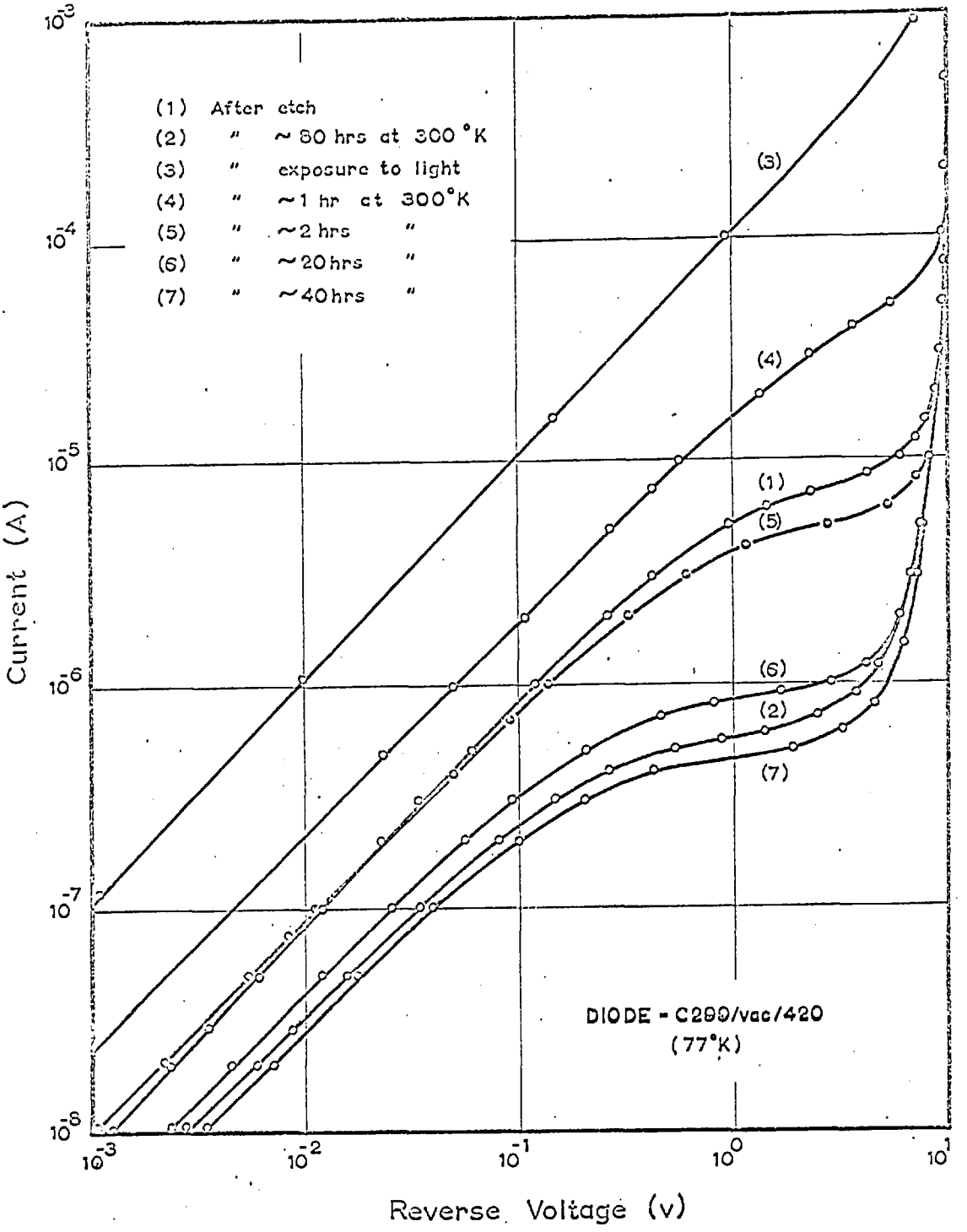


Figure 71. The reverse characteristics of an InSb p^+n diode demonstrating the presence of slow photosensitive surface states.

opened and the diode was exposed briefly to intense illumination. A large increase in the reverse current (see Figure 71) was observed and no decrease in this value was detected after six hours at 77°K. The large decreases in reverse current observed after warming the diode in darkness even for short periods (see Fig.71), indicated the presence of slow photo-sensitive surface states with, not surprisingly, extremely temperature dependent relaxation times.

The presence of undesirable surface states resulting from carelessness or accidental contamination during the chemical treatment could also be detected by an increase in diode capacity. In some instances increases as high as three times the expected capacity were observed, similar to the values reported by Lee and Kaminsky.(79)

It has also been observed, as mentioned in section 5.3.1. and in a recent publication,(223) that elimination of surface effects was much easier for diodes formed on B {111} surfaces. It has been shown(235) that many metals present at very low concentrations in pure acids deposit readily and, in some cases, irreversibly onto InSb. In view of the polarity dependent nature of InSb surfaces, as discussed in section 1.22 and Chapter 4, it is probable that the rate of deposition and the equilibrium concentrations of contaminants established on a surface are governed by its polarity type. Even if a junction surface departs from A {111} or B {111} during etching, it will be composed of surfaces such as {hhl} and {hll} which are polar in the same sense as the associated {111} surface. In such circumstances it would be surprising if polar surface effects were not observed

in diodes. Experiments with diodes formed on {100} surfaces were not sufficiently exhaustive to permit any relaible observations on the importance of surface effects.

In 1961 Davis[236] reported that low surface recombination velocities could be produced in InSb by bombarding the surface with ions or electrons from a high voltage discharge in nitrogen. This observation appears to be verified by lifetime measurements[5] on n-type InSb. To determine if the detrimental surface states on p-n junctions could be removed in a similar manner, a glass and Kovar crystat was designed and built. Diodes could be cooled to liquid nitrogen temperatures in a vacuum of less than 10^{-6} Torr. Facilities were provided for leaking purified gases into the system. Diodes previously etched to give good electrical characteristics were mounted in the cryostat and the diode surfaces were bombarded with nitrogen ions, hydrogen ions or electrons at both room temperature and liquid nitrogen temperature. Deterioration of the diode characteristic was rapid in all cases and no improvement resulted from extended periods of bombardment.

The above experiments and observations have demonstrated the sensitivity of diode surfaces and indicated something of the nature of surface states on InSb p-n junctions. However, in the context of this thesis it was considered of greater importance to establish a method of minimizing surface effects rather than a method for obtaining reproducible surface states which could be studied by means of p-n junctions. The results presented earlier indicate that a method for minimizing surface effects has been established but the observations of this section indicate that it

was not always successful. Even when extreme care was exercised in the etching, washing, drying and transferring procedures; the average number of etches required for each diode, in order to obtain characteristics showing no apparent effects of accidental surface contamination, was about four. Experiments on the effects of environmental conditions[237] suggest that this number could have been substantially reduced if the diode preparation procedures could have been carried out in clean room conditions.

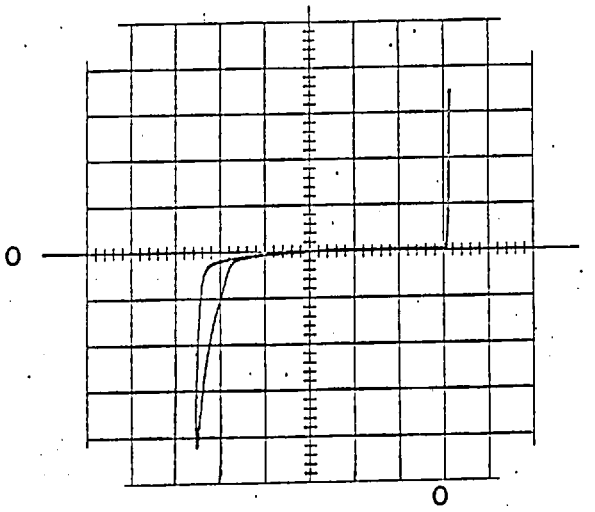
5.5.2. Anomalous Hardening Effect.

The hardening effect was introduced briefly in section 5.3.2. It is simply the tendency for InSb diodes to exhibit a low voltage breakdown which moves to higher voltages as the reverse current is increased. Hardening was accomplished with less danger of deterioration in the diode characteristics if an alternating current was used. It is possible that the high dissipation in the surface layer when hardening was attempted using DC currents tended to cause surface deterioration. For this reason hardening was generally performed on the curve tracer since this utilized an alternating voltage and also provided a means of continuously monitoring the characteristic. Not all diodes hardened in the same manner and the effects of hardening were not always the same. However, for the sake of brevity, the typical hardening process will be described and results not generally typical will be included where they give some additional indication of the possible causes of the hardening effect.

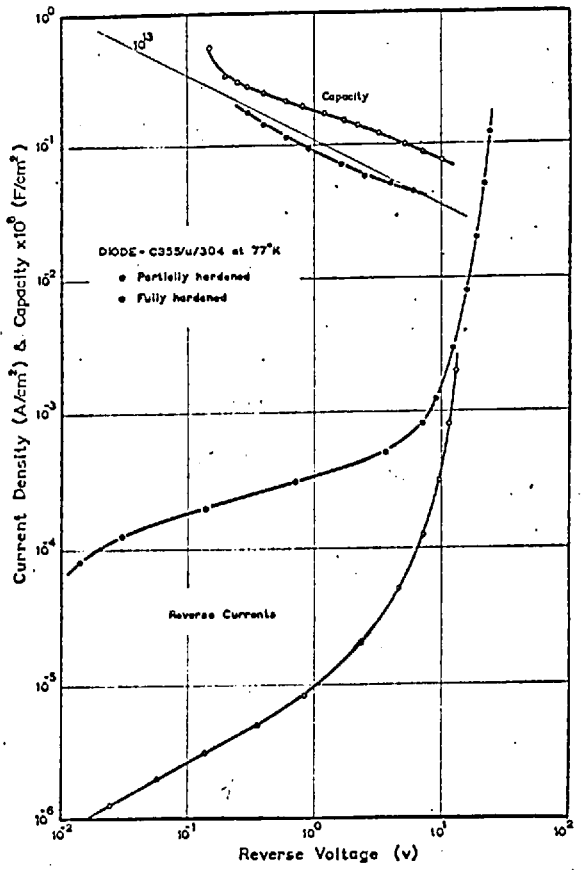
The current-voltage characteristic of a diode photographed

during hardening is shown in Figure 72(a). In general, currents less than 100μ A in the high reverse current region caused the breakdown to move rapidly but somewhat spasmodically out along the voltage axis. When the breakdown reached about three quarters of the limiting breakdown voltage it would literally snap up to give a characteristic very close to the limiting form. The hysteresis in the breakdown region shown in Figure 72(a) was often observed but it was not always so pronounced. This hysteresis in the hardening phenomena would suggest something in the nature of a surface layer of dipoles which move on the surface with a low mobility in the presence of an electric field.

It was mentioned in section 5.3.2. that hardening often affected the DC characteristics. In some instances reverse and forward currents were high before hardening and were reduced to values typical of the bulk following hardening but, in general, currents at low forward and reverse bias were increased by a small amount after hardening. This effect was exaggerated in the case of diodes formed on C355/u and Figure 72(b) shows the sort of change observed from partial hardening, where the low voltage currents were uneffected, to full hardening. The effects on capacity were also exaggerated in diodes formed on this pure material. The capacity before hardening was considerably higher than expected and it decreased towards the expected values after hardening as shown in Figure 72(b). It was further observed that in some cases when the reverse biased DC resistance of the diode was very high, the AC conductivity at 500 kc/s was very high and the capacity was high.



(a)



(b)

Figure 72. The effects of hardening (a) on the current-voltage characteristic during hardening (1 v/div horiz. and 2 x 10^{-6} A/div vert.) (b) on the capacity and reverse current of a $p^{+}n$ diode.

The hardening effect clearly arises from a surface layer with properties which can be considerably affected by chemical and ambient influences not within the control of these experiments. Furthermore the state or effect of this surface layer can be changed near the junction by applying a voltage to the junction and this altered state appears to be stable at 77°K.

Since the nature of this surface effect and those described in the preceeding section are not understood, it is important to ascertain their importance in relation to the reliability of the measurements reported earlier. It has not been possible to prove ipso facto that the characteristics reported for InSb diodes arise entirely from bulk properties. However, the results given in section 5.3. and 5.4. give strong evidence to suggest that bulk characteristics can be observed if the diode preparation procedure described in Chapter 3 is followed and repeated until it is apparent that surface effects have been minimized, and if the hardening procedure is carried out to further reduce the remaining effects to a minimum. The high degree of reproducibility in diode characteristics achieved by this procedure and the excellent agreement obtained between experiment and theory indicate with some certainty that surface effects in all but reverse current measurements can be consistently reduced to negligable levels.

5.6. Effects of High Dislocation Densities.

Dislocations are known[58] to affect the electrical properties of p-n junctions and for this reason experiments were carried out to determine what density of dislocations would begin to influence

diode properties and what effects would result from dislocation densities in excess of this value.

For initial experiments a slice of crystal C299/vac was chosen which had a donor density of $3.2 \times 10^{14} cm^{-3}$ and dislocation densities, measured on areas of about $10^{-2} cm^2$, ranging from $10^3 cm^{-2}$ to about $3 \times 10^4 cm^{-2}$. Dice from this material were photomicrographed both before and after alloying to permit an accurate count of the dislocations in the junction region. For dislocation densities up to $2.5 \times 10^4 cm^{-2}$ no effects which could be related to the dislocations were observed in the diode characteristics. In fact the combined results from four of these diodes were used in the earlier sections of this chapter.

Experiments were then performed on MCPB material. This InSb had a donor concentration of $10^{15} cm^{-3}$ and dislocation densities of about $10^5 cm^{-2}$. Figure 73 shows a DC characteristic, both for forward and reverse bias, typical of diodes formed on this material. These diodes required little or no hardening and in low forward and reverse bias they exhibited a resistive region which was relatively insensitive to chemical treatment. For reverse voltages in excess of 100 mv the current increased rapidly departing from the ohmic behaviour to obey a V^x law over two or three decades. The value of x was a constant for any diode and varied between 2 and 3 from diode to diode. When diodes were etched in CP4A the characteristics remained unchanged except for the resistive region where the currents were generally increased by about one order of magnitude.

The characteristic of a diode formed on similar low dislocation

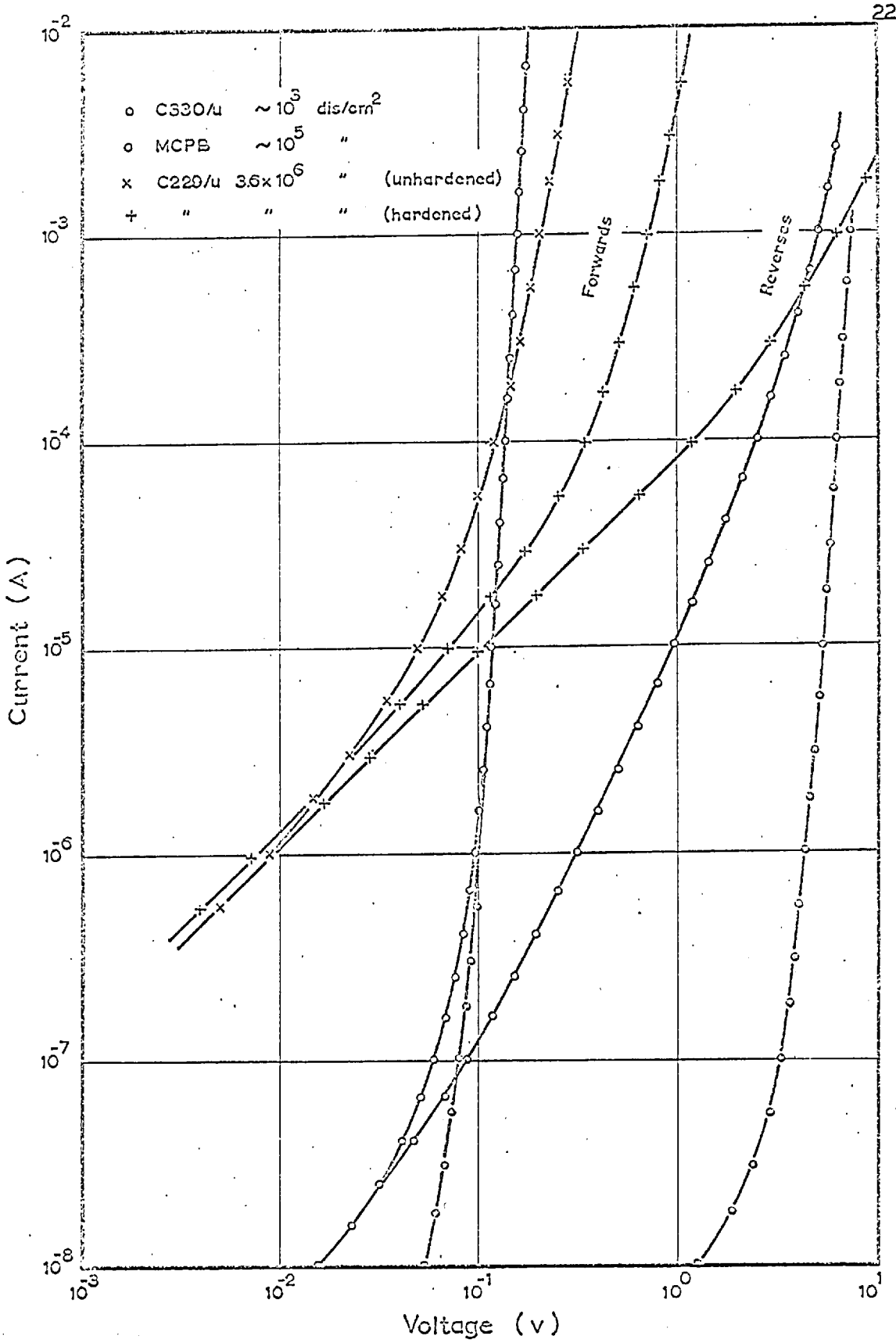


Figure 73. The effects of dislocation density and hardening on the forward and reverse current-voltage characteristics of InSb diodes.

density InSb is also plotted in Figure 73. Comparison of this characteristic with that of the MCPB diode shows that the dislocations produce a large effect in the reverse characteristic but they seriously affect the forward characteristic only at low currents. It is clear therefore that dislocations do affect the properties of InSb diodes but that the effects are not important for dislocation densities less than $3 \times 10^4 cm^{-2}$.

Pulse measurements were also made on the above diodes to determine the carrier lifetime. The lifetime was found to be independent of injection level up to currents of 10 A/cm^2 and it varied considerably from diode to diode. Typical values lay between 3×10^{-8} sec and 9×10^{-8} sec. It is thought that the low values of lifetime and the variation in values were due to dislocations and the variation in dislocation density from diode to diode. If this were the case it appears that dislocation densities in excess of $3 \times 10^4 cm^{-2}$ not only affect the electrical characteristics of diodes but also affect the carrier lifetimes.

The appearance of a well defined current-voltage relationship in the reverse characteristics of dislocated p-n junctions suggested the possibility that diodes could be used to investigate the electronic properties of dislocations. Apart from direct metallurgical observations, studies of the nature of dislocations in semiconductors[238] have been limited to measurements of their influence on carrier mobilities, concentrations and lifetimes. To investigate the possiblity of using p-n junctions in such studies and to determine the effects of high dislocation densities in p-n junctions, a specimen of crystal C229/u was plastically deformed

at 350°C by three point bending to produce single slip.[238] The control specimen (also used for diodes) had a donor concentration of $8 \times 10^{13} cm^{-3}$. The strained specimen had a final donor concentration of $10^{13} cm^{-3}$ and a dislocation density of $3.6 \times 10^{6} cm^{-2}$ on the {111} surface almost perpendicular to the slip plane. The dislocation density in the slip plane was about $2 \times 10^{5} cm^{-2}$. The strained specimen was cut into six dice: four perpendicular to the slip plane and two parallel to the slip plane. Diodes were formed on each of these dice.

Measurements on the diodes with junctions parallel to the slip plane showed that both junctions had metallurgical imperfections and thus no information was gained from these diodes. The measurements on diodes with junctions perpendicular to the slip plane showed some very interesting properties. A typical DC characteristic before and after hardening is plotted in Fig.73. These characteristics bore little resemblance to normal diode characteristics and it was concluded, therefore, that they were determined almost entirely by the presence of dislocations. The resistivity of the ohmic region at low voltages was greatly decreased and the reverse characteristic was essentially ohmic. Although it is not shown in Figure 73, the high current region of the characteristic indicated the presence of increased series resistance. This was one of the few observations which was obviously in agreement with present dislocation theory.

The effects of hardening were unexpected. Hardening caused the forward characteristic to shift with respect to the voltage axis in such a way as to indicate a substantial increase in the

diffusion potential. By limiting the amount of hardening at 77°K, the forward characteristic could be positioned stably at any position between the two extremes shown in Figure 73. The diode would revert to the unhardened state if the temperature was raised above 140°K even for a very short time. It would appear from these observations that the electron occupancy of a dislocation line can be controlled in the region of a p-n junction by the maximum voltage applied to the junction.

Measurements have shown the capacity of these diodes to be essentially independent of voltage but, since the values of capacity were generally about 1pF, the measurements were not considered to be reliable. Pulse measurements have indicated the hole lifetimes to be less than 10^{-8} sec. Since lifetimes are affected by dislocation densities of $10^{5}cm^{-2}$ and the recombination due to dislocations is directly proportional to the dislocation densities, lifetimes this short are to be expected.

The observations reported above can be considered only as initial results in a large area of investigation. However they are sufficient to indicate the very large effects which can arise from dislocations and to further indicate that some rethinking about the nature of dislocations is required before such measurements can be understood.

5.7. Preliminary Work on $n^{+}p$ Diodes.

Part of the objectives of this work was to assess the feasibility of transistor fabrication in InSb. Since only n-p-n structures utilise the high electron mobility in InSb and narrow

p-type base regions are necessary, the early work concentrated on the possibility of using a diffused layer for the base region. When this was shown by the diffusion work described in Appendix A and the work of Henneke[65] not to be entirely satisfactory, the investigations were concentrated on the formation of alloyed junctions. Because of the availability of n-type InSb, the development of alloying techniques and the initial measurements were performed on p^+n junctions. The results of this work form the main body of this thesis. However, interest in n^+p junctions and in the possibility of double-alloyed n-p-n transistor structures continued throughout the work and the results of some preliminary investigations with specific reference to n^+p junctions are described in this section.

5.7.1. Sn - InSb Liquidus.

When alloyed junction structures are to be prepared where the junction depth must be accurately controlled, it is necessary to know the liquidus curve of the appropriate metal-semiconductor system. In elemental semiconductors, where most of the binary phase diagrams of the metal-semiconductor systems are well established, this information is readily available. In compound semiconductors, however, the relevant ternary phase diagrams are for the most part unknown.

In the case of indium antimonide the choice of suitable metal solvents is limited. Since reasonably pure indium antimonide is intrinsic at room temperature, most devices operate only at low temperatures. Junctions must, therefore, be formed by metals with

either a low solid yield strength or the same thermal expansion coefficient as indium antimonide. Further, the metal must not form a compound, other than indium antimonide, with indium or antimony and in the pseudo-binary metal-InSb system and it must have a melting point well below that of indium antimonide. Metals which satisfy these conditions are few, but among them are indium and tin.

Indium, with small percentages of silver for ohmic contacts or with small amounts of zinc or cadmium for p^+n junctions, has been used throughout this work. Small amounts of selenium or tellurium used for n-type doping appear to change the properties of indium considerably and it was felt that tin, which is a donor[239] in InSb, might be more suitable for making n^+p junctions. However, before tin could be used to form transistor structures it was necessary to determine the solubility of InSb in molten tin.

Dice of InSb and pellets of pure tin were prepared by the methods described in Chapter 3. The pellets were weighed on a micro-balance to the nearest microgram. The dice were cleaned but not etched, in order to preserve the flat polished surface, and pellets were wetted to the dice by the radiation wetting technique. The dice with pellets wet on one or both surfaces were subsequently transferred to a tunnel furnace where they were alloyed in a hydrogen atmosphere at temperatures ranging from 248°C to 430°C. To prevent any oxidation of the tin during alloying, the hydrogen was purified by passing it through an activated charcoal trap at 78°K. Both the heating and cooling rates of the alloying furnace were kept below 5°C per minute in

order to produce planar junctions and to promote good regrowth. The maximum alloying temperature of the cycle was maintained for approximately ten minutes to ensure that equilibrium was achieved and that at the same time solid state diffusion was negligible. At the highest temperature used it is estimated that diffusion could cause an error of not more than 0.5 microns in the penetration depth.

After alloying the wetted area was photomicrographed to facilitate its measurement. The junctions were potted in epoxy resin and sectioned in the centre of the wetted area to minimize any misalignment errors. The junctions were delineated by etching for a few seconds in CP4A diluted with acetic acid. The mass of dissolved indium antimonide was calculated from the wetted area and the junction cross-sectional area. It was estimated that the mass thus determined was accurate to better than five per cent. The weight per cent of indium antimonide X(T) dissolved by the tin at the alloying temperatures was then calculated. The resulting values of X(T) are plotted in Figure 74.

In order to establish the validity of the experimental method, measurements on indium - indium antimonide junctions were made under similar conditions. In Figure 75 the results of these measurements are plotted on the known[6] liquidus curve of the In-Sb system. The agreement is well within the estimated experimental error.

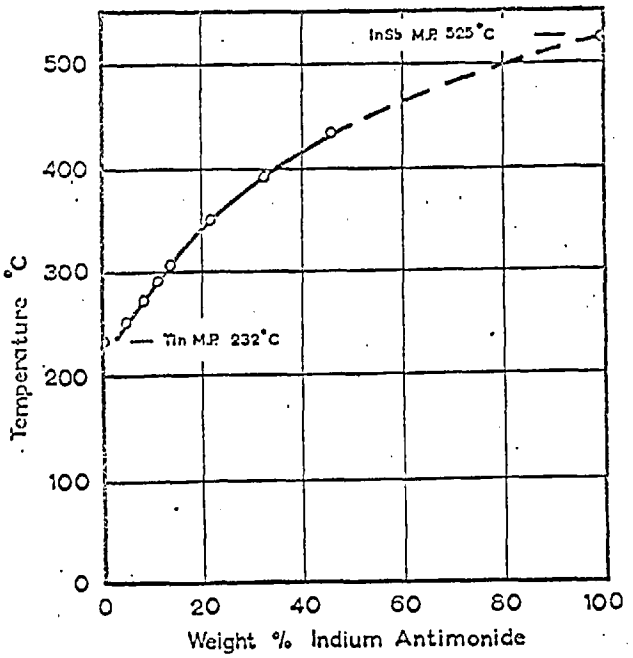


Figure 74. The tin-indium antimonide liquidus curve.

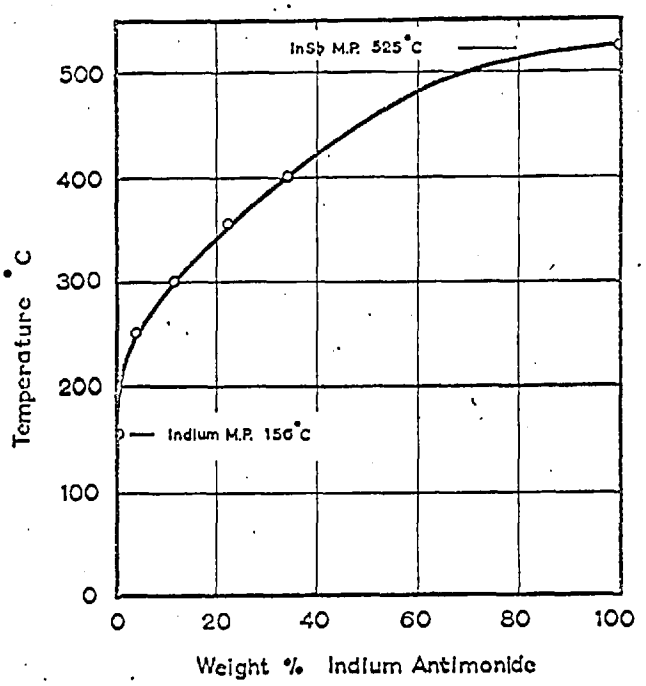


Figure 75. Measurements of the indium-indium antimonide liquidus plotted on the known In-Sb phase diagram[6].

5.7.2. Sn - InSb Junctions.

During the course of the entire work described in this thesis, certain observations have been made which indicate the presence of strain in tin - indium antimonide junctions. These are listed below.

(a) During sectioning the regrowth regions of tin - InSb junctions tended to crack much more readily than the remainder of the crystal. Similar behaviour was not observed in indium - InSb junctions.

(b) For the purpose of electrical measurements on p-n junctions formed using tin pellets on p-type indium antimonide, connections were made to the tin pellet using 0.002 inch diameter silver wire. In a number of these junctions small lateral forces applied to the pellet during the connection procedure were sufficient to cause a section of the original crystal at the edge of the regrowth region to chip out. This would seem to indicate that the crystal surrounding the regrowth region was strained in compression.

(c) On etching one of these Sn - InSb junctions for electrical tests the regrowth region with the tin pellet intact was completely separated from the crystal.

(d) Electrical tests on a small number of p-n junctions formed with tin have yielded lower breakdown voltages and higher reverse currents than similar indium junctions.

(e) Pure tin was used to alloy a large number of dice to headers. During the subsequent experiments, which in each case involved several temperature cycles between 300°K and 77°K, it

was found that approximately 10% of the dice became detached from the headers having cracked in the tin - indium antimonide junction region.

Although no particular effort has been made as yet to investigate these observations and explanations other than strain might be found for certain cases, there seems to be definite evidence of strain in at least some tin - indium antimonide junctions. In some respects this is surprising since tin, as a donor, must substitute to a large extent for indium which has a larger tetrahedral radius than tin.(18) Also the expansion coefficient of tin is much greater than that of indium antimonide(10) so that, even if the solid yield strength of tin were high enough to cause serious strain, it would be expected to cause tension at the crystal surface. The measurements of Kolm et al(240) have indicated a decrease in the lattice parameters of InSb containing small percentages of tin. This would also be expected to give rise to tension rather than compression in the situation described above. One point which might be relevant is the existance of a eutectic or peritectic at the high tin concentration end of the liquidus. Such a eutectic or peritectic would almost certainly have a higher solid yield strength than pure tin, however, the thermal expansion coefficient would not be expected to differ markedly from that of pure tin.

Although the reasons for the above observations are not known it can be concluded that the evidence of strain in at least some tin - indium antimonide junctions must cast doubt on the usefulness of tin for alloyed devices in indium antimonide.

CHAPTER 6

Conclusion.

It is neither desirable nor practical to restate all the conclusions drawn in the foregoing chapters. However, in this chapter some of the more important conclusions and contributions of this work will be reviewed briefly in the context of the objectives stated in Chapter 1.

Of first importance, because the possibility of any meaningful study of p-n junctions is determined by it, was the development of a reliable technique to form p-n junctions in InSb. The use of atomic or ionic hydrogen to obtain uniform wetting is considered to be a significant advance in the technology of InSb. It has been demonstrated that uniform wetting with indium can be achieved at temperatures as low as 200°C in the presence of atomic or ionic hydrogen generated by a hot tungsten filament or a high voltage discharge. It has further been demonstrated that this technique can be extended for use on other higher melting point semiconductors.

The electrical characteristics of InSb p^+n junctions formed by the above technique have been measured as a function of doping density, temperature, crystallographic orientation, dislocation density, surface treatment and metallurgical processing. The dependence on doping density and temperature has been interpreted in terms of the properties of the base crystals and the regrowth regions. No inherent effect of crystallographic orientation has

been observed but surface effects appear to be orientation dependent. Dislocation densities in excess of $3 \times 10^4 cm^{-2}$ appear to affect both the carrier lifetime and the characteristics of p-n junctions. Metallurgical effects on the electrical characteristics only arise if uniform wetting has not been achieved or if alloying temperatures rise at a rate in excess of $1^{\circ}C/sec$.

Consideration of medium level injection in p-n junctions has led to an analytic expression for the current-voltage relationship in this region. When combined with the existing expressions for high level and low level injection, the theory accurately describes the current-voltage behaviour of InSb p^+n diodes. It also gives an accurate description of the effects of doping densities, carrier lifetimes and mobilities, and temperature on the characteristics of such junctions.

Measurements of the space charge generation currents in reverse biased InSb diodes have indicated the presence of a trapping level at either 0.133 ev or 0.120 ev (extrapolated $0^{\circ}K$ values). Comparison with the trapping level obtained from independent lifetime measurements[4][5] suggests that the trapping level is 0.133 ev with a temperature dependence of -1.55×10^{-4} ev/ $^{\circ}K$. The values of carrier lifetime measured by the pulse response of InSb diodes are in reasonable agreement with those determined by p.e.m. and p.c. techniques.

The reverse breakdown of InSb p^+n junctions has been measured. It is orders of magnitude higher than simple considerations using bulk breakdown measurements would suggest. The ionization coefficient as a function of electric field was calculated from

these measurements and used to show that more thorough considerations indicate the values of breakdown voltage to be correct. A comparison has also been made with Si and Ge p-n junctions which confirms this conclusion.

An investigation of the effects of crystallographic direction on the formation of alloyed p-n junctions led to a number of important observations. It was found that both the {100} and the {111} surfaces are facet forming in InSb while only the {111} surfaces form facets in germanium. Direct evidence was obtained to show that A {111} surfaces dissolve more rapidly during alloying than do B {111} surfaces. A model for {111} surfaces similar to that proposed by Gatos and Lavine[27] has been used to explain the observations on {111} surfaces. A model has been proposed for the {100} surfaces which explains the appearance of {100} facets in polar semiconductors. Using these models it has also been possible to present an explanation for many of the metallurgical polar effects observed in III - V compounds.

A method has been developed for measuring the liquidus curves of metal-semiconductor systems using p-n junctions. The Sn - InSb liquidus has been measured between 250°C and 430°C using this technique.

The results of investigations of diffusion in InSb indicate that there are problems arising from high surface concentrations, non-error function diffusion and thermal conversion which must be understood and controlled before successful diffused transistors can be made. In this context, Henneke's[65] work must be

considered an impressive achievement. Double-alloyed transistors are certainly feasible in InSb. Since alloyed p-n junctions appear to exhibit bulk properties and the alloying process does not significantly affect the properties of the base crystal, it is possible that double-alloyed transistors could be used to answer some of the questions left unanswered by Henneke's alloy-diffused transistor. The effects of increasing electron lifetime as a result of increasing electron concentration in p-type base regions could also provide a basis for some interesting experiments not only with n-p-n transistors but also with n^+p diodes.

Appendix A - Diffusion and Thermal Conversion:

A.1. Materials and Techniques.

The principles of the p-n junction method in diffusion experiments have been treated by Fuller.(81) In these experiments boat grown InSb with donor concentrations of $10^{15}cm^{-3}$ for crystals A and B and $6 \times 10^{15}cm^{-3}$ for crystal C were used. The crystals were cut into specimens 10 x 3 x 1 mm and one large face was lapped to a flat mirror finish. Specimens were cleaned by degreasing in Analar trichlorethylene, etching in Analar HCl, washing throughly in 14MΩ -cm deionized water and drying under vacuum. The 5mm diameter silica diffusion tubes were cleaned by etching in HF, rinsing thoroughly in deionized water and firing at about 800°C under vacuum. The specimen and diffusant were then placed in the silica tube and evacuated for not less than one hour at a pressure of less than 10^{-5} Torr. The tube was then sealed off to give short ampoules 8 cm long or long ampoules 24 cm long. These ampoules were placed in a two temperature zone furnace at the desired temperature for the appropriate diffusion time.

After diffusion a flat, mirror finish, 5° bevel was ground on the polished surface of the specimen. The specimen was then mounted on a heat sink and cooled to less than 100°K in a dry nitrogen atmosphere. A pointed tungsten wire maintained at approximately room temperature was used to probe the bevelled surface. The thermoelectric voltage between the cold semiconductor and the hot probe was used to determine the conductivity type and/or the position of the p-n junction. The position of the

junction was marked on the bevel and the penetration depth was measured by a microscope interferometer.

A.2. Thermal Conversion.

During initial diffusion work it was found that specimens were occasionally converted completely from n-type to p-type in diffusion times as low as one hour. The following experiments were carried out either to determine the cause of this rapid uptake of impurity or to determine conditions under which it would be effectively suppressed. The procedures outlined in A.1 were adopted except where alterations were made for experimental purposes. Before bevelling and thermal probing, Hall measurements were made on the specimens. The results of these experiments are tabulated in table A.1.

In experiment I the effects of tube cleaning procedures and possible contamination from the vacuum system were examined. High purity quartz tubes were also tried but, since these tubes were 10 mm in diameter and some difficulty in sealing them off was experienced, the results obtained are doubtful. It is important to note that the inclusion of zinc seemed to prohibit conversion.

Previous experiments had shown that in short ampoules surface alloying was difficult to prevent, particularly in cadmium diffusion. Experiment II was carried out using cadmium diffusant in long ampoules in a two-zone furnace to determine if, under some conditions, both surface alloying and thermal conversion could be avoided. No such conditions were found.

In experiment I a temperature dependence of the impurity

Crystal No.	Temp. (°C) Low	Temp. (°C) High	Time (hrs.)	Hall conc. p-(cm^{-3})	Probed	Remarks
Experiment 1 - Short tubes						
$B6_1$	-	428	1	2.6×10^{15}	p	Control spec.
$B7_1$	-	428	1	2.6×10^{15}	p	Aqua regia etch
$B9_1$	-	428	1	2.6×10^{15}	p	" " "
$B8_1$	-	460	1	4.9×10^{15}	p	" " "
$B8_2$	-	460	1	4.2×10^{15}	p	" " "
$B6_2$	-	428	1	1.7×10^{15}	p	Double cold trap
$A5_1$	-	428	1	1.4×10^{15}	p	Quartz " "
$A5_2$	-	428	1	7.7×10^{15}	p	H.P.Quartz tube
$A4_2$	-	428	1	5.2×10^{15}	p	" " "
$A2_2$	-	428	1	Junction	15 μ	Zinc included
Experiment II - Long tubes, cadmium included						
$B7_3$	273	494	1	6.9×10^{15}	p	
$A1_1$	266	435	16	7.1×10^{15}	p	
$C2_1$	300	500	15	1.2×10^{16}	p	
$C2_2$	320	500	15	7.4×10^{16}	p&n	
$C2_3$	352	500	14	3.7×10^{16}	p	Sl.surface alloy
$C2_4$	375	500	$14\frac{1}{2}$	-	p&n	Surface Alloyed
Experiment III - Long tubes, no diffusant						
$A1_0$	273	494	$13\frac{1}{2}$	7.7×10^{15}	p	
$B9_2$	"	"	1	-	p	
$B7_2$	"	"	1	4.2×10^{15}	p	
$B10_1$	"	"	1	4.2×10^{15}	p	$\frac{1}{3}$ atm of H_2
$B10_2$	"	"	1	7.5×10^{15}	p	Spec supported in Tungsten coil

Table A.1 Results of Thermal Conversion Experiments.

uptake was observed. In experiment III a similar dependence on diffusion time was observed. Hydrogen was introduced into one ampoule to determine if the partial pressure in the tube would have any influence on the degree of conversion. None was observed. Finally, in an attempt to determine if conversion was affected through the vapor phase rather than through contact with the tube walls, the specimen was supported in a tungsten coil. It appears that only increased system contamination was achieved.

In conclusion, thermal conversion appears to be the result of a rapidly diffusing impurity which cannot easily be eliminated. Thermal conversion is both time and temperature dependent and can be greatly reduced by the presence of a diffusant such as zinc. In cadmium diffusion it does not seem possible to preclude both surface alloying and thermal conversion by the use of a two zone diffusion system.

Since this work Hulme[3] has published results of similar but much more exhaustive experiments on thermal conversion. The degree and nature of thermal conversion observed, agree with those reported here. Copper, Flourine, Lithium or sodium are suspected to be possible rapidly diffusing impurities.

A.3. Zinc and Cadmium Diffusion.

Both zinc and cadmium diffusions were carried out in short ampoules following the techniques described in A.1. In the zinc diffusions less than two milligrams of spectrographically pure zinc was included in the ampoules in order to avoid surface alloying. Measurements were done at a diffusion temperature of 368°C to

complement previous measurements done in the laboratory.[90] The results are shown in Figure 76. Since diffusion from a vapor source should give an error function impurity distribution,[80] the junction depth would be expected to vary as the square root of the diffusion time, t. The observed junction depth varied as $t^{0.62}$ and therefore non-error function diffusion was indicated.

Early cadmium diffusion experiments showed that even when the cadmium was limited to less than one milligram in the diffusion tube, surface alloying could not be avoided. To determine the feasibility of diffusing from indium-diluted cadmium sources, alloys of indium containing from 40% to 3% cadmium were made up and used as diffusion sources. All alloys with concentrations above 7% caused severe surface alloying unless the source was very limited. The results obtained by using small amounts of indium containing less than 7% Cd are plotted in Figure 76. The effects of alloy composition on diffusion are also illustrated. The pure Cd experiments were carried out at 428°C to avoid heavy surface alloying. These have been normalized to 500°C at 20 hours using the results obtained by Wilson.[90] It is apparent that there is some diffusion dependence on the cadmium vapor pressure in the ampoule.

In conclusion it appears that using a carefully measured amount of an accurately prepared alloy, reproducible diffusion of cadmium or zinc in InSb could be achieved. However if such diffused layers were used for transistor structures the high surface concentrations, which seem unavoidable, would result in tunnel emitters.

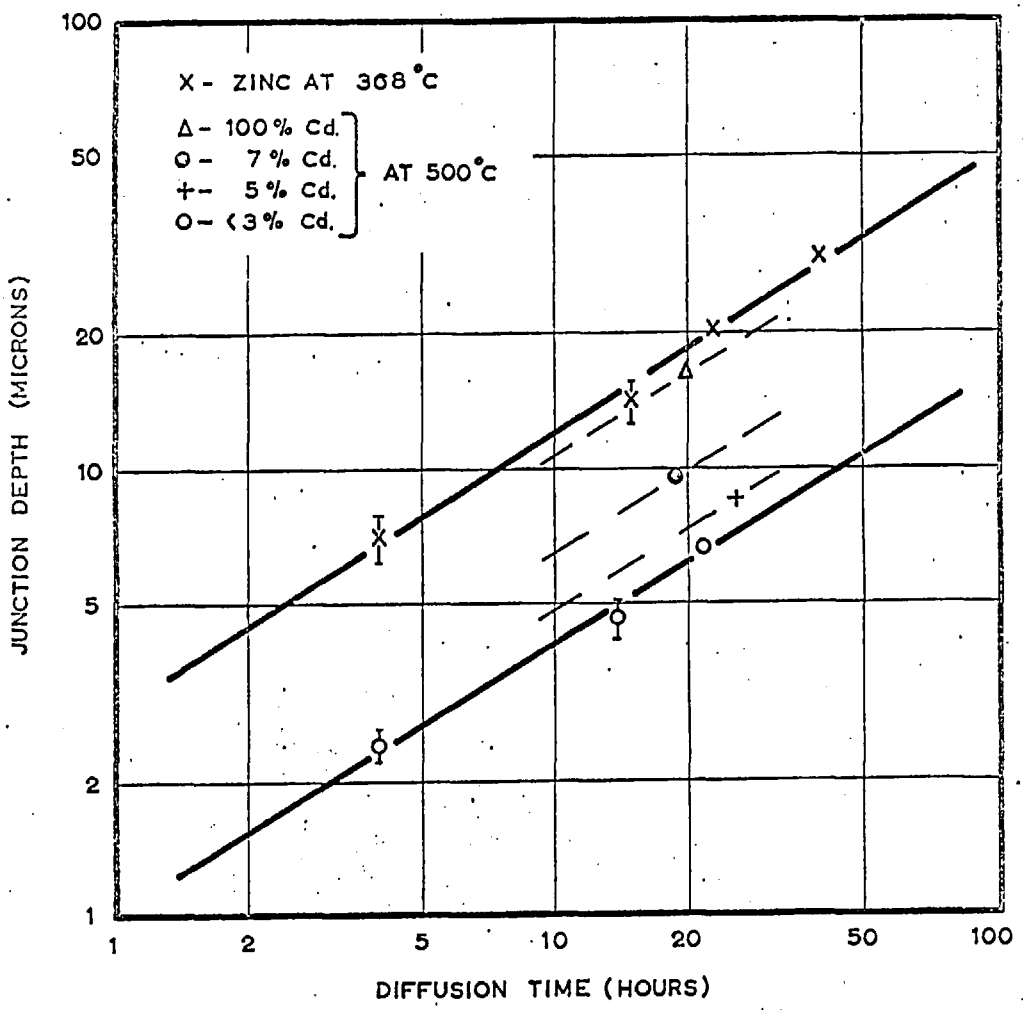


Figure 76. The diffusion of zinc and cadmium into InSb measured by the p-n junction technique as a function of diffusion time and, in the case of cadmium, as a function of diffusant composition.

Appendix B - Theory of Diffusion Currents.

B.1. General Formulation.

Equations (2.1) - (2.5) apply generally to the neutral n and p regions on either side of the space charge region assumed to be at $x = 0$.

Equations (2.1) and (2.2) can be solved for $d\psi/dx$,

$$\frac{d\psi}{dx} = \frac{J_p + qD_p \frac{dp}{dx}}{qp\mu_p} = \frac{J_n - qD_n \frac{dn}{dx}}{qn\mu_n} \quad \ldots\ldots (B.1.)$$

However, outside the space charge region, neutrality and the Einstein relation give from (B.1.)

$$\frac{dn}{dx} = \frac{dp}{dx} = \frac{J_n pD_p - J_p nD_n}{qD_n D_p (n + p)} \quad \ldots\ldots (B.2.)$$

If a single recombination level is assumed the continuity equations become:

$$\frac{dJ_n}{dx} = \frac{-dJ_p}{dx} = \frac{q(np-n_i^2)}{(n+n_t)\tau_p+(p+p_t)\tau_n} \quad \ldots\ldots (B.3.)$$

B.2. Low Level Case.

For low level injection in the n-region of a p-n junction $\Delta n = \Delta p \ll n_{no} > p_t$, n_t and p_{no}. Thus equation (B.3.) reduces to:

$$\frac{dJ_p}{dx} = \frac{q(p - p_{no})}{\tau_p} \quad \ldots\ldots (B.4.)$$

and equation (B.2.) reduces to:

$$\frac{dp}{dx} = - \frac{J_p}{qD_p} \qquad \ldots\ldots (B.5.)$$

By differentiating (B.5.) and substituting from (B.4.)

$$\frac{d^2p}{dx^2} = \frac{p - p_{no}}{\tau_p D_p} = \frac{p-p_{no}}{L_p^2} \qquad \ldots\ldots (B.6.)$$

where L_p is the diffusion length for holes. This equation is solved subject to the Boltzmann boundary conditions,

$$p - p_{no} = P_{no} \left(\exp \left(\frac{qV}{kT}\right) - 1\right) \qquad \text{at } x = 0$$

$$p - p_{no} = 0 \qquad \text{at } x = \infty$$

to give

$$p-p_{no} = p_{no}\left(\exp \frac{qV}{kT} - 1\right) \exp - \frac{x}{L_p} \qquad \ldots\ldots (B.7.)$$

By substituting (B.7.) into (B.5.)

$$J_p = \frac{qD_p}{L_p} p_{no} \left(\exp \frac{qV}{kT} - 1\right) \qquad \text{at } x = 0 \qquad \ldots\ldots (B.8.)$$

A similar treatment for the p-region gives:

$$J_n = q \frac{D_n}{L_n} n_{po} \left(\exp \frac{qV}{kT} - 1\right) \qquad \text{at } x = 0 \qquad \ldots\ldots (B.9.)$$

Since there must be overall current continuity

$$J = J_n + J_p = q \left[\frac{D_p}{L_p} p_{no} + \frac{D_n}{L_n} n_{po} \right] \left(\exp \frac{qV}{kT} - 1 \right) \qquad (B.10)$$

B.3. High level case.

For high level injection in the n-region of a p^+n junction $p = n = \Delta n = \Delta p > n_{no} > n_t, p_t, p_{no}$. Thus equation (B.3.) reduces to:

$$\frac{dJ_n}{dx} = \frac{-J_p}{dx} = \frac{qp}{\tau_p + \tau_n} \qquad \ldots\ldots (B.11)$$

and equation (B.2.) reduces to:

$$\frac{dp}{dx} = \frac{J_n D_p - J_p D_n}{2q\, D_n\, D_p} \qquad \ldots\ldots (B.12)$$

By differentiating (B.12) and substituting (B.11)

$$\frac{d^2p}{dx^2} = \frac{p(D_n + D_p)}{2D_n D_p (\tau_n + \tau_p)} \qquad \ldots\ldots (B.13)$$

This equation is solved subject to the high level Boltzmann boundary conditions,

$$p = n_i \exp \frac{qV}{2kT} \qquad \text{at } x = 0$$

$$p = p_{no} \simeq 0 \qquad \text{at } x = \infty$$

to give

$$p = n_i \exp \frac{qV}{2kT} \cdot \exp - \frac{x}{\sqrt{2}\, L_o} \qquad \ldots\ldots (B.14)$$

where $L_o^2 = \dfrac{D_n D_p (\tau_n + \tau_p)}{D_n + D_p}$

By using (B.14) to integrate (B.11) and recognizing that

$$J_p \simeq 0 \qquad \text{at } x = \infty$$

$$J_p = q \sqrt{\frac{D_n D_p}{(D_n+D_p)(\tau_n+\tau_p)}} \sqrt{2}\, n_i \exp \frac{qV}{2kT} \qquad \ldots \text{(B.15)}$$

at $x = 0$.

In the p^+ region $p_{po} \gg n_{no}$ and therefore low level conditions still apply and the electron current at $x = 0$ is given by (B.9.) The total current for a p^+n junction is, therefore,

$$J = J_n + J_p = q \sqrt{\frac{D_n}{\tau_n}}\, n_{po} \left(\exp \frac{qV}{kT} - 1\right) + q \sqrt{\frac{D_n D_p}{(D_n+D_p)(\tau_n+\tau_p)}}$$

$$\sqrt{2} n_i \exp \frac{qV}{2kT} \qquad \ldots \text{(B.16)}$$

For an n^+p junction the first term of (B.16) is given by (B.8.)

Appendix C - High Injection Step Response.

Consider the n region of a p^+n junction in steady state forward bias and assume that at $x = 0$, $J_p \gg J_n$. Then equation (B.12) from appendix B reduces to:

$$\frac{dp}{dx} = - \frac{J_p n}{qD_p(n+p)} \qquad \ldots \text{(C.1.)}$$

With the use of equation (C.1.) the hole continuity equation (2.4.) becomes:

$$\frac{dp}{dt} = (g-r) + D_p \frac{(n+p)d^2p}{n\,dx^2} \qquad \ldots\ldots (C.2.)$$

since it can be shown that

$$\frac{dp}{dx}\frac{d}{dx}\left(\frac{n+p}{n}\right) < \left(\frac{n+p}{n}\right)\frac{d^2p}{dx^2}$$

If the recombination rate is determined by a single trapping level then at injection levels where $\Delta n = \Delta p = p > n_t,\ p_t,\ p_{no}$.

$$(g-r) = \frac{np}{\tau_p n + \tau_n p} \qquad \ldots\ldots (C.3.)$$

Equation (C.2.) can then be reduced to

$$\frac{dp}{dt} = \frac{p}{\tau^*} + D^* \frac{d^2p}{dx^2} \qquad \ldots\ldots (C.4.)$$

where

$$\tau^* = \frac{n + p(K+1)}{K+1} \qquad D^* = \frac{D_p(K+2)}{K+1}$$

and

$K = \frac{\Delta p}{n_{no}}$ the ratio of injected minority carrier density to the majority carrier density.

Appendix D - Shot Tower.

The small spherical pellets used for alloying were formed using the shot tower illustrated in Figure 77. For successful formation of a specified size of pellet, the shape of the quartz nozzle was very important. A 5 mm diameter quartz tube was drawn

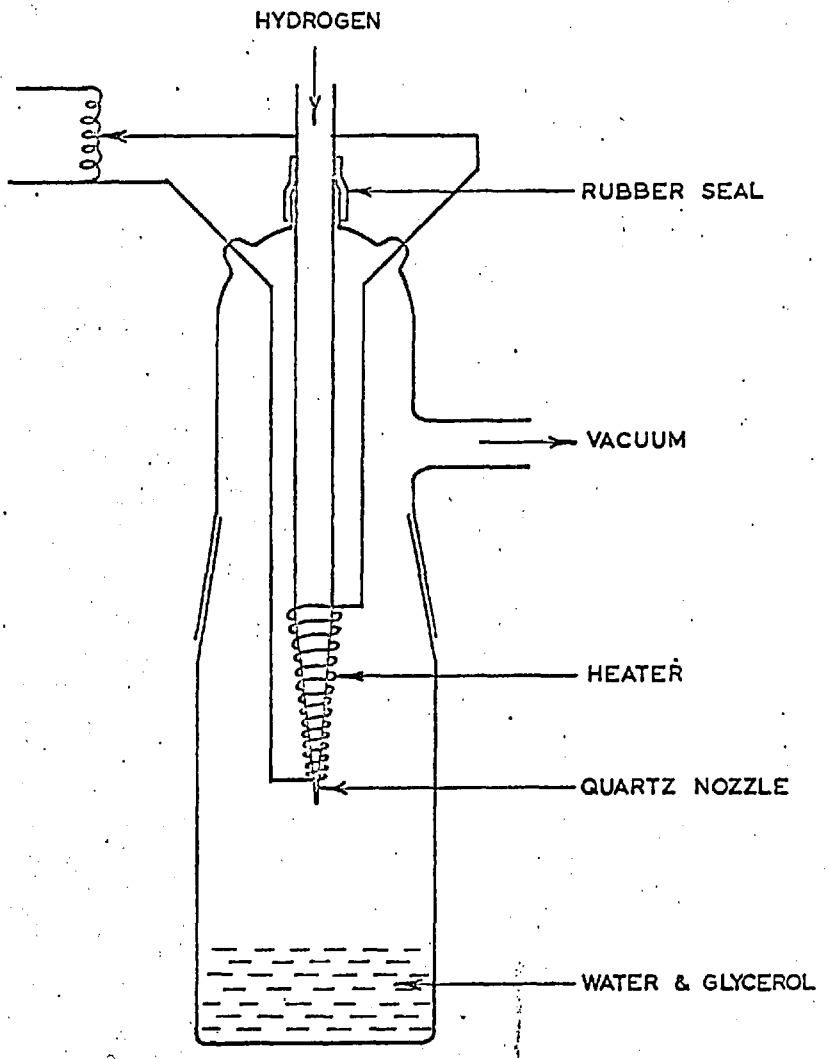


Figure 77. The shot tower for forming small spherical metal pellets.

down at one end to give a long uniform taper. This end was then ground back until the inside diameter at the tip was about 0.3 mm. Such a nozzle yielded a spectrum of pellet sizes varying from 10 to 1000 microns in diameter. The distribution was quite sharply peaked however at about 100 microns diameter.

To form a batch of pellets the nozzle was loaded with about 300 mg of indium or tin and the whole system was flushed several times with hydrogen. With the system hydrogen filled at a slightly positive pressure, the metal was heated about 100°C above the melting point and the vessel was quickly evacuated. The positive pressure of hydrogen in the quartz tube forced the molten metal through the nozzle, atomizing it into small pellets. The water and glycerol below the nozzle served to cool the pellets and reduce their velocity thus preventing the change in shape which would result if they were fired directly onto the bottom of the vessel. The alternative procedure of initially evacuating the system and then firing the molten charge by rapidly applying a positive hydrogen pressure was also used successfully. Since any metal oxide formed during the process remained in the nozzle after firing, it was not necessary to take stringent measures to avoid oxidation of the charge during heating and firing. The pellets were easily sorted for size and perfection by sieving and by selection under a microscope.

Appendix E - Capillary Alloying.

Capillary alloying[189][190][191] is a non-equilibrium alloying technique in which indium is heated to a desired temperature and forced up a fine capillary into contact with a wafer of semiconductor. The advantages of the technique arise largely from the fact that a fresh pure indium surface is being continually exposed until it touches the crystal and therefore wetting problems are greatly reduced. It was for this reason that the following experiments were carried out.

In all, five capillary alloying systems were designed and built. The first four used glass capillaries and from these the following points became apparent.

(a) Above the melting point indium oxidizes rapidly even if the oxygen concentrations are low.

(b) The compressibility of a gas or liquid used to force indium up a capillary greatly reduces the control over the indium column.

(c) Indium is itself compressible and therefore the reservoir from which indium is forced up the capillary must be kept small.

(d) The system must be made of material which indium does not wet if good control over the indium movement is to be realized.

With these considerations in mind the capillary alloying system shown in Figure 78 was built. The crystal holder was positioned by means of a micromanipulator and the small piston driven by a micrometer gave excellent control over the indium extrusion. The operating chamber was continually flushed by a

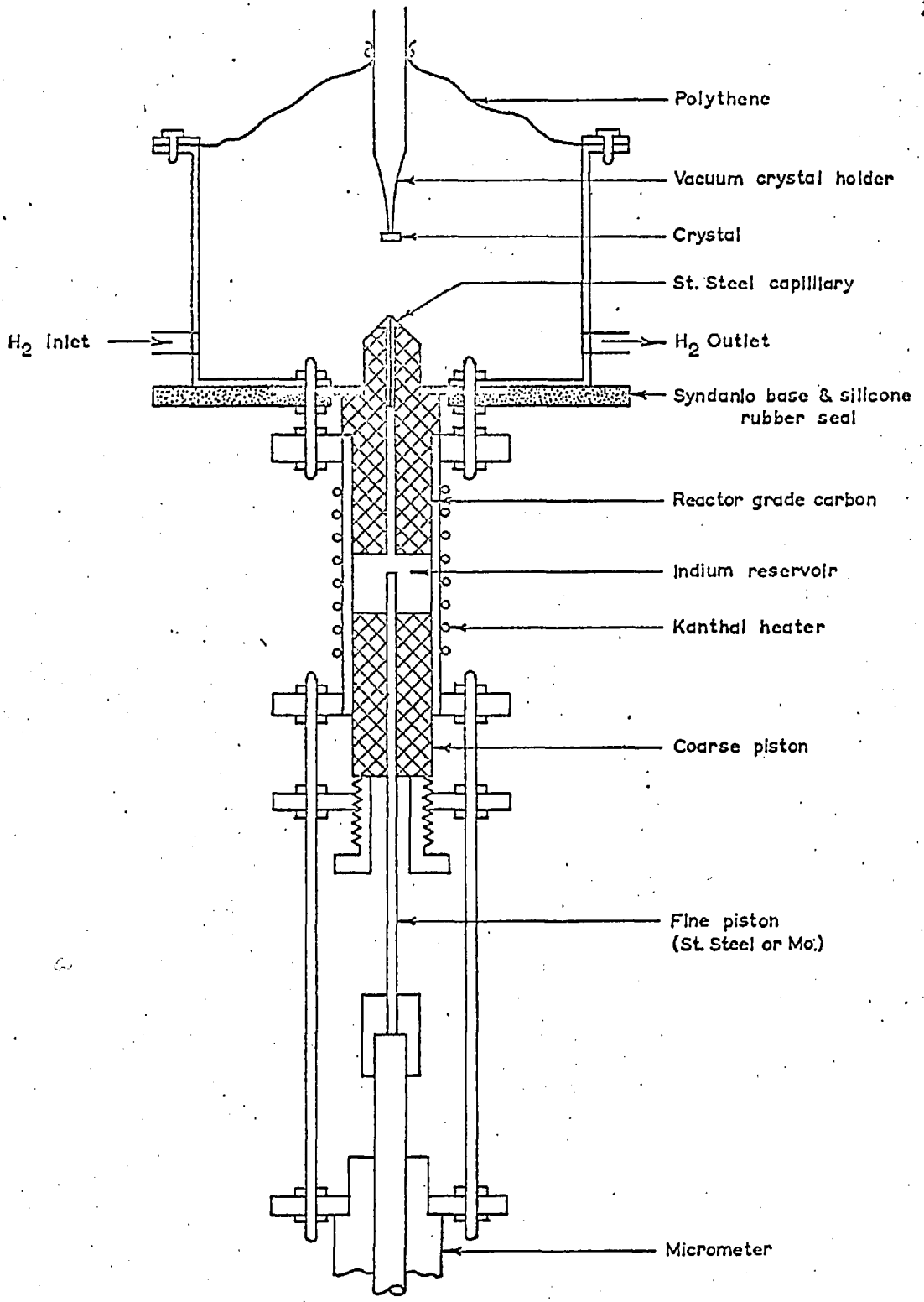
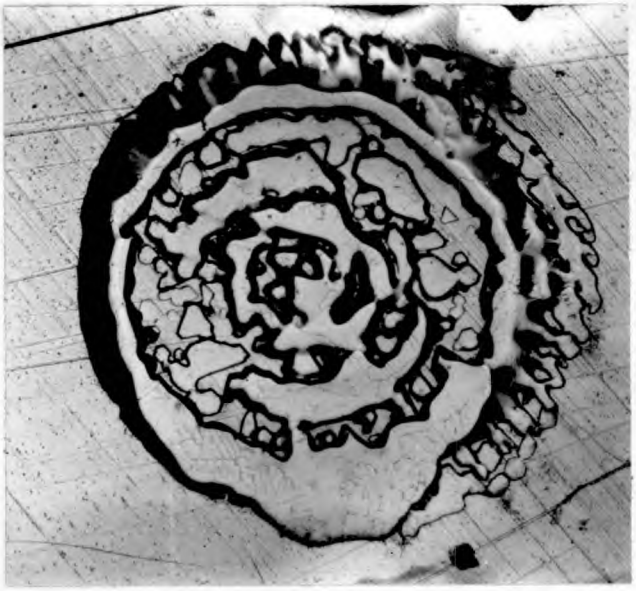


Figure 78. The capillary alloying apparatus.

flow of highly purified hydrogen maintained at a positive pressure by a silicon oil bubbler. Small quartz trays were included in the chamber so that any number of InSb dice could be used for an experiment without having to open the system to the air. The hydrogen was purified as follows. It was first passed through a continuously regenerated chromic chloride solution to remove any oxygen. The chromic chloride cell is described by Schuit and de Boer[192] who found that it removed oxygen below the limits of detection by a mass spectrometer. The hydrogen then passed through two series drying towers containing silica gel and magnesium per chlorate. Immediately before reaching the capillary alloying system it passed through an activated charcoal trap cooled to 77°K. Except for the activated charcoal trap which was frequently baked under vacuum, the rest of the system had a small flow of hydrogen at a slight positive pressure maintained during the entire course of experiments to ensure that no contamination could occur.

Alloying with this system was attempted on a large number of specimens. Indium temperatures ranged from 180°C to 200°C with contact times around 3 minutes. The junctions were examined by removing the indium in dilute HCl. In general the wetted areas enclosed concentric circle-like regions of unwetted crystal as shown in Figure 79(a). When indium was forced out past the crystal, causing a washing action, uniform wetting occurred as is shown in Figure 79(b). These results were easily reproduced and seemed independent of the procedures used for crystal cleaning or



(a)

(b)

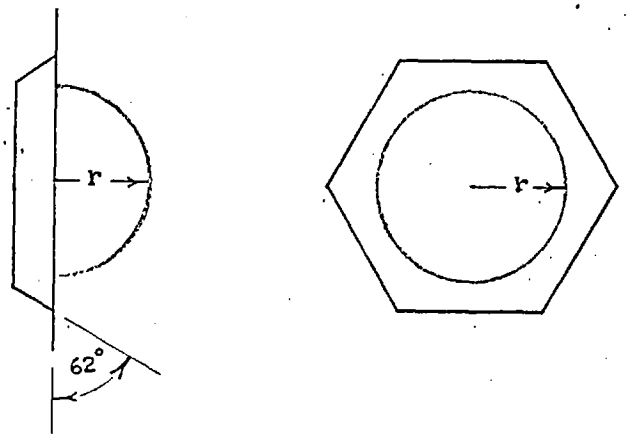
Figure 79. Capillary alloyed areas after the removal of the indium in dilute HCl (a) normal region showing concentric unwetted areas (x100) (b) completely wetted region after indium was flushed out towards the lower right hand corner (x150).

preparation. It was concluded therefore that the wetting difficulties were associated with the indium. This was further confirmed by the observation that when a sphere of indium was extruded at the tip of the capillary, and then drawn back into the capillary it appeared to have a very thin but tenacious skin. The concentric areas in Figure 79(a) are almost certainly associated with the manner in which this skin breaks up as the expanding indium surface is forced against the crystal. It seems most probable that, in spite of the precautions taken to purify the atmosphere, an oxide film was forming on the indium. More important however is the fact that a surface layer, much too thin to perceptibly change the shiny metallic lustre of the molten indium, was sufficient to prevent wetting on InSb.

Further experiments involving the use of fluxes were carried out and problems of a different nature arose which prevented successful alloying. However, in spite of the failure to achieve good wetting the conclusions reached above stimulated new thought on the techniques of alloying with the final result as described in Chapter 3.

Appendix F - Junction Area Correction

The following diode configuration is assumed:



The pellet is assumed to be hemispherical with a radius r and the regrowth region is assumed to be bounded by sides making an angle of 62° with the surface. This is an average angle for the {111} and {100} planes which form the boundaries. Since area corrections are generally less than 10% both of these assumptions are acceptable for first order corrections.

If the measured surface area of the alloyed region is designated A_1, the bottom area is designated A_2 and the area of the sides A_3; then by considering the pyramid whose base is A_1 it can be shown that the total junction area

$$A = A_2 + A_3 = 3.14\ A_1 - 0.636\ (V_{InSb})^{2/3} - 4.42\ (0.337\ A_1 - V_{InSb})^{2/3}$$

where $V_{InSb} = \frac{5.46 r^3 x}{1-2x}$, and x is the atom fraction of Sb corresponding to the In-Sb liquidus at the maximum temperature of the alloying cycle.

References

(1) J.Bardeen and W.H.Brattain, Phys.Rev.74, 230 (1948)

(2) H.Welker, Z. Naturforsch 11, 744 (1952)

(3) K.F.Hulme and J.B.Mullin, Solid State Electron. 5, 211 (1962)

(4) S.C.Choo, Ph.D. Thesis, University of London (1963)

(5) J.Hollis, Ph.D. Thesis, University of London (To be published)

(6) T.S.Liu and E.A.Peretti, Trans. Amer. Soc. Metals 44, 539 (1951)

(7) W.F.Schottky and M.B.Bever, Acta Met. 6, 320 (1958)

(8) K.F.Hulme, J. Electron. Control 6, 397 (1959)

(9) D.F.Gibbons, Phys. Rev. 112, 136 (1958)

(10) L.Bernstein and R.J.Beals, J.Appl. Phys. 32, 122 (1961)

(11) M.E.Fine, J.Appl. Phys. 24, 338 (1953)

(12) A.D.Stuckes, Phys. Rev. 107, 427 (1957)

(13) B.K.Agrawal and G.S.Verna, Physica 28, 599 (1962)

(14) R.Nii, Rev. Electrical Communication Labs. 8, 99 (1960)

(15) R.A.Smith, Semiconductors, Cambridge University Press (1959)

(16) A.S.Borschchevsku et al, Zh. Tekh. Fiz. 2, 1408 (1957)

(17) J.W.Allen, Phil. Mag. 4, 1046 (1959)

(18) C.Hilsum and A.C.Rose-Innes, Semiconducting III-V Compounds, Pergamon Press (1961)

(19) R.F.Broom, Proc. Phys. Soc.(Lond.) 71, 470 (1958)

(20) V.M.Goldschmidt, Trans. Faraday Soc. 25, 253 (1929)

(21) E.P.Warekois and P.H.Metzger, J.Appl. Phys. 30, 960 (1959)

(22) C.W.Mueller and N.H.Ditrick, Transistors 1, p.121, R.C.A. Laboratories (1956)

(23) J.B.Mullin, Compound Semiconductors, Vol.1, P.365, Reinhold Publishing Corp. (1962)

(24) J.F.Dewald, J.Electrochem. Soc. 104, 244 (1957)

(25) J.D.Venables and R.M.Broudy, J.Electrochem. Soc. 107, 296(1960)

(26) M.C.Lavine, A.J.Rosenberg and H.C.Gatos, J.Appl. Phys. 29, 1131 (1958)

(27) H.C.Gatos and M.C.Lavine, J.Electrochem. Soc. 107, 427 (1960)

(28) H.C.Gatos and M.C.Lavine, J.Phys. Chem. Sol. 14, 169 (1960)

(29) H.C.Gatos and M.C.Lavine, J.Appl. Phys. 31, 743 (1960)

(30) J.W.Allen, Phil. Mag. 2, 1475 (1957)

(31) Texas Instruments unpublished report.

(32) E.P.Warekois, M.C.Lavine and H.C.Gatos, J.Appl. Phys. 31, 1302 (1960)

(33) H.C.Gatos, M.C.Lavine and E.P.Warekois, J.Electrochem. Soc. 108, 645 (1961)

(34) E.N.Pugh and L.E.Samuels, J.Appl. Phys. 35, 1966 (1964)

(35) R.E.Hanneman, M.C.Finn and H.C.Gatos, J.Phys. Chem. Sol. 23, 1553 (1962)

(36) M.C.Finn and H.C.Gatos, Surface Science, 1, 361 (1964)

(37) D.Haneman, Brit, J.Appl. Phys. 16, 411 (1965)

(38) D.Haneman, G.J.Russell and H.K.Ip, Proc. Int. Conf. on the Physics of Semiconductors, p.1141 (1964)

(39) D.P.Miller, J.G.Harper and T.R.Perry, J.Electrochem. Soc. 108, 1123 (1961)

(40) F.H.Eisen, Phys. Rev. 135, A1394 (1964)

(41) D.Haneman, J.Phys. Chem. Sol. 14, 162 (1960)

(42) J.D.Venables and R.M.Broudy, J.Appl. Phys. 29, 1025 (1958)

(43) E.Peissker, P.Haasen and H.Alexander, Phil. Mag. 7, 1279 (1962)

(44) A.Willoughby and R.L.Bell, private communication.

(45) P.L.Moody, H.C.Gatos and M.C.Lavine, J.Appl. Phys. 31, 1696(1960)

(46) A.Steinemann and U.Zimmerli, Solid State Electron. 6, 597 (1963)

(47) H.C.Gatos, P.L.Moody and M.C.Lavine, J.Appl. Phys. 31, 212 (1960)

(48) R.K.Mueller and R.L.Jacobson, J.Appl.Phys. 32, 550 (1961)

(49) J.L.Richards, J.Appl. Phys. 31, 600 (1960)

(50) N.Nicholson and J.W.Faust, J.Electrochem. Soc. 110, 940 (1963)

(51) G.R.Booker, J.Appl. Phys. 33, 750 (1962)

(52) S.G.Ellis, J.Appl. Phys. 30, 947 (1959)

(53) R.E.Ewing and P.E.Greene, J.Electrochm. Soc. 111, 1266 (1964)

(54) F.V.Williams, J.Electrochem. Soc. 111, 886 (1964)

(55) M.T.Minamoto, J.Appl. Phys. 33, 1826 (1962)

(56) M.T.Minamoto and H.T.Malafi, J.Appl. Phys. 34, 1876 (1963)

(57) J.I.Pankove, J.Appl. Phys. 28, 1054 (1957)

(58) A.G.Chynoweth and G.L.Pearson, J.Appl. Phys. 29, 1103 (1958)

(59) V.Roberts and J.E.Quarrington, J.Electron. 1, 152 (1955)

(60) E.H.Putley, Proc. Phys. Soc. 73, 280 (1959)

(61) W.S.Boyle and A.D.Brailsford, Phys. Rev. 107, 903 (1957)

(62) S.Zwerdling, B.Lax and L.M.Roth, Phys.Rev. 108, 1402 (1957)

(63) S.D.Smith, T.S.Moss and K.W.Taylor, J.Phys. Chem. Sol. 11, 131 (1959)

(64) T.S.Moss, Progress in Semiconductors, 5, 191 (1960)

(65) H.L.Henneke, Solid State Electron. 3, 159 (1959)

(66) D.M.Bagguley, M.L.Robinson and R.A.Stradling, Phys. Rev. Letters 6, 143 (1963)

(67) J.Kolodziejczak, Proc.Int.Conf. on the Physics of Semiconductors p.1147 Paris (1964)

(68) R.A.Laff and H.Y.Fan, Phys. Rev. 121, 53 (1961)

(69) O.Madelung, Physics of III-V Compounds, John Wiley & Sons Inc. (1964)

(70) M.Glicksman and R.A.Powlus, Phys.Rev.121, 1659 (1961)

(71) A.G.Chynoweth and A.A.Murray, Phys. Rev. 123, 515 (1961)

(72) B.Ancker-Johnson, R.W.Cohen and M.Glicksman, Phys.Rev. 124, 1745 (1961)

(73) B.Ancker-Johnson, Phys. Rev. 134, A1465 (1964)

(74) I.Melngailis and R.H.Rediker, Proc. IRE, 50, 2428 (1962)

(75) M.C.Steele and M.Glicksman, J.Phys. Chem. Sol. 8, 242 (1959)

(76) M.C.Steele and M.Glicksman, Phys. Rev. 118, 474 (1960)

(77) J.B.Gunn, J.Electron. 2, 87 (1956)

(78) S.L.Miller, Phys. Rev. 99, 1234 (1955)

(79) C.A.Lee and G.Kaminsky, J.Appl. Phys. 31, 1717 (1960)

(80) F.M.Smits, Proc. IRE, 46, 1049 (1958)

(81) C.S.Fuller, Phys. Rev. 86, 136 (1952)

(82) E.Schillmann, Z. Naturforsch 11a, 472 (1956)

(83) F.H.Eisen and C.E.Birchenall, Acta Met. 5, 265 (1957)

(84) B.I.Boltaks and G.S.Kulikov, Zh. Tekh. Fiz. 27, 82 (1957)

(85) B.I.Boltaks and Yu A. Gutorov, Fiz. Tv. Tela. 1, 761 (1959)

(86) K.F.Hulme and J.E.Kemp, J.Phys. Chem.Sol. 10, 335 (1959)

(87) B.Goldstein, Properties of Elemental and Compound Semiconductors, p.155, Interscience Pub. N.Y. (1960)

(88) F.A.Cunnell and C.H.Gooch, J.Phys. Chem.Sol. 15, 127 (1960)

(89) B.Goldstein, Phys. Rev. 118, 1024 (1960)

(90) R.B.Wilson and E.L.Heasell, Proc.Phys.Soc. 79, 403 (1962)

(91) F.A.Cunnell and C.H.Gooch, Nature 188, 1096 (1960)

(92) R.L.Longini, Solid State Electron. 5, 127 (1962)

(93) D.Shaw, P.Jones and D.Hazelby, Proc.Phys.Soc. 80, 167 (1962)

(94) S.M.Sze and L.Y.Wei, Phys. Rev. 124, 84 (1961)

(95) R.W.Orth and L.A.K.Watt, J.Phys. Chem.Sol. 26, 197 (1965)

(96) B.I.Boltaks and V.I.Skolov, Fiz. Tv. Tela. 5, 1077 (1963)

(97) B.I.Boltaks and V.I.Skolov, Fiz. Tv. Tela. 6, 771 (1964)

(98) L.A.Watt and W.S.Chen, Bull. Am. Phys.Soc. 7, 89 (1962)

(99) M.T.Minamoto and C.M.Allen, Solid State Electron. 5,263 (1962)

(100) J.R.Knight, D.Effer and P.R.Evans, Solid State Electron. 8, 178 (1965)

(101) J.B.Mullin, C.V.D. Conference on Electronics Materials Res. Birmingham (1964)

(102) D.A.Jenny, Proc. I.R.E. 46, 959 (1958)

(103) G.R.Mitchell, A.E.Goldberg and S.W.Kurnick, Phys. Rev. 97, 239 (1955)

(104) R.M.Talley and D.P.Enright, Phys. Rev. 95, 1092 (1954)

(105) R.Gremmelmaier, Z. Naturforsch. 10a, 501 (1955)

(106) R.Gremmelmaier and H.Welker, Z. Naturforsch. 11a, 420 (1956)

(107) P.Rappaport, R.C.A.Review, 20, 373 (1959)

(108) J.Lowen and R.H.Rediker, J.Electrochem. Soc. 107, 26 (1960)

(109) J.Halpern and R.H.Rediker, Proc.I.R.E. 48, 1780 (1960)

(110) R.H.Rediker and T.M.Quist, Solid State Electron. 6, 657 (1963)

(111) H.Welker and R.Gremmelmaier, Proc. I.E.E. 106B, 850 (1959)

(112) A.Herzog, R.R.Haberecht, and A.E.Middleton, J.Electrochem.Soc. 105, 533 (1958)

(113) J.Mandelkorn, Proc. I.R.E. 47, 2012 (1959)

(114) W.Sharpless, Bell Syst. Tech. J. 38, 259 (1959)

(115) R.N.Hall, J.H.Racette and H.Ehrenriech, Phys. Rev. Lett. 4, 456 (1960)

(116) R.N.Hall, Proc.Int.Conf. on Semiconductors, p.193, Prague(1961)

(117) R.N.Hall and J.H.Racette, J.Appl. Phys. 32, 2078 (1961)

(118) R.L.Anderson, Solid State Electron. 5, 241 (1962)

(119) R.L.Anderson, I.B.M. J.Res. 4, 283 (1960)

(120) F.A.Pizzanello, J.Electrochem. Soc. 109, 226 (1962)

(121) J.R.Dale and M.J.Josh, Solid State Electron. 8, 1 (1965)

(122) M.I.Nathan et al, Appl. Phys. Lett. 1, 62 (1962)

(123) R.N.Hall, et al, Phys. Rev. Lett. 9, 366 (1962)

(124) T.M.Quist, et al, Appl. Phys. Lett. 1, 91 (1962)

(125) R.J.Rhelan, et al, Appl. Phys. Lett. 3, 143 (1963)

(126) K.Weiser and R.S.Levitt, Appl. Phys. Lett. 2, 178 (1963)

(127) I.Melngailis, Appl. Phys. Lett. 2, 176 (1963)

(128) I.Melngailis, A.J.Strauss and R.H.Rediker, Proc. IEEE. 51, 1154 (1963)

(129) F.B.Alexander, et al, Appl. Phys. Lett. 4, 13 (1964)

(130) N.Holonyak and S.F.Bevacqua, Appl. Phys.Lett. 1, 82 (1962)

(131) G.Burns and M.I.Nathan, Proc. IEEE. 52, 770 (1964)

(132) W.N.Reynolds and M.T.Lilburne, Proc.Phys.Soc. 71, 416 (1958)

(133) D.Filatley and D.Stolnitz, Solid State Electron. 8, 255 (1965)

(134) M.E.Jones and E.C.Wurst, I.R.E. Conv. Rec. 9, pt.3, 26 (1961)

(135) D.G.Avery, D.W.Goodwin and A.E.Rennie, J.Sci.Inst. 34, 394(1957)

(136) V.V.Galavanov and N.A.Erokhina, Sov.Phys.Solid State, 1, 1096 (1960)

(137) M.E.Lasser, P.Cholet and E.C.Wurst, J.Opt. Soc.Am. 48, 468 (1958)

(138) C.A.Lee and G.Kaminsky, J.Appl. Phys. 30, 2021 (1959)

(139) C.A.Lee and G.Kaminsky, J.Appl. Phys. 31, 1717 (1960)

(140) Y.Marfaing, Compt. Rend. 250, 3608 (1960)

(141) A.G.Chynoweth and R.A.Logan, Phys.Rev. 118, 1470 (1960)

(142) R.L.Batdorf, G.C.Dacey, R.L.Wallace and D.J.Walsh, J.Appl.Phys. 31, 613 (1960)

(143) A.R.Calawa, R.H.Rediker, B.Lax and A.L.Mc Whorter, Phys.Rev. Lett. 5, 55 (1960)

(144) A.G.Chynoweth, R.A.Logan and P.A.Wolff, Phys.Rev. Lett. 5, 548 (1960)

(145) A.G.Chynoweth, G.H.Wannier, R.A.Logan and D.E.Thomas, Phys. Rev.Lett. 5, 57 (1960)

(146) P.N.Butcher, J.A.Hulbert and K.F.Hulme, J.Phys. Chem. Sol. 21, 320 (1961)

(147) L.Esaki and R.R.Haering, J.Appl. Phys. 33, 2466 (1962)

(148) K.F.Hulme, Brit. J. Appl. Phys. 12, 651 (1961)

(149) Y.Marfaing, Comp. Rend. 253, 626 (1961)

(150) Y.Marfaing, Solid State Electron. 7, 1 (1964)

(151) H.C.Gorton, A.R.Zacaroli, F.J.Reid and C.S.Peet, J.Electrochem. Soc. 108, 354 (1961)

(152) V.V.Galavanov, Radiotechnica i Electronica 9, 1417 (1964)

(153) I.Melngailis and R.H.Rediker, J.Appl. Phys. 33, 1892 (1962)

(154) I.Melngailis, R.J.Phelan and R.H.Rediker, Appl.Phys.Lett. 5, 99 (1964)

(155) C.M.Allen, P.R.Liegey and B.Salzberg, Proc. IEEE. 51, 856(1963)

(156) R.K.Mueller and R.L.Jacobson, J.Appl.Phys. 35, 1524 (1964)

(157) K.N.Maffitt and R.K.Mueller, J.Appl. Phys. 35, 1563 (1964)

(158) C.T.Sah, R.N.Noyce and W.Shockley, Proc.I.R.E. 45, 1228 (1957)

(159) H.J.Stocker, J.Appl. Phys. 32, 322 (1961)

(160) H.D.Barber and E.L.Heasell, Solid State Electron. 8, 113 (1965)

(161) B.R.Pagel and R.L.Petritz, J.Appl.Phys. 32, 1901 (1961)

(162) G.R.Pruett and R.L.Petritz, Proc.I.R.E. 47, 1524 (1959)

(163) L.B.Valdes, The Physical Theory of Transistors, McGraw-Hill (1961)

(164) A.K.Jonscher, Principles of Semiconductor Device Operation, G. Bell and Sons (Lond.) (1960)

(165) A.Nussbaum, Semiconductor Device Physics, Prentice Hall (1962)

(166) R.N.Hall, Phys. Rev. 87, 387 (1952)

(167) W.Shockley and W.T.Read, Phys. Rev. 87, 835 (1952)

(168) W.Shockley, Bell System Tech. J. 28, 435 (1949)

(169) J.J.Sparks, J.Electron. Control, 16, 153 (1964)

(170) P.E.Gray, D.Dewitt, A.R.Boothroyd and J.F.Gibbons, SEEC Notes 2, PEM, John Wiley and Sons (1962)

(171) J.L.Moll, Physics of Semiconductors, McGraw-Hill (1964)

(172) P.H.Passau and M. van Styvendael, Solid State Phys. in Electron and Telecommunications, 1, 407 Academic Press (1960)

(173) R.W.Lade, Proc. IEEE, 52, 743 (1964)

(174) W.B.Berry, Proc. IEEE, 53, 188 (1965)

(175) J.S.Blakemore, Semiconductor Statistics, Pergamon Press (1962)

(176) R.N.Hall, Proc. IRE, 40, 1512 (1952)

(177) N.H.Fletcher, Proc. IRE, 45, 862 (1957)

(178) W. van Roosbroeck, Phys. Rev. 91, 1079 (1953)

(179) N.H.Fletcher, J.Electronics, 2, 609 (1957)

(180) R.W.Lade and C.G.Poncelet, Proc. IEEE, 52, 629 (1964)

(181) K.G.McKay and K.B.McAfee, Phys. Rev. 91, 1079 (1953)

(182) S.L.Miller, Phys. Rev. 105, 1248 (1957)

(183) S.L.Miller and J.J.Ebers, Bell System Tech. J. 34, 883 (1955)

(184) C.Zener, Proc. Roy. Soc.(Lond.) 145A, 523 (1934)

(185) K.B.McAfee, E.J.Ryder, W.Shockley and M.Sparks, Phys.Rev. 83, 650 (1951)

(186) D.Dewitt and A.L.Rossoff, Transistor Electronics, McGraw-Hill (1957)

(187) B.Lax and S.F.Neustadter, J.Appl. Phys. 25, 1148 (1954)

(188) J.S.Saby and W.C.Dunlap, Phys. Rev. 90, 630 (1953)

(189) R.C.Ingraham and R.E.Hunt, IRE Wescon. Conv. Record, 3, 50(1959)

(190) K.Lehovec, et al, J.Electrochem. Soc. 108, 241 (1961)

(191) K.Lehovec, US Pat. No. 2,893,901.

(192) G.C.A.Schuit and N.H.de Boer, Rec.Trav.Chim. 70, 1067 (1951) ibid 72, 909 (1952)

(193) K.J.Laider, J.Phys. Colloid. Chem. 55, 1067 (1951)

(194) A.K.Jonscher, J.Electron. Control, 5, 226 (1958)

(195) A.R.Bennet and R.L.Longini, Phys. Rev. 116, 53 (1959)

(196) M.F.Millea and C.T.Tomizuka, J.Appl.Phys. 27, 96 (1956)

(197) W.D.Edwards, Can. J. Phys. 38, 439 (1960)

(198) J.H. Wernick, J.Metals, 205, 1169 (1957)

(199) J.B.Mullin and D.T.J.Hurle, private communication.

(200) R.C.Sangster, Compound Semiconductors, Vol.1, p.241, Reinhold Pub. Corp. (1962)

(201) L.Pauling, The nature of the Chemical Bond, Cornell University Press N.Y. (1948)

(202) D.B.Holt, J.Appl. Phys. 31, 2231 (1960)

(203) H.C.Gatos, J.Appl. Phys. 32, 1232 (1961)

(204) D.Haneman, Compound Semiconductors Vol.1, p.423, Reinhold Pub. Corp. (1962)

(205) J.W.Cahn and R.E.Hanneman, Surface Science, 1, 387 (1964)

(206) W.A.Tiller, J.Appl. Phys. 29, 611 (1958)

(207) W.K.Burton, N.Cabrera and F.C.Frank, Phil. Trans. Roy. Soc. 243A, 299 (1951)

(208) D.T.J.Hurle, Progress in Materials Science, Vol.10, Pergamon Press (1962)

(209) A.J.Strauss, Solid State Electron. 5, 97 (1962)

(210) O. Lindberg and J.W.Faust Jr. Compound Semiconductors, Vol.1, p. 294, Reinhold Pub. Corp. (1962)

(211) W.Bardsley, J.S.Boulton and D.T.Hurle, Solid State Electron. 5, 395 (1962)

(212) D.T.J.Hurle, O.Jones and J.B.Mullin, Solid State Electron. 3, 317 (1961)

(213) G.Joos, Theoretical Physics, p.574, Blackie & Son Ltd.(1958)

(214) L.V.Azaroff, Introduction to Solids, McGraw-Hill (1960)

(215) W.P.Allred and R.T.Bate, J.Electrochem. Soc. 108, 258 (1961)

(216) E.Billig, Semiconductors and Phosphors, p.2, Friedr.Vieweg und Sohn, Braunschweig (1958)

(217) W.C.Dash, J.Appl. Phys. 30, 459 (1959)

(218) E.Billig, Proc.Roy. Soc. (London) A235, 37 (1956)

(219) H.G.van Bueren, J.Hornstra and P.Penning, Nuovo Cimento, 7, 646 (1958)

(220) A.Ghani, D.I.C. Thesis, Imperial College, London, (1965)

(221) M.J.Hampshire, Elect.Eng.Dept. Memo 190, University of Birmingham (1964)

(222) K.T.Rogers, A.J.Dawes and R.L.Bell, J.Appl.Phys. 36, 1493(1965)

(223) H.D.Barber and E.L.Heasell, J.Appl.Phys. 36, 176 (1965)

(224) F.W.G.Rose, Semiconductors and Phosphors, Verlag Friedr. Vieweg & Sohn, Braunschweig (1958)

(225) H.Yoshinaga and R.A.Oetjen, Phys. Rev. 101, 526 (1956)

(226) J.L.Moll and R.van Overstraeten, Solid State Electron. 6, 147 (1963)

(227) A.C.Prior, J.Electron. Control, 4, 165 (1958)

(228) J.Yamashita, Progress in Semiconductors, 4, 63 Heywood (1960)

(229) K.G.McKay, Phys. Rev. 94, 877 (1954)

(230) R.W.Cunningham, Boeing Research Labs. Report D1-82-0306 (1963)

(231) J.J.Sparkes and J.R.W.Smith, J.Electron. Control, 12, 177(1962)

(232) A.G.Chynoweth and K.G.McKay, Phys. Rev. 102, 369 (1956)

(233) D.J.Rose, Phys. Rev. 105, 413 (1957)

(234) A.C.English, Solid State Electron. 6, 511 (1963)

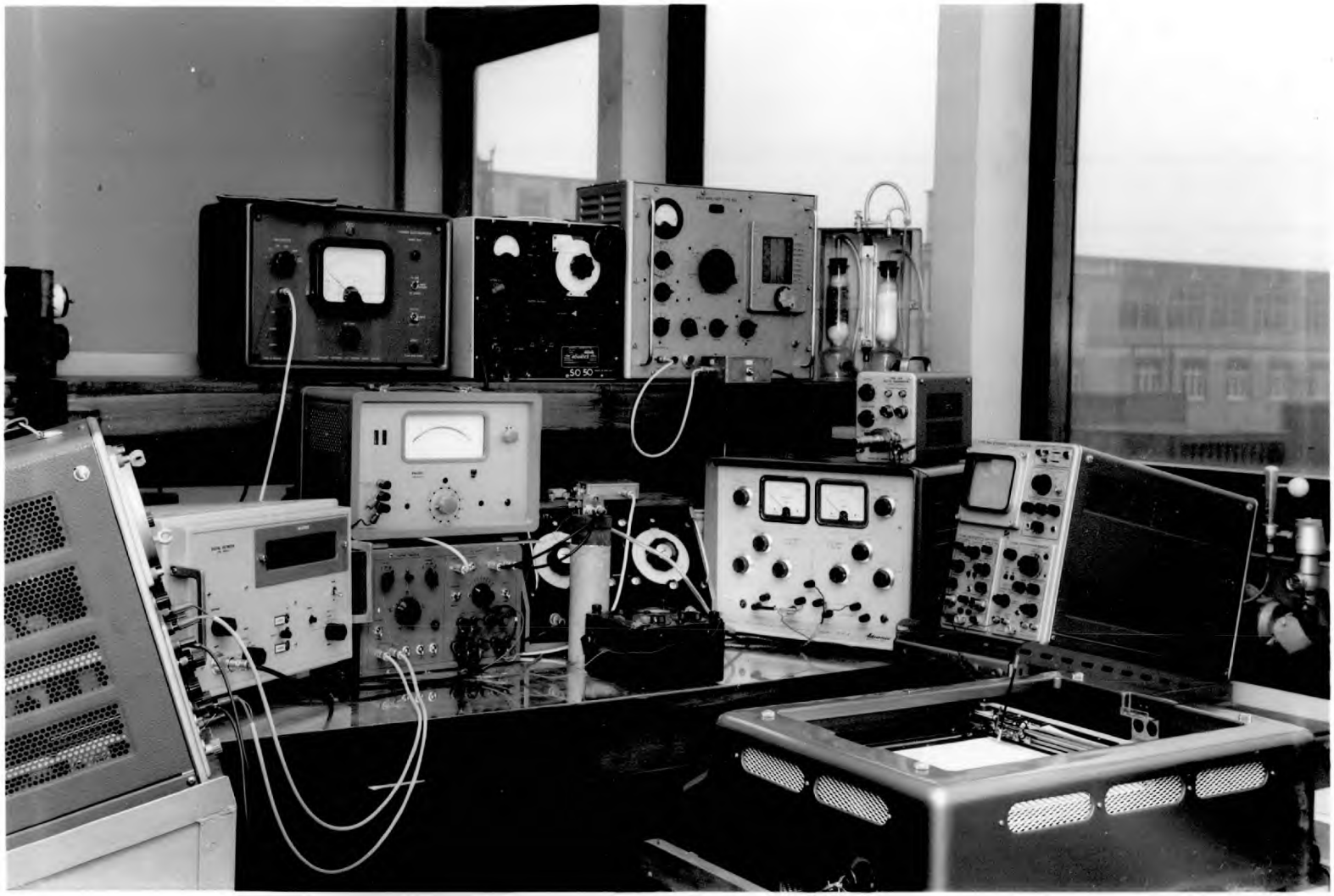
(235) G.B.Larrabee, J.Electrochem. Soc. 108, 1130 (1961)

(236) J.L.Davis, Bull. Amer. Phys. Soc. 6, 18A (1961)

(237) J.F.Pudvin, Proc. IEE. 106B, 1125 (1959)

(238) J.T.Edmond, Proc. Phys. Soc. 76, 802 (1960)

(240) C.Kolm, S.A.Kulin and B.L.Averbach, Phys.Rev. 108, 965 (1957)



Electrical Measuring System

Solid-State Electronics Pergamon Press 1965. Vol. 8, pp. 113–117. Printed in Great Britain

A TECHNIQUE FOR MAKING ALLOY *p–n* JUNCTIONS IN InSb

H. D. BARBER* and E. L. HEASELL*

Electrical Engineering Dept., Imperial College, London S.W.7

(*Received* 12 *June* 1964; *in revised form* 7 *July* 1964)

Abstract—A radiation wetting technique coupled with a quasi equilibrium alloying procedure has been developed which gives uniform wetting and good junctions on indium antimonide at alloying temperatures as low as 200°C. The technique allows controlled formation of alloyed junctions of very low penetration such as would be required in the formation of alloy-diffused transistors in indium antimonide. The technique also permits the formation of flat junctions on both {111} and {$\bar{1}\bar{1}\bar{1}$} surfaces. Possible explanations for the success of the technique are discussed.

Résumé—Une technique de mouillage à rayonnement associée à un procédé d'alliage de quasi-équilibre ont été développés pour donner un mouillage uniforme et de bonnes jonctions sur l'antimoniure d'indium à des températures d'alliages aussi basses que 200°C. Cette technique permet la formation controllée des jonctions alliées de faibles pénétrations telles que celles requises dans la formation des transistors alliés-diffusés dans l'antimoniure d'indium. Cette technique permet aussi la formation de jonctions plates dans les surfaces (111) et ($\bar{1}\bar{1}\bar{1}$). Des raisons possibles pour le succés de cette méthode sont discutées.

Zusammenfassung—Ein auf Strahlung beruhendes Anfeuchtungsverfahren kombiniert mit einem Legierungsprozess im Quasi-Gleichgewicht wurde entwickelt. Man erhält gleichförmige Anfeuchtung und gute Übergänge auf Indium-Antimonid bei Legierungstemperaturen von nur 200°C. Das Verfahren ermöglicht die Steuerung der Bildung von legierten Übergängen bei sehr niedrigen Eindringtiefen, wie sie bei der Herstellung von legierten Transistoren durch Legierungsdiffusion erforderlich ist. Das Verfahren lässt ausserdem die Bildung flacher Übergänge auf {111}- und {$\bar{1}\bar{1}\bar{1}$}-Flächen zu. Schliesslich wird versucht, eine Erklärung für den Erfolg des Verfahrens zu finden.

INTRODUCTION

ATTEMPTS to use established techniques[1–3] in alloying indium to indium antimonide have disclosed a number of difficulties. It has proved necessary to develop a modification of the usual alloy techniques in order to circumvent these.

Experiments in this laboratory using a capilliary alloying system[3,4] have shown that wetting of indium antimonide surfaces by indium is strongly inhibited by very small amounts of oxide. It has also been found that indium above its melting point oxidises rapidly and that this oxide is particularly effective in preventing wetting. Even when a fresh indium surface is exposed from the capilliary in a hydrogen atmosphere, where oxygen and water vapor content are kept below the level of detection, the oxygen held in solution in indium is sufficient to prevent uniform wetting at temperatures below 300°C.

Indium antimonide melts at 525°C[5] and thus, unless very large penetrations can be tolerated, alloying temperatures much lower than that at which hydrogen becomes an effective reducing agent are desirable. This is particularly true if shallow junctions are required as in making contacts to diffused layers.

In order to obtain wetting using conventional strip heater techniques, either the temperature has to be raised rapidly to values above 300°C in order to reduce the effects of oxidation or a flux is required. In the absence of a flux the wetting area obtained at 300°C is small. Thus, although the wetted area increases with increasing temperature, large penetrations are unavoidable. Fluxes may promote spreading and not yield uniform wetting. Moreover, they are undesirable because of the danger of contamination and inclusion in the junction region.

An additional problem is posed by the polar nature of the zinc-blende lattice in ⟨111⟩ directions. It has been shown[6] that using conventional alloying techniques it may not be possible to obtain p–n junctions having good metallurgical or electrical properties on both {111} and {$\bar{1}\bar{1}\bar{1}$} surfaces. (The {111} surfaces are taken here as those ideally terminating in triply bonded antimony atoms and {$\bar{1}\bar{1}\bar{1}$} surfaces as those ideally terminating in indium atoms.)

In view of these difficulties the following technique for obtaining uniform wetting, controlled alloying at temperatures as low as 200°C and consequent good p–n junctions or either {111} or {$\bar{1}\bar{1}\bar{1}$} surfaces has been developed.

TECHNIQUE

In this technique for making alloy junctions in indium antimonide wetting and alloying are effected in two separate processes. Uniform wetting with negligible penetration of the crystal is obtained by radiation heating for a very short period of time in the system illustrated in Fig. 1. The crystal with a pellet of indium on top is placed on a heat sink immediately under a tungsten filament.

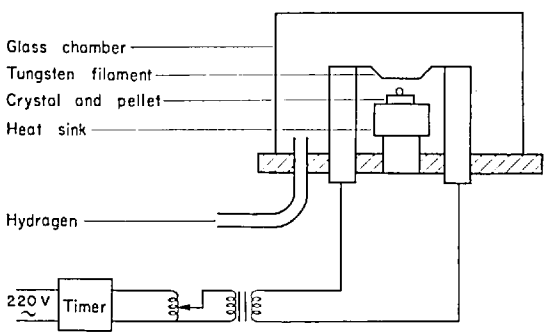


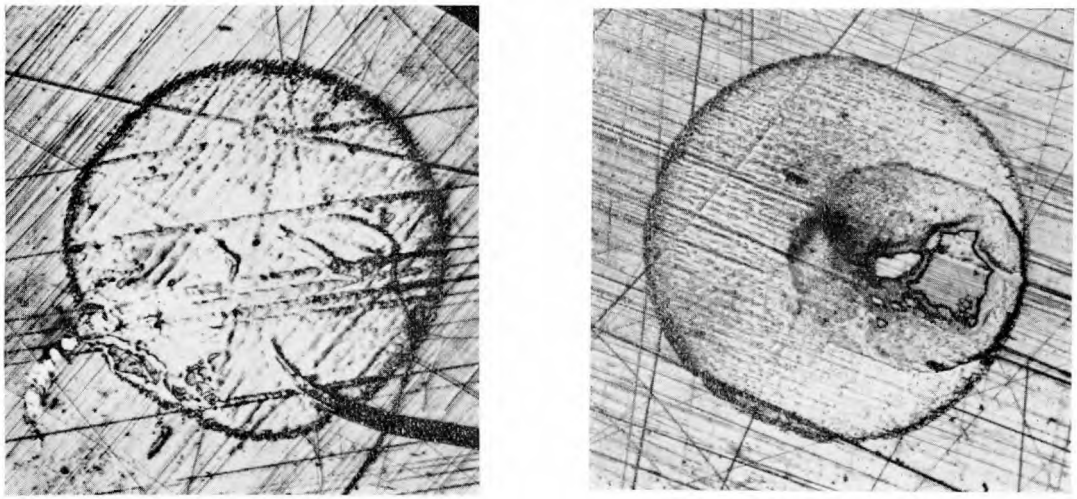
FIG. 1. Radiation wetting system.

The filament is heated to white heat in an atmosphere of hydrogen or forming gas for a period ranging from one to four seconds. The heating current is accurately controlled and the heating cycle is controlled to better than 1 per cent by means of an automatic timer. The pellet appears to melt instantly under the radiation and then collapses onto the crystal. The heating cycle is so adjusted that the indium freezes onto the crystal in a hemispherical shape. With wetting thus accomplished the alloying procedure is greatly simplified, since the extreme precautions required to keep the alloying surfaces and the atmosphere clean are no longer necessary. The crystal and wetted pellet are placed in a furnace in air or hydrogen and the temperature raised at 5°C per minute up to the required alloying temperature. Usually the furnace is maintained at this temperature for 5–10 min, although good diodes can also be made if cooling starts the instant the maximum temperature is reached. During the cooling the temperature drop is again limited to 5°C per minute to promote good re-growth of the crystal.

TYPICAL RESULTS

The junctions chosen to illustrate the results achieved by this technique were made on single crystal indium antimonide having a dislocation density less than 1000/cm² as revealed by etch pit counts. The crystals were cut on {111} surfaces to within $\pm 1°$ and polished on a one micron diamond lap. Since the surface scratches on a polished crystal make it easier to compare the wetted or alloyed region with the original surface, several of the examples have been chosen from dice not even lightly etched before the wetting procedure. The pellets used were spherical, less than 500 μ dia. and composed of 99% indium, 1% cadmium. Pellets of indium, 7% cadmium and pure tin have also been used successfully.

Radiation wetting generally produces uniform wetting over the entire area under the dot. Figure 2(a) shows such a wetted region after removal of the pellet in dilute HCl. However, it is not uncommon to find a spot or cluster of spots as shown in Fig. 2(b) which have not wet during radiation. These spots have been related to the point of contact between the pellet and the crystal before wetting. When these spots occur they are, in general, small and thus are quickly and easily dissolved during the alloying cycle. Measurements of the pellet penetration after radiation wetting have shown that it is limited to a few microns at most. This is demonstrated in Fig. 2(a) and (b) by the fact that some of the small scratches (1 μ) and all of the larger scratches ($\sim 3\,\mu$), can be traced through the wetted area. It is important to note, however, that except for the spots already referred to in connection with Fig. 2(b), none of the original



(a) (b)

FIG. 2. Wetted areas after removal of the indium pellets in dilute HCl (a) uniformly wetted (× 150) (b) small unwetted areas remaining. (× 100)

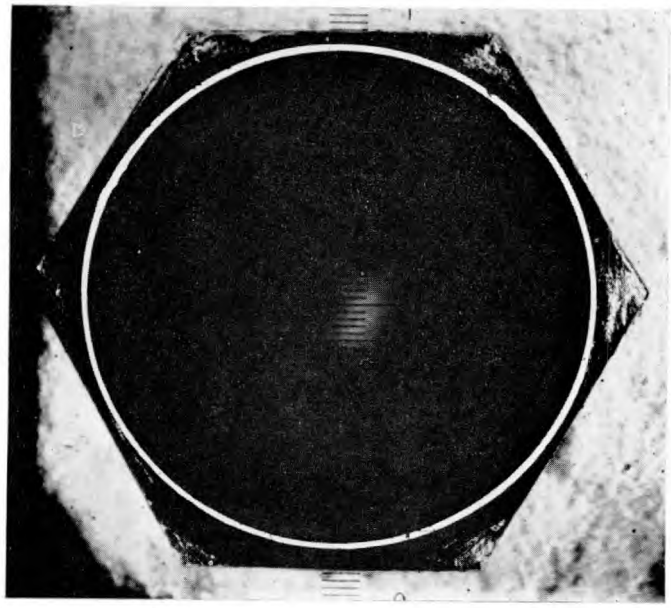


FIG. 3. Hexagonal wetted area bounded by {111} planes after alloying. Original circular wetted area indicated in white. (× 100)

[*facing p.* 114

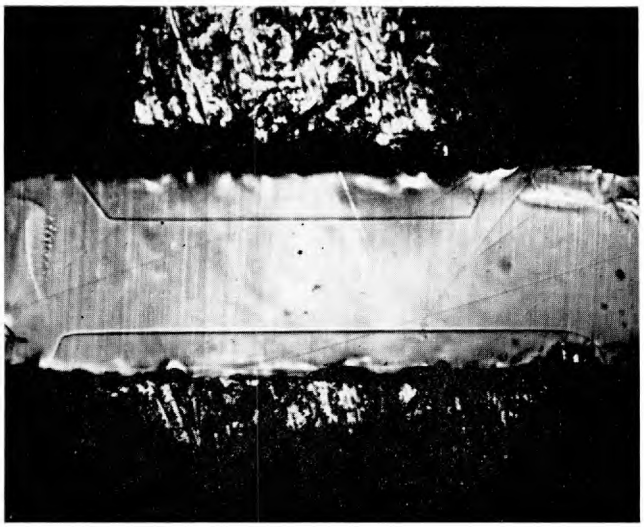


FIG. 4. Section of an InSb double junction showing flat junctions on both $\{\bar{1}\bar{1}\bar{1}\}$ (top) and $\{111\}$ surfaces. ($\times 100$)

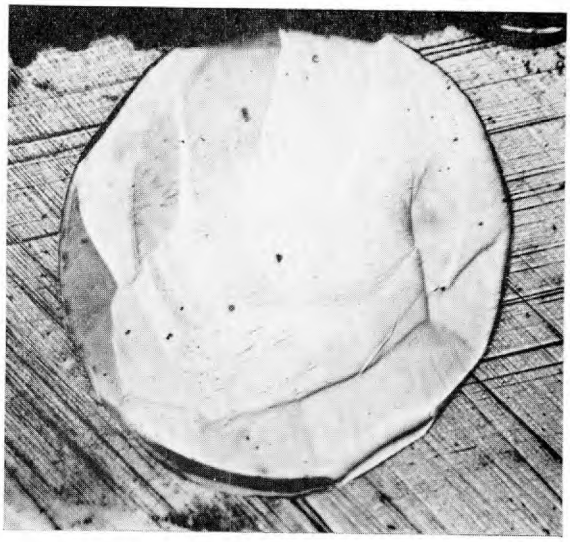


FIG. 5. Surface regrowth after alloying at 250°C and removing the pellet in dilute HCl. ($\times 110$).

crystal surface is visible and this indicates that uniform wetting has occurred.

When the alloying cycle is performed in air, a tenacious oxide forms on the surface of the indium pellet and prevents it from spreading or changing shape. The actual wetted area does change, however, during the alloying as can be seen in Fig. 3. The change of wetted area is due to horizontal dissolution. If care is taken to obtain the same wetting and alloying conditions the final wetted area is quite reproducible. Sections such as that in Fig. 4 have shown that during alloying the melt proceeds into the crystal bounded by planar {111} surfaces on either {111} or $\{\bar{1}\bar{1}\bar{1}\}$ surfaces. From Fig. 4 and the surface of a typical regrowth shown in Fig. 5, it can be seen that the regrowth is extensive and uniformly single crystal. Back reflection Laue photographs of the regrowth region have further confirmed its single crystal nature and shown it to be of the same orientation as the wafer.

As a check on the possible control over alloying penetration using this technique, thirty diodes were prepared using pellets weighed to the nearest microgram and alloyed at temperatures ranging from 250°C to 428°C in either air or hydrogen. Each diode was sectioned to determine the alloy penetration. The mass of dissolved crystal was then calculated using the depth of penetration thus obtained and the wetted area. Values of x, the atom fraction of indium in the liquid at the alloying temperature were then calculated from the following expression:

$$m_1 = \frac{(1-x)M_1}{(2x-1)M_2} m_2$$

where m_1 is the mass of InSb dissolved, m_2 is the mass of the indium pellet, M_1 is the molecular weight of InSb and M_2 is the atomic weight of indium. These values are shown in Fig. 6 plotted on the known phase diagram of InSb.[7,8] In the calculations no allowance was made for the fact that a considerable amount of the indium formed an oxide during air alloying and hence the difference between the curves for air alloying and hydrogen alloying. The points obtained by hydrogen alloying are in extremely good agreement with the phase diagram.

The most critical test of the metallurgical perfection of a p–n junction is its electrical characteristic. Junctions formed by radiation wetting without further alloying gave high reverse currents with a soft breakdown at voltages between 4–6 V. Junctions which were alloyed after radiation wetting gave good electrical characteristics

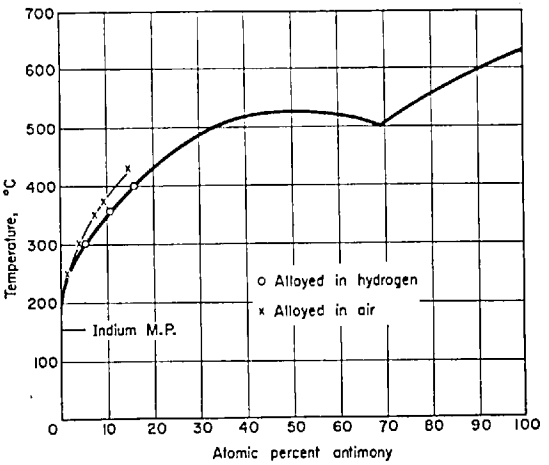


FIG. 6. Measured values of In–Sb liquidus as determined by alloy penetration during alloying in hydrogen or air compared with the known In–Sb phase diagram.[6,7]

even when alloying temperatures were as low as 200°C. A typical current–voltage characteristic for a junction alloyed at 250°C is shown in Fig. 7. Although reverse currents are about one order of magnitude higher than that expected from conventional diffusion and recombination theories,[9] experiments indicate that these are surface and not bulk currents. The forward characteristics can be fitted quite well using conventional diode theory and the coefficients obtained agree reasonably with expected values.

DISCUSSION

The radiation wetting process is the most important step in this technique for obtaining good alloyed p–n junctions. Wetting and spreading of a liquid on a surface is determined to a large extent by the surface energies of the liquid and the solid.[10] In the case of molten indium on indium-antimonide it appears that oxides on the indium cause a large increase in its surface energy and hence preclude any wetting unless the temperature is raised to a very high value or a flux is used.

During radiation wetting the pellet is closer to the heater than the crystal and thus it becomes very

hot. Initially it melts and surface tension pulls it up into a sphere leaving only a small area of contact on the relatively cold crystal. Because of this small contact area heat losses to the crystal are small and the pellet becomes very hot. At this elevated temperature it is possible that some of the surface oxide is taken into solution by the indium. However, more important is the fact that

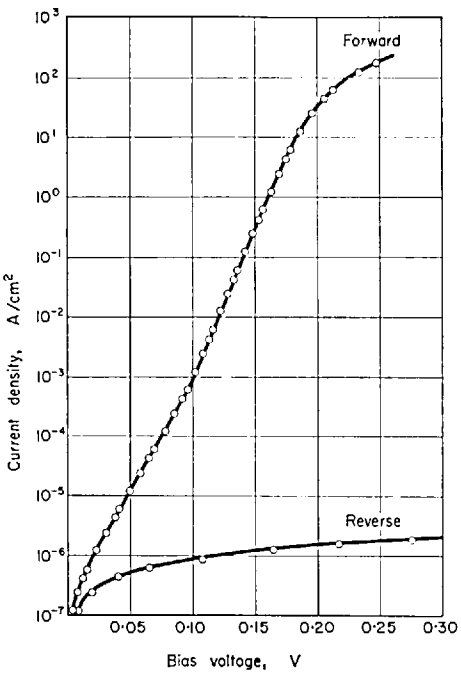


FIG. 7. Current–voltage characteristic of an InSb alloyed p–n junction at 77°K. Alloyed at 250°C on a $\{\bar{1}\bar{1}\bar{1}\}$ surface, wafer 1×10^{14} cm^{-3} n-type, alloyed pellet In–1% Cd, area 10^{-3} cm^2.

a hot tungsten filament in hydrogen generates considerable amounts of atomic hydrogen.[11] Under the conditions used in this system it is estimated that more than 0·05 per cent of the hydrogen flow is converted into atomic hydrogen, which probably acts as an effective reducing agent at temperatures as low as 200°C.

From these considerations it appears that during radiation two processes which promote wetting are proceeding simultaneously; the surface oxide is being removed and the temperature of the pellet is rising. Because these processes are proceeding quickly the last point on the pellet surface where the critical wetting conditions are reached is at the point of contact between the pellet and the crystal. This gives rise to a highly unstable situation and thus when wetting starts the pellet literally collapses onto the crystal. The thermal contact with the crystal is then very good and the pellet temperature drops to the crystal temperature. It is now clear why a small unwetted area of crystal may be left at the original point of contact, since not all of the contact area may have reached the critical wetting condition when the wetting is nucleated. Since good diodes can be made by alloying at temperatures as low as 200°C it is probable that the crystal temperature is always below 200°C during normal radiation wetting conditions. This also substantiates the conclusions drawn from Fig. 2(a) and (b) that penetration during the wetting procedure must indeed be small and that the wetting itself must be uniform.

The slow rate of change of temperature associated with the alloy cycle ensures that at all times the melt is essentially in equilibrium with the solid and hence flat junctions are obtained on both $\{111\}$ and $\{\bar{1}\bar{1}\bar{1}\}$ surfaces. If this precaution is not observed the different dissolution rates of $\{111\}$ and $\{\bar{1}\bar{1}\bar{1}\}$ surfaces will give rise to diffusion controlled round junctions on one or both surfaces.[12] These round junctions give similar characteristics to those obtained for diodes made by radiation wetting only. The soft reverse characteristic is probably due to local breakdowns in regions of high irregularity and to imperfection in the p–n junction. The electrical characteristics of flat junctions obtained by this technique are comparable with the best reported in the literature.[13] The departures of these characteristics from those theoretically predicted are probably due to surface effects. This indicates that junctions formed by this technique possess a good degree of perfection in the p–n interface and in the regrowth region.

Acknowledgement—The authors wish to express their thanks to Mr. J. CONRADI who undertook part of the original sectioning and metallographic preparation and to Dr. J. C. ANDERSON for his support during the work. One of us (H.D.B.) also wishes to acknowledge the financial support of the Athlone Fellowship Committee and the National Research Council of Canada.

REFERENCES

1. J. M. GOLDEY, *Transistor Technology*, Vol. 3, p. 231, Van Nostrand, N.Y. (1958).
2. N. P. BURCHAM, *Transistor Technology*, Vol. 3, p. 175, Van Nostrand, N.Y. (1958).
3. K. LEHOVEC *et al.*, *J. Electrochem. Soc.* **108**, 241 (1961).
4. R. C. INGRAHAM and R. E. HUNT, *Inst. Radio Engrs. Wescon Conv. Records*, Vol. 3, Pt. 3, 50 (1959).
5. K. F. HULME and J. B. MULLIN, *Solid-State Electron.* **5**, 211 (1962).
6. M. T. MINAMOTO, *J. Appl. Phys.* **33**, 1826 (1962).
7. T. S. LIU and E. A. PERETTI, *Trans. Amer. Soc. Metals* **44**, 534 (1952).
8. R. N. HALL, *J. Electrochem. Soc.* **110**, 385 (1963).
9. C. T. SAH, R. N. NOYCE and W. SHOCKLEY, *Proc. Inst. Radio Engrs.* **45**, 1228 (1957).
10. A. BIONDI, *Chem. Rev.* **52**, 417 (1953).
11. K. J. LAIDLER, *J. Phys. Colloid Chem.* **55**, 1067 (1951).
12. H. D. BARBER and E. L. HEASELL, to be published.
13. H. J. STOCKER, *J. Appl. Phys.* **32**, 322 (1961).

Reprinted from JOURNAL OF APPLIED PHYSICS, Vol. 36, No. 1, 176–180, January 1965

Polarity Effects in InSb Alloyed p–n Junctions

H. D. BARBER AND E. L. HEASELL
Electrical Engineering Department, Imperial College, London, S. W. 7, England
(Received 9 June 1964; in final form 18 August 1964)

The shape dependence of the profiles of alloy junctions formed on {111} and {$\bar{1}\bar{1}\bar{1}$} surfaces of InSb has been studied as a function of alloying conditions. Polarity-dependent effects are found which may be attributed to different dissolution rates for {111} and {$\bar{1}\bar{1}\bar{1}$} surfaces. Direct evidence of this difference is presented. By proper choice of alloying conditions it is possible to obtain planar junctions on both types of surface. The rectification ratios of such junctions prepared on either surface range from 10^7 to 2×10^8, at 250 mV. The polar differences in the current–voltage characteristics of these junctions are small and tentatively attributed to polarity-dependent surface effects.

INTRODUCTION

THE lack of inversion symmetry in the ⟨111⟩ direction of III–V compounds gives rise to a difference in a number of the properties of {111} or {$\bar{1}\bar{1}\bar{1}$} surfaces. After Dewald[1] the following convention with regard to A and B surfaces will be followed throughout this paper. The A or {$\bar{1}\bar{1}\bar{1}$} surface is defined as that ideally terminating in group III atoms triply bonded to the matrix and the B or {111} surface as that terminating in triply bonded group V atoms. Although the definition of {111} and {$\bar{1}\bar{1}\bar{1}$} surfaces is arbitrary and other authors[2,3] have used the opposite convention, the surfaces have been positively identified by x-ray experiments[4–6] and thus whatever convention is chosen it can be unambiguously applied.

It has been shown that A surfaces are more noble[3] and are hence less easily attacked chemically[3,7,8] or oxidized anodically.[1,9] Unless inhibitors are used etch pits develop only on A surfaces.[3,7] Crystal growth is faster on A surfaces,[10,11] and crystals are more perfect when grown from B ended seeds.[12–14]

Minamoto[2,15] has shown that these differences can give rise to pronounced effects in the properties of alloyed p–n junctions formed on one or other surface. Since p–n junction electrical characteristics are sensitive to the nature of the crystal surface[16] as well as to the presence of crystal defects in the junction region,[17,18] it is not unreasonable to assume that their electrical characteristics may be affected by the metallurgical and chemical phenomena associated with the different {111}type surfaces. In view of this and of the work of Minamoto[2,15] experiments have been carried out to determine what effects these phenomena may have on p–n junctions, whether these effects result from all or any of the polarity phenomena and what conditions are required to minimize these effects. In the light of these experiments crystal growth and dissolution are considered and a qualitative explanation is presented for the observed polarity effects.

EXPERIMENTAL PROCEDURE

Pulled crystals of indium antimonide having donor densities from 9×10^{13} to 10^{15} cm^{-3} and mobilities from 8×10^5 to 3×10^5 cm^2/V·sec at 77°K were used. Dislocation densities (as determined from etch-pit counts) were in general less than 1000 cm^{-2}. The crystals were cut on {111} surfaces to within ±1° by a high-speed slitting wheel. After fine polishing the slices were briefly etched to identify A and B surfaces and then cut into dice by an air abrasive cutting machine. All dice were given a further brief etch before alloying to ensure that the crystal surface was clean and to reveal any damage which might have occurred during preparation and handling. Spherical pellets of 99% In–1% Cd with diameters ranging from 0.2 to 0.5 mm were used for alloying.

The p–n junctions were prepared by two methods. For nonequilibrium experiments alloying was carried out in a hydrogen atmosphere on a strip heater of low thermal capacity. In this system the minimum time required for the temperature to rise from 25°C to its maximum value and to fall back to the indium freezing

[1] J. F. Dewald, J. Electrochem. Soc. 104, 244 (1957).
[2] M. T. Minamoto, J. Appl. Phys. 33, 1826 (1962).
[3] H. C. Gatos and M. C. Lavine, J. Electrochem. Soc. 107, 427 (1960).
[4] E. P. Warekois and P. H. Metzger, J. Appl. Phys. 30, 960 (1959).
[5] J. G. White and W. C. Roth, J. Appl. Phys. 30, 946 (1959).
[6] E. P. Warekois, M. C. Lavine, A. N. Mariano, and H. C. Gatos, J. Appl. Phys. 33, 690 (1962).
[7] H. C. Gatos and M. C. Lavine, J. Phys. Chem. Solids 14, 169 (1960).
[8] M. C. Lavine, A. J. Rosenberg, and H. C. Gatos, J. Appl. Phys. 29, 1131 (1958).
[9] J. D. Venables and R. M. Broudy, J. Electrochem. Soc. 107, 296 (1960).
[10] A. Steinemann and U. Zimmerli, Solid-State Electron. 6, 597 (1963).
[11] S. G. Ellis, J. Appl. Phys. 30, 947 (1959).
[12] H. C. Gatos, P. L. Moody, and M. C. Lavine, J. Appl. Phys. 31, 212 (1960).
[13] R. K. Mueller and R. L. Jacobson, J. Appl. Phys. 32, 550 (1961).
[14] K. F. Hulme and J. B. Mullin, Solid-State Electron. 5, 211 (1962).
[15] M. T. Minamoto and H. T. Malafi, J. Appl. Phys. 34, 1876 (1963).
[16] See, for example, A. L. McWhorter and R. H. Kingston, Proc. IRE 42, 1376 (1954); W. T. Eriksen, H. Statz, and G. A. J. De Mars, J. Appl. Phys. 28, 133 (1957).
[17] A. G. Chynoweth and G. L. Pearson, J. Appl. Phys. 28, 133 (1957).
[18] J. I. Pankove, J. Appl. Phys. 28, 1054 (1957).

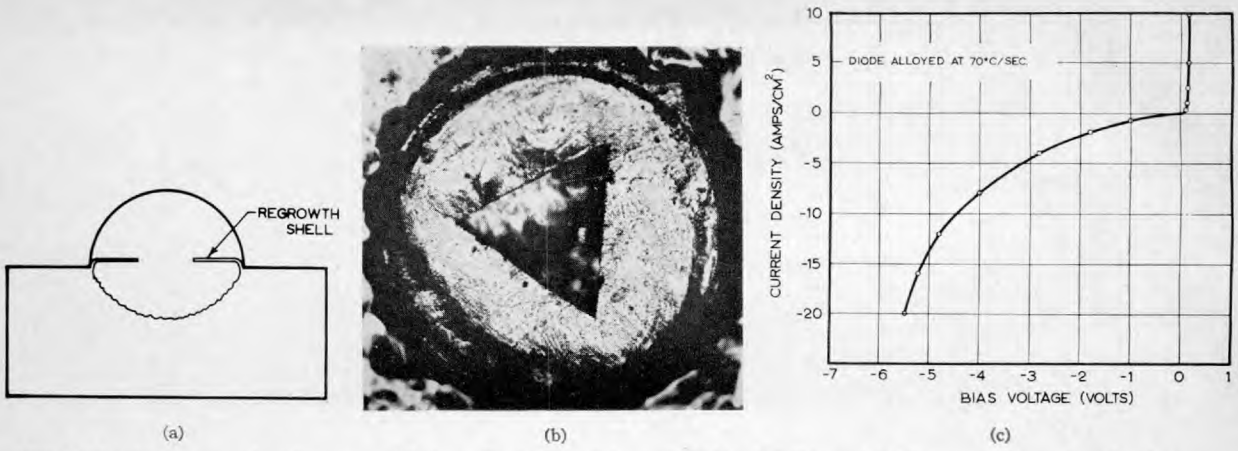


FIG. 1. Junction prepared under nonequilibrium alloying conditions, $(dT/dt)>50°$ C/sec. (a) Diagrammatic section showing round junction and regrowth shell. (b) Regrowth shell showing a triangular hole bounded by $\{\bar{2}\bar{1}\bar{1}\}$ surfaces. (c) Electrical characteristic at 77°K.

temperature was twelve seconds. The maximum temperatures used were varied between 300° and 400°C and the rate of temperature rise was controlled at values between 5° and 70° C/sec. For quasiequilibrium alloying the pellet was first wet onto the crystal without allowing any significant penetration and then the crystal and pellet were alloyed on a strip heater or in a suitably programmed furnace. Under quasiequilibrium alloying conditions the rate of temperature rise was controlled between 0.1° and 10°C/sec.

Junctions formed under the above conditions were examined in one or more of the following ways:

(1) The wetted area on the crystal surface was examined for any shape dependence on the alloying cycle.

(2) The indium dot was removed in warm dilute HCl and the surface of the regrowth region examined.

(3) Junctions were sectioned, polished, and etched to reveal the shape and perfection of the regrowth region and the extent of alloy penetration.

(4) Junctions were mounted on TO5 headers, etched, and tested electrically by measuring the direct current characteristics. The diodes were etched in diluted CP4 ($HF:HNO_3:HAc=3:5:5$) followed by a further light etch designed to leave the surface in a stoichiometric condition.

Results

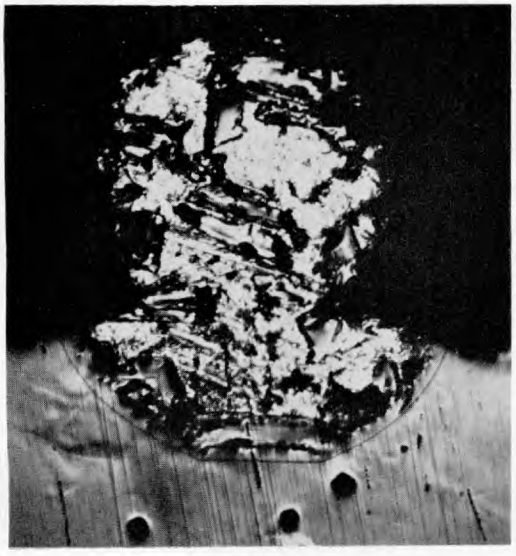
Over 170 diodes were made under the conditions described above. All were examined for any shape dependence of the surface wetting area, about seventy were sectioned and about thirty were etched to allow examination of the regrowth surface. For convenience the results will be described under three conditions of alloying.

(1) Nonequilibrium alloying, $(dT/dt)>50°$ C/sec. The regrowth regions produced under these conditions were round on both A and B surfaces, Fig. 1(a). The regrowth was of limited extent and the regrowth surface very uneven. In every case a single-crystal shell was formed above the original crystal surface, Fig. 1(a), and if this shell did not completely cover the regrowth region, the open portions were, in general, bounded by edges identified as $\{\bar{2}\bar{1}\bar{1}\}$ surfaces, Fig. 1(b). Since $\langle 211\rangle$ directions favor dendritic growth it is possible that this shell grew by a dendritic mechanism. Surface wetting areas were always round and the electrical characteristics, Fig. 1(c), were uniformly bad for both surfaces.

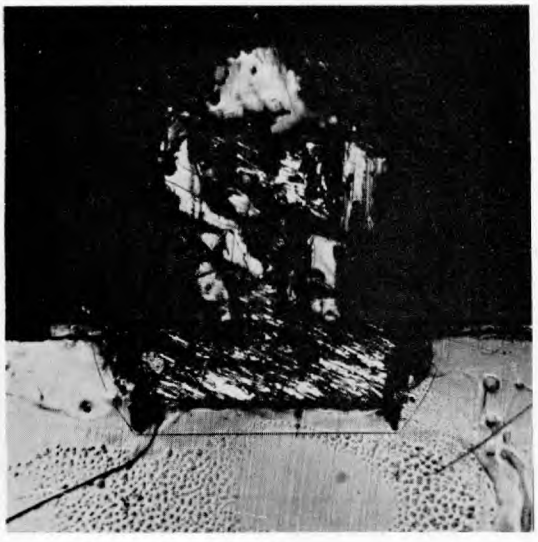
(2) Fast quasi-equilibrium alloying, $5°C/sec<dT/dt<10°C/sec$. In this region experiments were carried out using both alloying without pre-wetting and alloying with pre-wetting. In the first method wetting began only when the temperature reached approximately 300°C and then only took place over a limited area. By the completion of the alloy cycle the pellet had invariably assumed a pear-like shape with the neck in contact with the crystal. The final wetted area in all cases was circular and was larger than the pellet contact area. The spreading of the wetted area was greater on B surfaces than on A surfaces. Junctions formed on A surfaces were, in general, round with a slight flattening at the bottom as shown in Fig. 2(a) while those formed on B surfaces were flat with only slight rounding near the edges, Fig. 2(b). The regrowth was poor on both surfaces but there were usually larger areas of good regrowth on the B surfaces. Electrical measurements were not made on any of these junctions but it was assumed that the observations of Minamoto[2] would apply to this case.

When the pellets were pre-wetted, junctions on both surfaces were flat at the center and rounded at the edges, Fig. 3. However, in general, the wetted area and the flat area of the junction were larger on the B surface. Surface wetting areas on both surfaces tended to be round with slight flattening on three sides. The flattening was found to be associated with the three B planes intersecting the crystal surface.

(3) Slow quasiequilibrium alloying, $(dT/dt)<1°C/sec$. A few diodes were alloyed at a rate of 1°C/sec. None of these diodes were examined electrically but sections showed flat junctions on both A and B surfaces. The wetted area on the B surfaces was considerably larger than that on the A surfaces and often tended to be triangular, Fig. 4. The major side faces were again



(a)



(b)

FIG. 2. Junction prepared under fast quasiequilibrium alloying conditions with no pre-wetting. (a) Section of junction alloyed at 10°C/sec on the A surface. (b) Section of junction alloyed at 10°C/sec on the B surface.

found to be associated with the intersection of B planes with the crystal surface.

Most of the junctions examined were prepared under conditions in which the rates of rise and fall of temperature during the alloy cycle were less than 5° C/min. Flat junctions were obtained on both A and B surfaces

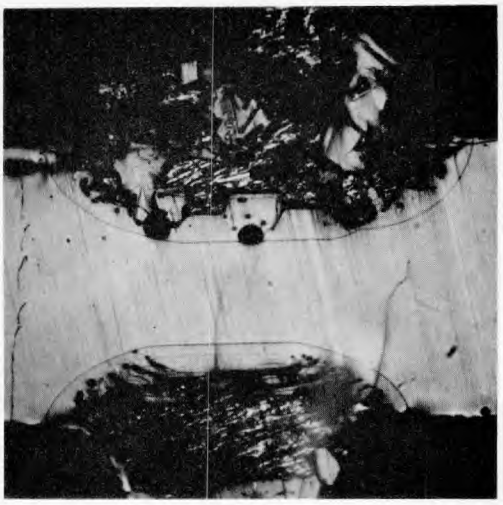


FIG. 3. Section of a double junction alloyed at 8°C/sec (A surface uppermost) after pre-wetting.

without exception. A double alloyed section is shown in Fig. 5(a). Although some variation in the shape of the wetted areas occurred, about 50% were almost perfectly hexagonal, Fig. 5(b).

Measured voltage–current characteristics on either surface showed rectification ratios typically of the order 10^7 and as large as 2×10^8. These ratios are between one and three orders of magnitude better than those previously reported in the literature.[19] Even after etching until the diodes assumed a configuration such as that of Fig. 5(c), the breakdown region seemed slightly softer and the reverse currents somewhat higher for diodes fabricated on A surfaces (Fig. 6). No consistent difference in breakdown voltages were observed for diodes made on either surface. The forward current densities

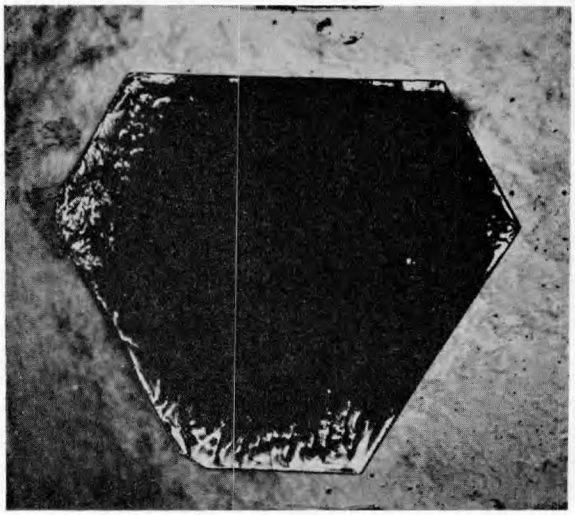
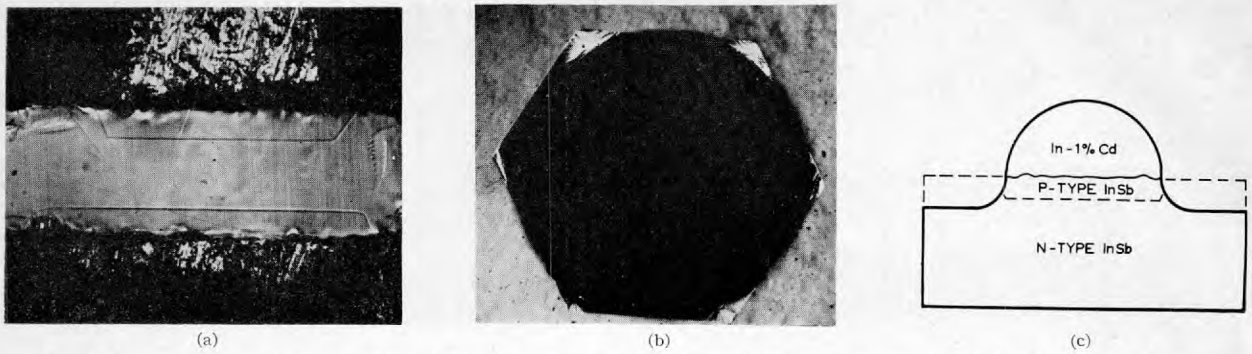


FIG. 4. Wetted area of junction alloyed at ~1°C/sec. The well developed sides are associated with B planes.

[19] H. J. Stocker, J. Appl. Phys. **32**, 322 (1961).



(a) (b) (c)

FIG. 5. Junction prepared by pre-wetting followed by slow alloying $(dT/dt) \sim 5°C/min$. (a) Section of double junction showing flat junctions on both A (upper) and B surfaces. (b) Hexagonal wetted area. (c) Diode configuration after thorough etching.

at low bias voltages were consistently higher for diodes formed on A surfaces, Fig. 7. The excess forward current could be decreased by careful etching but in the limit was still greater than that observed in diodes made on the B surface, even when the B surface diodes exhibited higher reverse current (Fig. 7).

DISCUSSION

In the simplest case the rate of dissolution of a substance in chemical[7] and electrochemical[1] reactions is limited by one or both of two restraints: One restraint arises from the dissociation energy associated with breaking the interatomic bonds and taking the freed atom into solution, the other from the need for the solute atoms to diffuse into the volume of the solvent. These same two mechanisms can determine the rate of metallic etching or alloying. Thus under conditions which depart far from equilibrium, i.e., when the rate of temperature rise during alloying is too great, the process of dissolution will be independent of crystal orientation and the solid–solute interface will show a diffusion-limited profile. When alloying is so limited the dissolved volume tends to take the shape of a diminished mirror image of the solute volume. This effect has been observed in junctions prepared under nonequilibrium conditions and explains the junction profile shown in Fig. 1(a).

The tendency for wetted areas to be bounded by a triangle of B surfaces as shown in Fig. 3 and to spread laterally more on the B than on the A surfaces gives direct evidence for a higher dissolution rate on A surfaces since the more quickly dissolving surface would tend to extinction in this situation. This evidence coupled with that of other authors[10,11] points to a considerable difference in the ease both of dissolution and of growth between A and B surfaces. One would then expect that under particular alloying conditions the dissolution process on one surface could be diffusion limited while that on the other could be rate limited. This effect has been observed in junctions prepared without

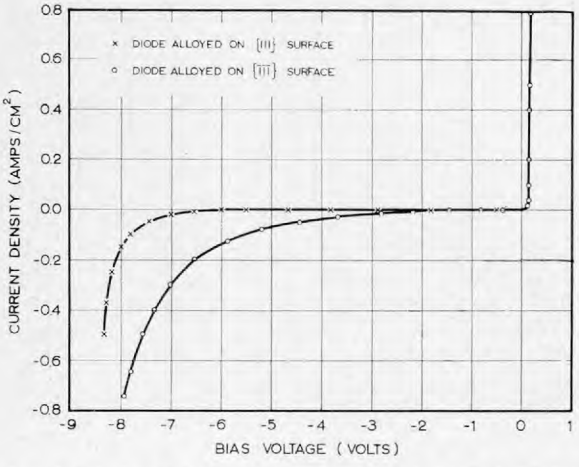


FIG. 6. Electrical characteristics at 77°K of diodes formed on A and B surfaces of 3.3×10^{14} n-type InSb. Diode areas are approximately 10^{-3} cm².

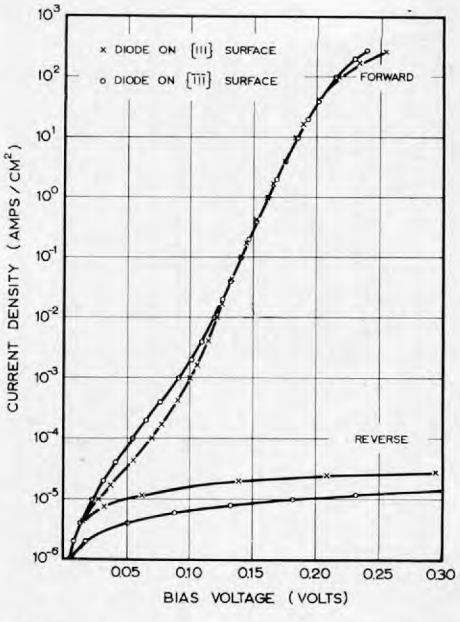


FIG. 7. Electrical characteristics at 77°K of diodes formed on A and B surfaces of 9×10^{13} n-type InSb for small applied voltages. Diode areas approximately 10^{-3} cm².

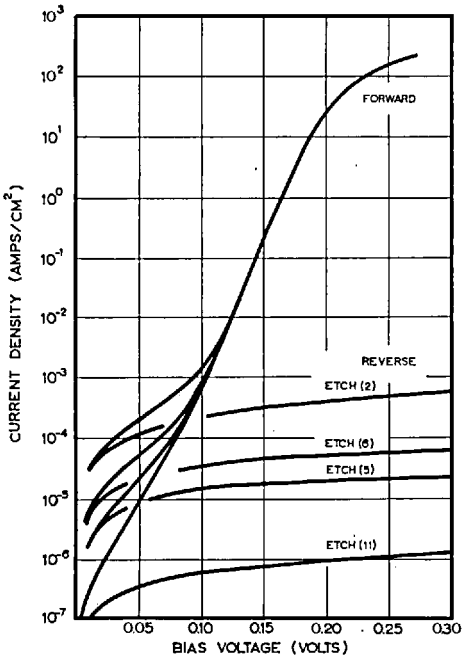


FIG. 8. Electrical characteristics at 77°K of a diode subjected to repeated etching. The etching procedure was the same at each stage. Etching stages not shown yielded reverse currents slightly in excess of those observed after etch (2). Area before etch (1) 3.20×10^{-3} cm², after etch (11) area $=1.05\times10^{-3}$ cm².

pre-wetting under the fast quasiequilibrium alloying conditions, Figs. 2(a) and (b). In this situation the contact angle between the pellet and the crystal is always less than 90° and hence over the contact area the pellet appears to be an infinite source of solvent. With no apparent shape limitations applied by the solvent volume dissolution proceeds limited only by either of the two restraints. In this case alloying on the B surface is largely rate limited while that on the A surface is diffusion limited. When pre-wetting is carried out, however, the contact angle between the pellet and the crystal is always greater than 90°. Even under conditions where rate limitation would occur on both surfaces for a truly infinite source, in this situation the shape of the solvent volume would cause diffusion limited profiles near the edges where the amount of solvent is limited. Junctions prepared with pre-wetting under fast quasiequilibrium alloying conditions have shown this tendency to be both diffusion limited and rate limited at the same time on both surfaces, Fig. 3.

In the light of this discussion it becomes apparent that if the rate of temperature rise were decreased until the diffusion rates into the solvent were unimportant, the alloy region on the A surface would also be bounded by flat {111} type planes. Under such conditions both A and B surfaces would give flat solute–solid interfaces, the resulting junction interfaces would be flat, Fig. 5(a), and the surface wetting boundaries would tend to be hexagonal, Fig. 5(b).

The electrical properties of p–n junctions formed by alloying to either surface are less readily explained. In the case of the rounded, diffusion-limited junctions it is easily imaged that if, on some fine enough scale, they consisted or arrays of {111} type faces, then reverse characteristics would be soft due to localized field concentrations.

When true planar junctions are formed on either surface, the difference in the observed current–voltage characteristics is so small that on a scale fifty times more sensitive than that used by Minamoto[2] no difference would be detected. The residual difference in electrical characteristic is probably due to polarity-dependent surface effects which almost certainly mask whatever true bulk effects may exist. Figure 8 illustrates the extreme sensitivity of diode characteristics to chemical treatment where no significant change of junction area is involved. The diode characteristics are also sensitive to ambient atmosphere and may under certain conditions display a large sensitivity to external illumination. This latter observation is well illustrated in the work reported by Mueller and Jacobson.[20]

It has been shown[21] that many metals present in small quantities in normal pure acids deposit very readily on InSb. In view of the polarity-dependent chemical behavior of surfaces it is likely that the rate of deposition and the equilibrium concentrations established at the surface are governed by its polarity type. Even when the junction surface departs from the {111} or $\{\bar{1}\bar{1}\bar{1}\}$ as illustrated in Fig. 5(c) it will be composed of surfaces such as $\{hh1\}$ and $\{h11\}$ which are polar in the same sense as the associated {111} or $\{\bar{1}\bar{1}\bar{1}\}$. In such circumstances differences in electrical characteristics could arise readily from polar surface phenomena. Thus although the slight differences observed in electrical characteristics of junctions prepared on opposite faces could be due to polarity-dependent effects in the regrowth region, the evidence suggests that these differences are more probably due to residual surface effects which are themselves polarity-dependent.

[20] R. K. Mueller and R. L. Jacobson, J. Appl. Phys. 35, 1524 (1964).
[21] G. B. Larrabee, J. Electrochem. Soc. 108, 1130 (1961).