ELECTRICAL PROPERTIES OF SILICON FILMS ON SAPPHIRE SUBSTRATES

A thesis presented for the degree of Doctor of Philosophy in the

University of London

by

ALEXANDER BELL MARR ELLIOT

Department of Electrical Engineering, Imperial College of

Science and Technology

and   Post Office Research Centre

      Martlesham Heath, Ipswich

January 1976

To Jan, Karen and Fraser

ACKNOWLEDGEMENTS

ABSTRACT

A critical evaluation of published information on the properties of heteroepitaxial silicon layers on sapphire substrates is given, including a detailed comparative study of all the available data on the electrical properties of these layers. This review is followed by an account of an investigation of the electrical properties of thin layers as a function of depth in the film using two independent techniques, both of which employ the progressive penetration of a depletion region to vary the thickness of the layer measured.

The first technique uses a deep depletion MIS Hall effect structure to measure the variation in carrier concentration and mobility through the epitaxial layer. Additional information is also obtained on the properties of the silicon-silicon dioxide and the silicon-sapphire interfaces. Because the MIS techniques is not capable of measuring the properties of all the films through to the sapphire interface a second technique, using a junction field-effect structure, is also used to obtain further carrier concentration and mobility in formation. The junction field-effect technique is not capable of measuring the properties of the region adjacent to the silicon dioxide interface and is therefore complementary to the MIS technique. It also provides information on the variation of carrier lifetime through the thickness of the film.

A concise discussion of the effect of variations in electrical properties on the performance, yield and reliability of MOS circuits fabricated in silicon-on-sapphire is given in the conclusions.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

In conventional monolithic semiconductor integrated circuit technology the individual components in a circuit are isolated from each other by the depletion region surrounding a reverse-biased p-n junction. Whereas this isolation technique has been adequate for a large number of applications there is a requirement, particularly in the higher frequency applications, for an alternative isolation method which avoids the parasitic capacitance associated with the reverse-biased p-n junction. A variety of 'dielectric isolation' processes have been investigated and one of the simplest and most promising approaches has been to deposit a thin heteroepitaxial film of silicon on a suitable insulating substrate. Single-crystal α- alumina, (hereafter referred to as sapphire,) has been the most commonly used substrate material, (Allison et al 1969) although some work has also been done using magnesium aluminate spinel, (Cullen et al 1969).

Most of the circuits fabricated in silicon on sapphire have been of the MOS type because they are more tolerant than bipolar circuits to the crystalline defects which inevitably occur in an imperfectly matched heteroepitaxial system. The silicon-on sapphire system has a number of advantages for MOS circuits. Firstly, although the individual transistors in MOS circuits do not require specific isolation there is always a possibility of spurious interconnection by means of inversion layers forming under thick oxide MOS structures. The use of separate islands of silicon, each isolated by the dielectric substrate, for all the circuit components completely eliminates the possibility of any such spurious interconnection. Secondly, the component packing density is improved because the 'channel-stopping' diffusions which are sometimes used to prevent the thick-oxide parasitic interconnection are

unecessary and also because diffused regions can be placed closer together without the possibility of depletion regions 'punching-through' to the adjacent diffusions. A third advantage is that if the films are only 1 to 2 microns thick and the source and drain are diffused through to the sapphire substrate significant speed improvements over bulk silicon circuits are obtained. The improvements result both directly from the reduction of junction capacitances and also from the ability to use gated silicon resistors (Wilcock 1971) as load components in place of the conventional load MOST.

When complementary circuits are considered further advantages accrue. By using a two-stage epitaxial process or a one-stage epitaxial process and a diffusion, separate islands of n and p-type silicon can readily be made for the p and n-channel transistors repectively in a complementary circuit. Moreover a new design of complementary circuit using silicon islands of a single conductivity type is also possible (Boleky 1970) if deep-depletion transistors of the type described by Heiman (1966) are used.

The two special silicon-on sapphire devices mentioned above, namely the gated resistor and the deep-depletion transistor, operate by conduction in the bulk of the film and hence their characteristics will be influenced by any variations in the carrier transport properties through the film. Although information on the variation of carrier transport properties in silicon on sapphire films of the order of 10 microns thick is available in the literature (Dumin and Robinson 1966 and 1968) and also some information on the average properties of thinner films there is little information on the variations in properties of films only 1 to 2 microns thick. The work reported herein was undertaken in an attempt to provide this missing information.

Before proceeding to the main experimental investigation a critical review of

all the published information on the electrical and structural properties of silicon on sapphire layers was conducted and this is presented in Chapter 2. Also included in this review is any information on the silicon on magnesium aluminate spinel system which is thought to be relevant to the silicon on sapphire system.

The profiling technique employed in the thicker films was to gradually thin the film by mechanical or chemical means and to make Hall and conductivity measurements at each step either using a Hall bar or by the van der Pauw (1958) technique. However if the film is only 1.5 μm thick with a donor concentration of 5×10^{15} cm^{-2} there are only 7.5×10^{11} donors per square cm of film and trapping of carriers at both silicon interfaces is likely to significantly influence the transport properties. Because the sapphire interface states will remain invariant during a set of measurements they will not influence the calculation of the properties of the removed layer unless the film is sufficiently thin for the space charge layer to penetrate the entire film. However the new air interface which is exposed at each removal step is unlikely to exhibit a constant surface state density and carrier density errors in excess of 10 per cent at a film thickness of 1.5 μm are probable for fairly moderate surface state density variations of 7.5×10^{10} cm^{-2}. Therefore, even if accurate thinning techniques could be developed for 1.5 μm thick films, the errors involved in using the conventional Hall techniques could become very large as the sapphire interface is approached.

An alternative method of varying the thickness of a conducting layer is by progressively driving a depletion region into the film. This method of thickness modulation which has the advantage that the depletion layer boundary is free of interface states is the basis of both the profiling techniques used in this investigation.

The experimental structure used in the first technique is a 300μm square deep depletion MIS transistor and the depletion region is driven in from the oxide interface by applying the appropriate gate potential. The MIS Hall and conductivity measurements as a function of gate voltage are combined with C-V measurements to provide both the carrier concentration and mobility profiles. These measurements and results are presented in Chapter 3.

Chapter 4 contains work done using the second technique which employs a large junction field-effect transistor with a channel width of 600μm and a channel length of 40μm. The carrier concentration profile is obtained from an analysis of the C-V characteristics of the reverse-biased gate junction and the mobility profile from these characteristic together with channel conductivity measurements as a function of gate voltage. Information on the variations in minority carrier lifetime through the film are also obtained from the gate leakage current as a function of the reverse-bias voltage.

In the final chapter the interpretation of the bulk and interface properties measured in Chapter 3 and 4 are discussed and their influence on the performance of silicon on sapphire MOST devices and integrated circuits is evaluated.

CHAPTER 2

PREPARATION AND PROPERTIES OF

HETEROEPITAXIAL FILMS OF SILICON ON

SAPPHIRE

Since 1964 workers at a number of laboratories have been studying the properties of thin single-crystal silicon films grown on amorphous and single-crystal insulating substrates. The majority of the investigators have used sapphire as a substrate but recently there has been an increasing interest in magnesium-aluminium spinels. Because the experimental work in this thesis uses silicon films on sapphire the major emphasis in the following review will be placed on the silicon—on-sapphire system but reference will be made to results obtained from other silicon on insulator composites where they contribute to an understanding of the silicon-on-sapphire system.

2.1 Deposition Techniques

A substrate surface with a relatively low density of crystalline and topographical faults is a prerequisite for the growth of high quality layers by any deposition technique since disturbances of the periodic field at the substrate surface may give rise to defects in the silicon layer during the early stages of growth. The fault density on the substrate surface is determined both by the perfection of the substrate material and by the quality of the surface preparation but, since P. B. Hart et al (1967) have shown that layers grown on highly faulted flame-fusion material have similar properties to layers grown on more perfect crystals pulled from the melt by the Czochralski technique, it would appear that neither the surface dislocation densities of the order of 10^6 cm^{-2} nor the low-angle grain boundaries which are both present in flame-fusion material (Stephens and Alford 1964) are a limiting factor in determining layer quality. However, a number of workers, (e.g. Robinson and Mueller 1966) have shown that the properties of the silicon layers are critically dependent on the substrate surface preparation.

The first stage in surface preparation is a mechanical polishing with successively finer grades of carborundum or diamond powders. Attempts to remove the crystal damage near the surface which results from this polishing by wet chemical processes have not been entirely successful. Among the etches that have been used are phosphoric acid at $400^{\circ}C$-$500^{\circ}C$ (Scheuplin and Gibbs 1960), lead fluoride at $850^{\circ}C$ (Robinson and Mueller 1966), and potassium bisulphate at $450^{\circ}C$, but not one of these has produced consistently smooth surfaces. Borax at $1000^{\circ}C$, (Robinson and Mueller 1966) and vanadium pentoxide at $900^{\circ}C$ (Faktor et al 1967), have reproducibly yielded smooth surfaces but both processes leave surface residues which are difficult to remove.

Gas phase polishing processes have been considerably more successful than the liquid etches. Sulphur fluorides in a helium carrier gas at $1350^{\circ}C$ (Manasevit and Morritz 1967), hydrogen chloride in hydrogen at $1400^{\circ}C$ (Chu et al 1965) and hydrogen itself at temperatures up to $1500^{\circ}C$ (Robinson and Mueller 1966) have all been used to obtain smooth flat polished surfaces free from surface residues. These processes have the further advantage that they can be used in situ, immediately prior to the growth of the silicon layer, and the hydrogen treatment has tended to become a standard pre-deposition treatment employed by the majority of workers in this field. Silicon vapour has also been used to polish sapphire. Reynolds and Elliot (1966) used vacuum evaporated silicon, Filby (1966) used a dilute mixture of silane in hydrogen and an H_2-$SiHCl_3$. gas mixture was used by Tsunoda (1966). In the vapour transport systems a temperature in excess of $1300^{\circ}C$ is necessary to obtain smooth flat surfaces at sapphire removal rates in excess of 0·1 μm/min but similar results can be achieved with vacuum evaporated silicon at $1150^{\circ}C$. Chang (1971) has achieved polishing action at $925^{\circ}C$, using a vacuum sublimation

source, but the silicon arrival rate was only 60 $\overset{o}{A}$/min. He has also demonstrated by L.E.E.D-Auger studies that a heat treatment for 30 seconds at $1400^{o}C$ in vacuum removes carbonaceous contamination from the surface and causes surface atoms to rearrange themselves into an equilibrium structure which remains stable on cooling.

The majority of the silicon-on-sapphire research has been with vapour transport systems using quartz reaction vessels and the processes used were either -

1.   Hydrogen reduction of silicon tetrachloride or trichlorosilane (Joyce et al 1965 and Manasevit et al 1965), or

2.   Pyrolysis of silane at reduced pressures (Joyce et al 1965) or diluted in hydrogen (Mueller and Robinson 1964 and Manasevit et al 1965).

The films prepared by the pyrolysis of silane were found to be crystallographically and electrically superior to those prepared by the hydrogen reduction processes (Zaminer 1968) because the hydrogen chloride reaction product of the latter processes etches the sapphire surface producing a rough multi-faceted surface before complete silicon coverage is achieved. Since there is not a unique substrate-overlayer crystallographic relationship in the silicon-on-sapphire system (see Section 2.2), the multi-faceted surface results in the formation of nuclei of different orientations which generate crystal faults when they coalesce.

In the atmospheric pressure silane system a dilute mixture of silane in hydrogen, typically one part per thousand, is made to flow over a sapphire substrate which is heated to a temperature in the range $1000^{o}C$ to $1200^{o}C$ on an r.f. heated susceptor. Whereas in the reduced pressure silane system, the silane flow to the reaction vessel is adjusted to maintain a pressure of 1 torr in the vessel and the susceptor has to be resistively heated because r.f. heating at this pressure would result in a gas discharge. Substrate

13

temperatures are again in the range 1000°C to 1200°C. In both systems the silane thermally decomposes on the hot sapphire surface depositing silicon atoms which are sufficiently mobile to form a single-crystal film. Doping of the film can be achieved by the introduction of small quantities of phosphine or diborane into the gas stream.

Using the Hall mobility of majority carriers as an index of the layer quality, Dumin (1967) has shown that the optimum growth temperature for (111) layers on (0001) sapphire is 1200°C and for (100) layers on (1$\bar{1}$02) sapphire is 1115°C. The films were grown at rates between 0·3 μm/min to 3 μm/min to a thickness of 9 to 10 μm and the mobility maximum was 80-100% of bulk silicon mobility for the (111) films and 60% of bulk silicon mobility for the (100) films. The (1$\bar{1}$02) results have been substantiated by Mercier (1970) who reported an optimum temperature of 1110°C for 5 μm thick (100) layers grown at a rate of 0·1 μm/min. When growing very thin layers (0·15 μm), deposition rates as high as 10 μm/min at 1150°C are reported to be necessary in order to achieve optimum layer quality (Dumin et al 1970).

Mataré (1969) has predicted that the faults formed by the coalescence of misaligned nuclei can be reduced by using a faster initial growth rate and thereby generating a higher density of smaller nuclei which can re-align more readily during coalescence. Experimental work by Hart et al (1967) has demonstrated the improvemnents that can be made by applying this technique to thicker films but it is not generally employed in the growth of the 1-2 μm thick films used for vertical junction MOS device structures.

Since the use of hydrogen as a carrier gas inhibits the silane dissociation and aids the transfer of aluminium from the substrate to the growing film, Chiang and Richman (1970) and Mercier (1971) have recently investigated the use of helium as an alternative carrier gas. Epitaxial layers were grown at 900°C on (1$\bar{1}$02) sapphire at a growth rate comparable to

that normally employed at 1100°C with a hydrogen carrier gas but the film orientation was (110) and no mobility values were measured. Films grown at higher temperatures were (100) orientation and had mobilities comparable to those obtained with a hydrogen carrier gas. From his observations of the nucleation processes in a helium carrier gas system Mercier gives the decrease in density and increase in size of the nuclei as one of the reasons for the lower epitaxial temperature but this is contradictory to the predictions of Mataré. The observed reduction of the contact angle to less than 90°, which shows the increased influence of the substrate, may be the more significant difference in the nucleation behaviour.

Vacuum deposited layers with properties comparable to layers grown from silane have been prepared by evaporation and by sublimation in ion-pumped and diffusion-pumped systems which were evacuated to pressures in the range 2×10^{-5} torr to 1×10^{-9} torr, (Reynolds and Elliot 1967, Naber and O'Neal 1968, Weisberg and Miller 1968 and Itoh et al 1968a). Lawson and Jefkins (1970) observed no significant differences in film quality in films prepared over the pressure range 2×10^{-5} torr to 10^{-7} torr. Epitaxial layers can be grown at much lower temperatures in a vacuum system than they can in a vapour transport system; e.g. good single-crystal layers can be grown on a 900°C substrate at a deposition rate of 0·1 μm/min. The substrate temperature necessary to obtain optimum layer quality in vacuum evaporated films is a function of the deposition rate (Lawson and Jefkins 1970), but it is found that if the substrate temperature for deposition ($T_D$$^{\circ}$C) is expressed as a fraction of the substrate temperature necessary for the impinging silicon to etch the substrate ($T_E$$^{\circ}$C) then the optimum layer quality in films 5 μm to 12 μm thick is achieved when (T_D/T_E) is equal to 0·91 for (0001) substrates and 0·92 for (1$\bar{1}$02) substrates. The optimum quality for films on (0001)

substrates corresponds to a majority carrier mobility of up to 75% of the bulk silicon value but the results were not very reproducible. The ($1\bar{1}02$) results were reproducible but the optimum mobility was only 45% of the bulk silicon mobility.

A variety of evaporation and sublimation sources have been used but electron bombardment of a silicon charge on a water cooled copper hearth has been the most common system because it provides a high deposition rate together with minimal source contamination. This is achieved by always having a solid skin of silicon in contact with the hearth and by using electrostatic or magnetic focussing of the electron beam to eliminate any direct rectilinear path between the hot filament and either the source or the substrate. Deposition rates are generally controlled by arranging for the silicon flux to fall on a vibrating quartz crystal and a signal proportional to the rate of change of vibrational frequency is used to control the filament current in the electron beam evaporator. Doping facilities were not included in many of the vacuum systems but Itoh and his co-workers (1968a) described an antimony source which enables carrier concentrations of n-type films to be controlled in the range 3×10^{16} cm^{-3} to $1 \cdot 5 \times 10^{17}$ cm^{-3} and similar results were achieved by Lawson and Jefkins (1970) using an auxiliary phosphorus-doped silicon sublimation source.

Silane vapour transport systems have generally been preferred to vacuum systems because –

    1.   the capital cost is less

    2.   a large throughput can be achieved without the use of complicated substrate assemblies and

    3.   growth-rate control and doping control are simpler.

2.2 Crystalline Properties

A large density of crystalline defects is to be expected in heteroepitaxial systems like silicon-on-sapphire and silicon-on-spinel which have significant mismatches between their substrate and overlay values of thermal expansion coefficient and lattice spacing.

Manasevit and Forbes (1966) have shown that by fitting three silicon unit cells to two spinel unit cells a lattice mismatch of only 1.9% is possible in the silicon-on-spinel system. In silicon-on-sapphire the mismatch is anisotropic because the silicon structure is cubic and the sapphire is rhombohedral-hexagonal and it is as high as 12% in one direction with (100) silicon on ($1\bar{1}02$) sapphire, (Nolder and Cadoff 1965). The mismatch of thermal expansion coefficients is also greater on sapphire than on spinel.

Other major fault-inducing factors are the reaction between the deposit and the substrate, (Reynolds and Elliot 1966), the nature of the nucleation and growth processes, (Blank and Russell 1966) and the crystallographic and topographical defects in the sapphire surface. All of the above factors, with the exception of the thermal mismatch, will introduce faults at the sapphire or spinel interface, many of which will be eliminated as the film growth proceeds, (Pashley 1965). Therefore a high density of faults is to be expected at the insulator interface and the density will decrease as the distance from the interface increases.

X-ray diffraction, chemical etching followed by replica electron microscopy and transmission electron microscopy techniques have all been employed to investigate the defect structures, (Bicknel et al 1966, Nolder et al 1965, Schlotterer and Zaminer 1966 and many others) and the general conclusions to be drawn from the results is that the following faults occur in varying densities and proportions depending upon the growth system and the growth conditions:-

(1) low angle grain boundaries,

(2) stacking faults and microtwin lamellae and

(3) dislocations.   17

The low angle grain boundaries, which are common in layers grown on sapphire substrates cut from boules prepared by the flame-fusion process, are simply a replication of the structure present in the substrate. Many of these boundaries can be eliminated by annealing in hydrogen at $1250^{\circ}C$ to $1350^{\circ}C$, (Robinson and Mueller 1966) but this anneal will exacerbate the aluminium autodoping problem discussed in section 2.4.

The stacking faults and microtwin lamellae in layers on sapphire have been studied by Nolder et al (1965) and Bicknell et al (1966) and those in layers on spinel by Schlotterer and Zaminer (1966). Primary and secondary twins about $\{111\}$ planes have been observed in both systems but there is insufficient quantitative information to compare the densities of defects in the two systems. Virtually all the lamellar faults can be shown to originate at the substrate interface, hence they are clearly a product of the growth process and do not result from the stress generated by the thermal expansion mismatch between the substrate and the overlay.

Zaminer (1968) has shown that the higher reaction rate between silicon tetrachloride and spinel, as compared with silane and spinel, gives rise to a higher density of twins in the former system and Nolder et al (1965) have obtained similar results for silicon-on-sapphire layers but it is not known whether the increased twin densities are caused by residual reaction products on the substrate suface or by topographical defects resulting from the etching. Seiter et al (1965) have shown that the coalescence of 'out of phase' nuclei in heteroepitaxial systems which do not have a 1:1 atomic match is likely to generate stacking faults.

The dislocation densities measured from etch-pit studies have ranged from $10^{7} cm^{-2}$ to $10^{9} cm^{-2}$ with once again no clear difference between layers on sapphire and layers on spinel. The lower densities were observed on layers a few tens of microns thick. Linnington (1970) has studied the variation in dislocation density through films a few microns thick by transmission electron

microscopy and he observed as much as two orders of magnitude increase in density between the top surface of the layer and the substrate interface.

Nearly all of the reported structure studies were concerned with 'as-deposited' films and therefore little is known of the structure of films after annealing treatments in various ambients. Booker and Stickler (1965) have shown that when damaged bulk silicon surfaces are annealed in oxygen oxide platelets can form and similar results are to be expected in these highly-faulted silicon-on-sapphire layers. Furthermore, if the oxide platelets do form, aluminium will segregate into them since it has a segregation coefficient greater than 10^3, (Attala and Tannenbaum 1960). Although no direct evidence of this phenomena in silicon-on-sapphire has been reported, it has been observed that when the films used in this work were subjected to long anneals in oxidising ambients not only did the aluminium concentration fall but also the mobility decreased and this is presumably due to the generation of new scattering centres which may well be the oxide platelets described above.

The major significance of this section is that it has indicated that variations in the free carrier density and in the scattering properties through the thickness of the film must be given full consideration when the electrical results reported in this thesis are analysed.

2.3 Stress

When a thin silicon film on a sapphire or a spinel substrate is cooled from the deposition temperature to room temperature a compressive stress is induced in the silicon film because the thermal expansion coefficient of the silicon is less than that of the substrate. The following expressions for the interfacial and surface stresses in the silicon film have been derived by Jefkins (1970),

$$(S_1) \text{ interface} = \frac{(1 + 3 \epsilon p^2 + 4\epsilon p^3) \, Y_1 \, (\alpha_2 - \alpha_1) \, \Delta T}{(1 + 4\epsilon p + 6\epsilon p^2 + 4\epsilon p^3 + \epsilon^2 p^4)(1-\nu_1)} \quad \dots (2.3.1)$$

and

$$(S_1) \text{ surface} = \frac{(1 - 3\epsilon p^2 - 2\epsilon p^3) \, Y_1 \, (\alpha_2 - \alpha_1) \, \Delta T}{(1 + 4\epsilon p + 6\epsilon p^2 + 4\epsilon p^3 + \epsilon^2 p^4)(1 - \nu_1)} \quad \dots (2.3.2)$$

where the subscript 1 denotes silicon, the subscript 2 denotes the substrate, α, x, Y and ν are the thermal expansion coefficient, material thickness, Youngs modulus and Poissons ratio respectively and the following substitutions have been made,

$$p = x_1/x_2$$

and

$$\epsilon = \frac{Y_1 \, (1 - \nu_2)}{Y_2 \, (1 - \nu_1)}$$

All the specimens used in this work satisfy the condition $P < 0.002$ and therefore equations (2.3.1) and (2.3.2) reduce to

$$(S_1) \text{ interface} \simeq (S_1) \text{ surface} \simeq \frac{Y_1 \, (\alpha_2 - \alpha_1) \, \Delta T}{1 - \nu_1} \quad \dots (2.3.3)$$

The above analysis is only strictly valid if the thermal expansion coefficients are isotropic in the plane of the interface. Whereas this is true for silicon, spinel and the (0001) sapphire plane, it is not true for the ($1\bar{1}02$) sapphire plane which is the orientation of all the substrates used in this work. The minimum value of the thermal expansion coefficient in the ($1\bar{1}02$) sapphire plane is 8.31×10^{-6} K^{-1} and it occurs in the direction perpendicular to the (0001) plane, which in the case of the ($1\bar{1}02$) plane means it is also perpendicular to the ($2\bar{1}\bar{1}0$) plane, (see the accompanying stereographic projection). The maximum expansion coefficient is in the direction perpendicular to the plane marked X in the stereographic projection below and its value is 8.82×10^{-6} K^{-1}.
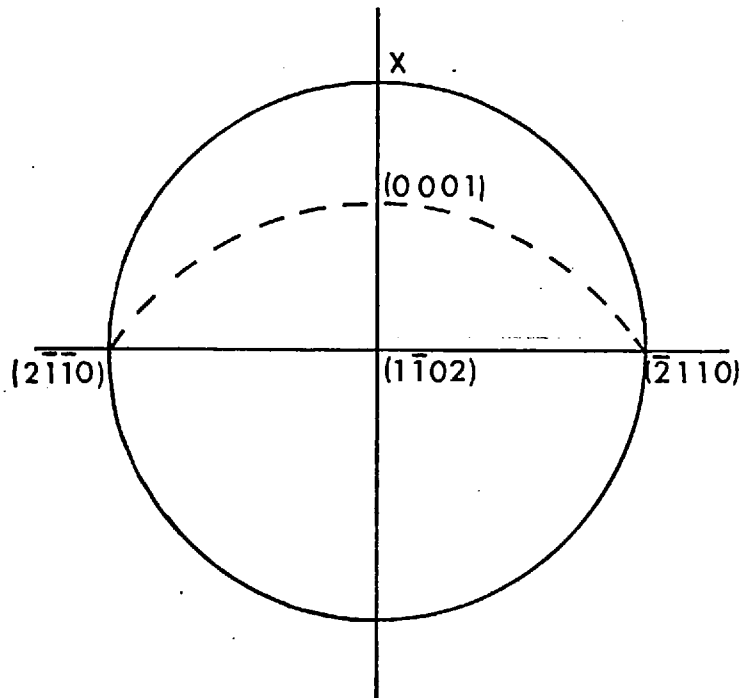
Since the difference between the maximum and minimum values of the expansion coefficient is only of the order of 6 per cent, an approximate analysis based on a mean value of 8.6×10^{-6} K^{-1} should enable a sufficiently accurate estimate of the stress levels to be determined. Equation (2.3.3) also requires a value of Youngs modulus for silicon and this can be determined from the relationship.

$$\frac{1}{Y} = S_{11} - 2 (S_{11} - S_{12} - \frac{1}{2} S_{44})(\ell_1^2 \ell_2^2 + \ell_2^2 \ell_3^2 + \ell_3^2 \ell_1^2) \ldots (2.3.4)$$

where the S_{ij} are compliance coefficients and the ℓ_k are directional casines. This equation is valid for any cubic system (Nye 1957) and the value of the compliance coefficients for silicon at $20^\circ C$ are

$$S_{11} = 7.68 \times 10^{-12} \ m^2 \ N^{-1}$$

$$S_{12} = -2.14 \times 10^{-12} \ m^2 \ N^{-1}$$

$$S_{44} = 12.56 \times 10^{-12} \ m^2 \ N^{-1}, \text{ (Runyan 1965)}.$$

which means that Young modulus for silicon is also anisotropic in the (100) plane. The second term in equation (2.3.4) is always negative therefore Young modulus is a minimum in the $<100>$ direction where the directional cosine term reduces to zero. Substitution of this minimum value of $1.3 \times 10^{11} Nm^{-2}$ in equation (2.3.3) determines a lower limit for the stress in this plane and an upper limit can be determined from the maximum Youngs modulus value of $1.69 \times 10^{11} \ Nm^{-2}$ If these values are substituted into equation (2.3.3) together with a value of 0.25 for Poissons ratio, (Runyan 1965) and 4.1×10^{-6} K^{-1} for the mean thermal expansion coefficient for silicon (Nan and Yi-Huan 1964) then the upper and lower stress limits shown as dashed lines in Figure 2.3.1 can be predicted for a film grown at $1050^\circ C$.

GROWTH TEMPERATURE = 1050 °C

STRESS, Nm^{-2}

UPPER LIMIT

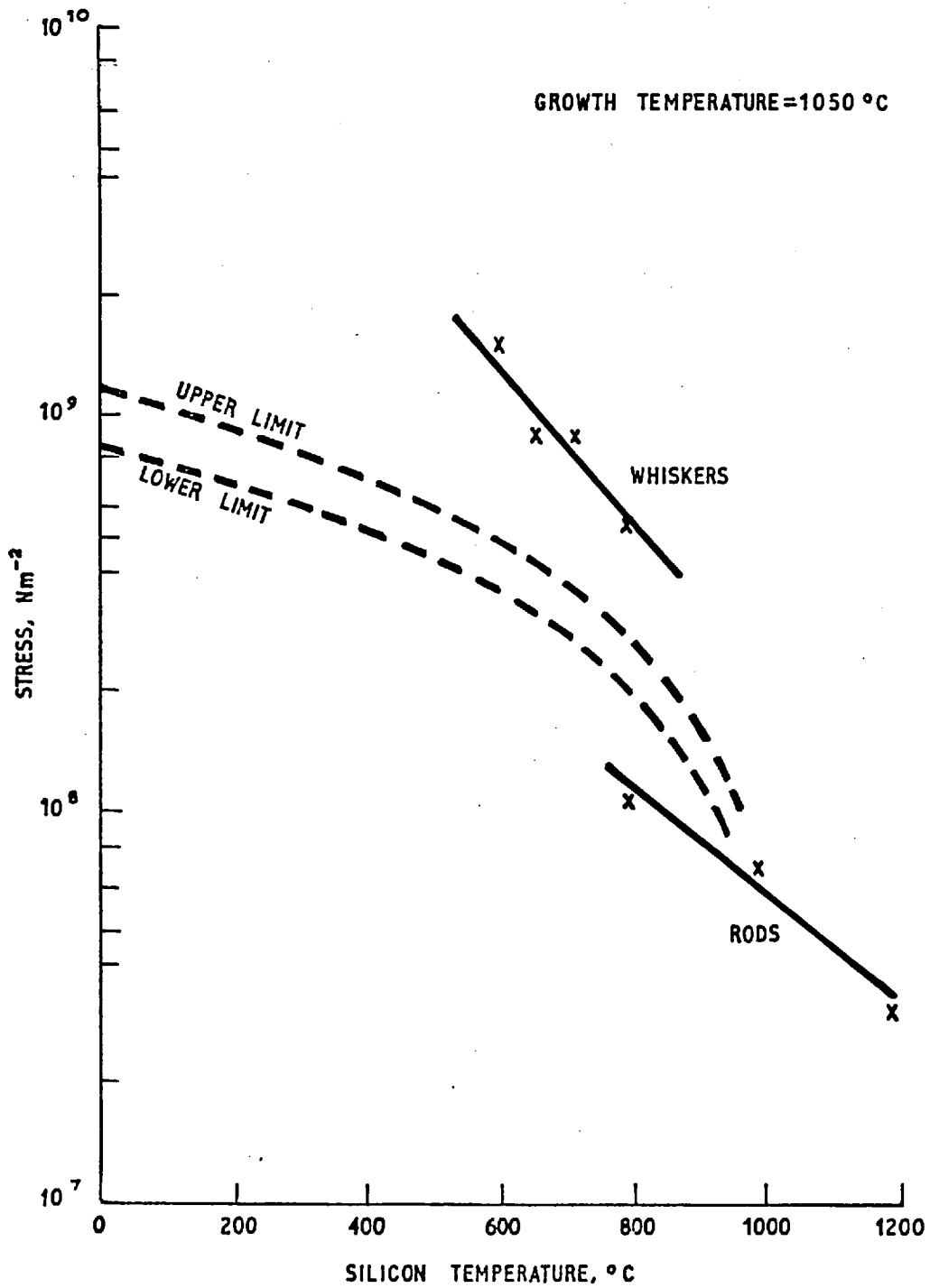LOWER LIMIT

WHISKERS

RODS

SILICON TEMPERATURE, °C

FIG. 2.3.1. COMPARISON OF CALCULATED STRESS
LEVELS IN SILICON-ON-SAPPHIRE FILMS AND
EXPERIMENTAL DATA ON THE ELASTIC LIMIT FOR SILICON

Also shown in Figure 2.3.1 as full lines are experimental results for the elastic limit as measured on vapour grown whiskers and on rods cut from bulk silicon, (Pearson et al 1957). It follows from these results that, if the film had elastic properties similar to the bulk silicon rods, a significant degree of plastic deformation would occur during the cooling, thereby relieving at least part of the stress. But crystallographic studies showed no evidence of dislocation arrays arising from plastic deformation, (see previous section) therefore the elastic properties of the films must resemble more closely those of the vapour grown whiskers.

No stress measurements of (100) silicon films on ($1\overline{1}02$) sapphire are reported in the literature. Ang and Manasevit (1965) have used a cantilever beam technique to measure the compressive stress in a $1.8\mu m$ thick (111) film on a $16\mu m$ thick (0001) sapphire ribbon. The film was grown from silicon tetrachloride at $1150^{\circ}C$ and the calculated stress value, making a correction for the partial deposit on the reverse side of the ribbon, was $9 \times 10^{8}Nm^{-2}$. Dumin (1965) has also measured stress values in the range $10^{8} - 10^{9}$ Nm^{-2} for (111) films on (0001) substrates. All these results are in fairly good agreement with the room temperature values in Figure 2.3.1, although some discrepancy would be expected because of the small differences in Y_1, α_2 and ΔT. This evidence substantiates the crystallographic evidence that there appears to be little stress relaxation due to plastic deformation. Dumin analysed the deformation of $1-46\mu m$ thick silicon films on 0.95 cm diameter sapphire discs, whose thicknesses were in the range $127-508\mu m$, using linear beam theory. The deformations were isotropic as expected. An anomalously large increase in stress with increasing substrate

thickness was observed. Since the value of p for all the layers was less than 0.06 only a very weak dependence of stress on substrate thickness is to be expected, (Jefkins 1970).

The stress in silicon films on alumina-rich spinel substrates has been studied by Schlotterer (1968) and by Robinson and Dumin (1968). Robinson and Dumin examined 2-25 μm thick silicon films on 0.025 cm and 0.050 cm thick spinel. The spinel composition was in the range $Mg\, O : Al_2\, O_3 = 1 : 3.0 - 3.5$ and (111), (110) and (100) orientations were investigated. All the layers exhibited an anomalous anisotropic deformation with the maximum anistropy occurring on (111) specimens. Maximum stress values in the range $1-4 \times 10^9\, N\, m^{-2}$ were observed, which is up to five times the values obtained on sapphire. Schlotterer observed isotropic deformation on (111) and (100) spinel ($Mg\, O : Al_2\, O_3 = 1 : 2 - 3.5$) with a typical stress value of $8.8 \times 10^8\, N\, m^{-2}$ on (100) substrates. Seiter (1967) has also demonstrated that the stress is isotropic by examining the stress dependent etch figures associated with crystal defects. The differences between the Robinson and Dumin and the Schlotterer and Seiter results may originate from differences in the quality of the spinel substrates used, since exsolution of α-alumina at crystal defects is a frequent problem with alumina-rich spinel.

2.4 Autodoping

One of the primary drawbacks of both the silicon-on-sapphire and the silicon-on-spinel systems is that the reduction of the substrate during film growth at elevated temperatures releases aluminium acceptors into the growing film. This reduction proceeds due to the reaction

$$2\ Si + Al_2O_3 \longrightarrow Al_2O + 2\ SiO \qquad\qquad (2.4.1)$$
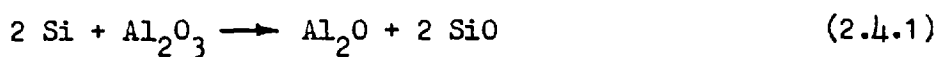
$$\text{(von Grube et al 1949)}$$

provided that the reaction products can escape. A further reaction with the hydrogen carrier gas is also possible

$$2\ H_2 + Al_2O_3 \longrightarrow Al_2O + 2\ H_2O \qquad\qquad (2.4.2)$$
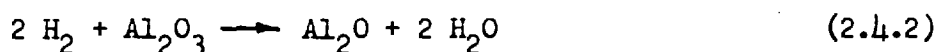
$$\text{(Manasevit et al 1965)}$$

Both of the above reactions have been used to remove the substrate surface damage left by the mechanical polishing, (Reynolds and Elliot 1966, Filby 1966 and Robinson and Mueller 1966).

The presence of aluminium in the silicon film has been positively identified by emission spectroscopic analysis (Dumin and Robinson 1966). The mechanism by which the aluminium sub-oxide is incorporated into the growing film will be discussed later.

Clearly to reduce autodoping it is desirable to grow the film at as low a temperature as possible but if the temperature is too low the crystalline quality of the film will suffer due to the lack of mobility of depositing atoms. Therefore, at any given growth rate, a temperature must be chosen which achieves a compromise between minimising the auto-doping and maximising the crystalline perfection. The compromise will depend on the application for which the film is to be used.

The relationship between the level of autodoping, as defined by the hole concentration determined from Hall measurements on undoped layers, and the growth temperature has been investigated by Robinson and Dumin (1968) for growth on both sapphire and $M_gO:3.5 Al_2O_3$ spinel substrates in a silane system and by Reynolds and Elliot (1967) for vacuum deposition on a sapphire substrate and the results are shown in Figure 2.4.1.

The ordinate in the figure is expressed as a density per unit film area because both sets of authors considered the dopant source to be a limited source produced during the nucleation period. Also included in Figure 2.4.1 are single temperature points which were obtained by other workers and althoug these confirm the general form of the results obtained by the above authors they also indicate the fairly wide spread in values that have been obtained under nominally similar growth conditions.

Comparing first the sapphire and the spinel results obtained with the silane system it is apparent that the autodoping on spinel at any given temperature is approximately an order of magnitude less than it is on sapphire. The less reactive nature of the spinel surface has been demonstrated directly by Schlotterer (1967).

A comparison between the vacuum-deposited and the silane-grown films shows that the temperature at which a given autodoping level occurs is approximately $150-250^\circ K$ lower for the vacuum-deposited layers. Some insight into why this should be the case can be obtained by considering the differences in the growth processes. During the nucleation stage of vacuum-deposited films some of the depositing silicon atoms will react with the substrate according to equation (2.4.1), but since the Al_2O reaction product which sublimes in a vacuum will not return to the growing film, the aluminium in the growing film must result from reaction at nuclei perimeters or from surface migration. When the surface is completely covered further aluminium autodoping will be
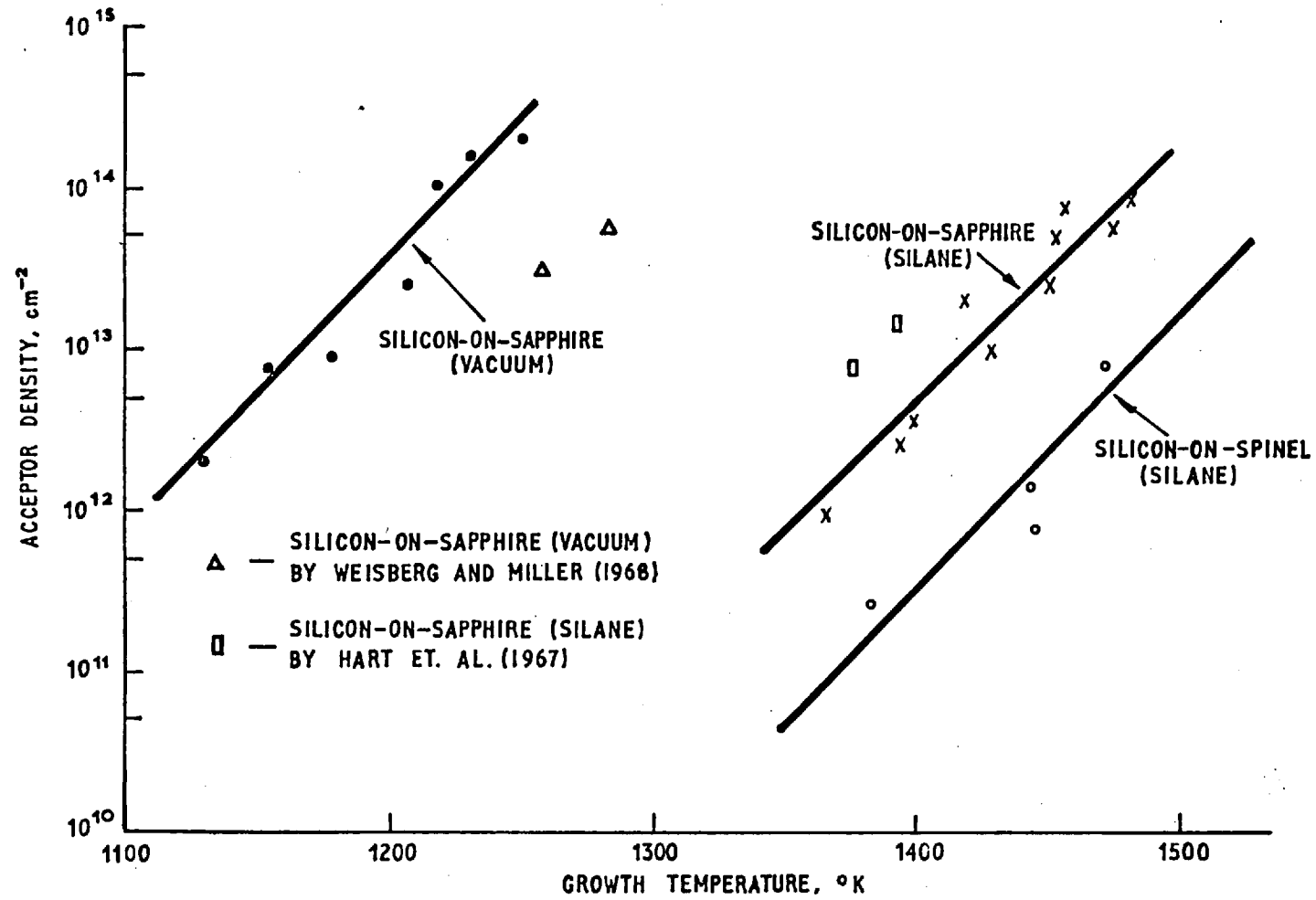
27

FIG. 2.4.1. ALUMINIUM AUTODOPING AS A FUNCTION OF THE FILM GROWTH TEMPERATURE

28

limited by the diffusion of the reaction products away from the interface. A possible complication in practical vacuum systems is the existence of partial pressures of hydrogen in the range 10^{-7} to 10^{-8} torr, (Lawson and Jekfins 1970 and Weisberg and Miller 1968). These partial pressures correspond to hydrogen arrival rates of a monolayer in a few tens of seconds. But at temperatures around $1000^{\circ}C$ the mean time of stay of hydrogen atoms is likely to be much too short for a significant number to react with the alumina.

A superficial consideration of the situation may induce one to believe that the silane system would be expected to give rise to higher autodoping levels because the volatile reaction products can be transported back to the surface and because the hydrogen carrier gas is present at atmospheric pressure but the more complex growth process that is believed to occur in the case of growth from silane invalidates this argument. It is probable that the silane molecule is only partially decomposed leaving SiH_x species on the alumina surface. Additional bonding between hydrogen atoms in the SiH_x and oxygen atoms in the substrate would make the SiH_x species less mobile than silicon atoms and therefore could account for the higher epitaxy temperatures in the silane system. If the growth proceeded in this manner the autodoping caused by the silicon reaction is also likely to be reduced because of the lower concentration of free silicon.

If the back of the slice is sealed off with a silicon film prior to the epitaxial growth on the top surface then the level of autodoping is reduced, (Hart et al 1967 and Mercier 1970). This transfer from the back surface may be due to either a hydrogen or a silane reaction. For a given crystalline quality the autodoping level in vacuum-deposited and silane-grown films are similar because the lower epitaxial temperature compensates for the inherently faster autodoping reaction rate.

Recently Chiang and Richman (1970) and Mercier (1971) have used helium as the silane carrier gas in an attempt to suppress the autodoping contribution from the hydrogen carrier gas although some hydrogen will still be present as a product of the silane pyrolysis.

The results of Mercier showed that "the aluminium transfer was present at a comparable level in the same high-temperature range", but that epitaxy could be achieved at lower temperatures because the helium did not limit the silane dissociation in the way that the hydrogen did. The result indicates that the silicon reaction is probably the major one in the autodoping process.

From the dependence of the carrier concentration on the growth temperature Dumin (1967) has calculated the heat of reaction of the autodoping process as 835 kJ/mole on (0001) sapphire and 1045 kJ/mole on ($1\overline{1}02$) sapphire. Although he attributed these values to the reaction in equation (2.4.2) it has been shown above that the reaction in equation (2.4.1) is probably dominant in autodoping.

After preparing a film it must undergo a number of high temperature thermal treatments during device processing and these treatments can lead to significant doping changes in the film. If a film is heated in hydrogen, additional aluminium is generated by the reduction of sapphire, (eg a 300 Ω-cm n-type film heated for 5 minutes at 1175°C was converted to 0.6 Ω-cm p-type (Dumin 1967)), and n-type silicon films have shown p-type regions at the top surface and at the sapphire interface after hydrogen heat treatments. This indicates that there is aluminium transfer both via the gas phase and from generation at the sapphire interface due to diffusion of hydrogen through the silicon.

Heat treatments in oxygen can remove aluminium acceptors in two ways; either by the out diffusion of aluminium and its consequent segregation into the oxide layer, or by precipitation within the silicon film onto one of the many crystal faults. Ross and Warfield (1969) have shown that the variation of

acceptor density as a function of oxidation time can be explained entirely by out diffusion and segregation in the oxide if the diffusion constant for aluminium in silicon-on-sapphire is four times the value for bulk silicon. The effect of prolonged oxidation on the carrier mobility is discussed in section 2.6.

Low temperature (450°C) heat treatments in an inert atmosphere after oxidation can form silicon-oxygen complexes which act as donors but these are easily removed by a 30 minute treatment at 1000°C and further heat treatment up to 1250°C has little effect on either the carrier concentration or the mobility, (Dumin 1967).

2.5 Deep Levels and Sapphire Interface States

In the previous section the effect of growth parameters on the density of shallow aluminium acceptors was discussed, but this is not the only way in which the growth conditions effect the carrier concentration. Deep donor and acceptor levels have been identified by Heiman (1967b) and by Dumin (1970). Deep acceptors increase the resistivity of n-type films by trapping the electrons released from shallow donors and deep donors play a similar role in p-type films.

Dumin (1970) investigated the resistivity of both n-type and p-type films grown from silane reservoirs with calibrated impurity contents. The films were grown at a number of different temperatures and their thicknesses ranged from $0.1\mu m$ to $30\mu m$. The resistivity was found to increase as the thickness decreased. Only part of the increase could be accounted for by mobility reductions and the associated decrease in the carrier concentration was believed to be due to compensation by increasing densities of deep donor and acceptor traps associated with crystal defects. Comparisons of the Fermi levels calculated from the resistivity data with the values determined from a graphical solution of the charge neutrality equation enabled estimates of the energy level of the trapping states to be made. The values obtained were 0.25eV from the conduction band edge for the deep acceptor and 0.3eV from the valence band edge for the deep donor. In the thickest films the density of deep levels is of the order of 10^{14} cm^{-3} and in the $0.1\mu m$ films the value may be as large as 10^{17} to 10^{18} cm^{-3}.

Optical absorption and photoconductivity measurements give results which are consistent with the deep donor and acceptor levels predicted by the analysis of the resistivity data. Measurements of the Hall voltage were made in conjunction with the photoconductivity measurements to ensure that majority carriers were responsible for the enhanced conductivity in both the n-type

and p-type films, since Heiman had initially reported only the existence of deep acceptors. A larger photoconductivity response was observed on the lower mobility layers, presumably because of the higher density of crystal defects. The exact physical nature of the deep levels has not yet been established. There is probably more than one type of deep level present since HCl gettering sometimes produces improvements in lifetime whereas in other layers no significant improvements are observed, (Robinson and Heiman 1971). The deep levels reported by Dumin were reduced in density by HCl gettering and had energy levels very similar to those reported for copper, (Schibli and Milnes 1967), but no positive identification was possible. Since the deep levels which were not reduced by HCl gettering were observed in films with appreciable autodoping, it was tentatively postulated by Robinson and Heiman that they may be associated with an aluminium oxygen complex.

Additional less direct evidence for the presence of deep acceptor levels was obtained from the early saturation of the drain characteristics of N-channel MOS transistors. The deep acceptor sites become charged under the positive gate bias leading to a charge density in the depletion region much greater than the acceptor density measured from the Hall measurements and since the knee in the MOS drain characteristic occurs at the voltage $V_k = V_g + (Q_{ss} + Q_s)/C_i$ where Q_s is the charge per unit area in the depletion region the enhanced depletion layer charge accounts for the early saturation of the characteristic.

An electronic layer can be generated at the sapphire interface in some films as a result of certain post-deposition thermal treatments. In 'contaminated' films a 15 minute heat treatment in hydrogen or moisture at 500-1000°C will generate a layer with electron densities in the range $10^{11} - 10^{13}$ cm^{-2}. Heiman (1967) has proposed the following mechanism for the generation of this layer: unintentional impurities (eg sodium) deposited during device processing can diffuse rapidly through the silicon layer and are incorporated

as positively ionised impurities in a highly disordered 'glassy' layer
adjacent to the sapphire. The precise role of the hydrogen or moisture
is not specified but its presence may be necessary to remove the anion
of a sodium salt. Heiman's hypothesis of a 'glassy' layer is not an
essential part of this mechanism since the sodium ions could simply be
trapped at defects in the sapphire-silicon interface. Support for the
hypothesis was obtained by the observation of interface electronic layers
after hydrogen heat treatment of films intentionally contaminated with
potassium.

2.6   Mobility

The two parameters that are of primary importance when specifying
the electrical properties of a semiconducting film are the majority
carrier mobility and carrier concentration. Both these parameters
are closely dependent on the deposition rate, the deposition
temperature, the film thickness and any post-deposition heat treatment
which the film may have received. Therefore for any set of
measurements to be really meaningful all these conditions must be
stated. A number of authors have measured the majority carrier
mobility as a function of the carrier concentration while holding all
the other conditions constant. Their results for both p-type and
n-type films of (111) and (100) orientations are shown in Figures
2.6.1 to 2.6.4. Results for silicon-on-spinel are also included in
these Figures because they provide additional insight into the
significance of various scattering mechanisms in thin silicon films.
The results are 'mean' values for the total thickness of the films.
Since crystalline defects grow out of the film as the growth proceeds
and there is autodoping from the substrates, marked variations in
both the mobility and carrier concentration with position in the film
are to be expected.

With compressive stress of the order of 10^9-10^{10} dynes/cm^2 pronounced
piezo-resistance effects are to be expected for hole conduction in (111)
films and for electron conduction in (100) films, (Schlotterer 1968).
Calculated mobilities for bulk silicon with a comprehensive stress of
8×10^9 dynes/cm^2 are shown in the mobility versus carrier concentration
Figures for p-type (111) films and n-type (100) films. The p-type
mobility is enhanced by up to 65 per cent and the n-type mobility
reduced by up to 57 per cent. Since these mobilities were calculated

35
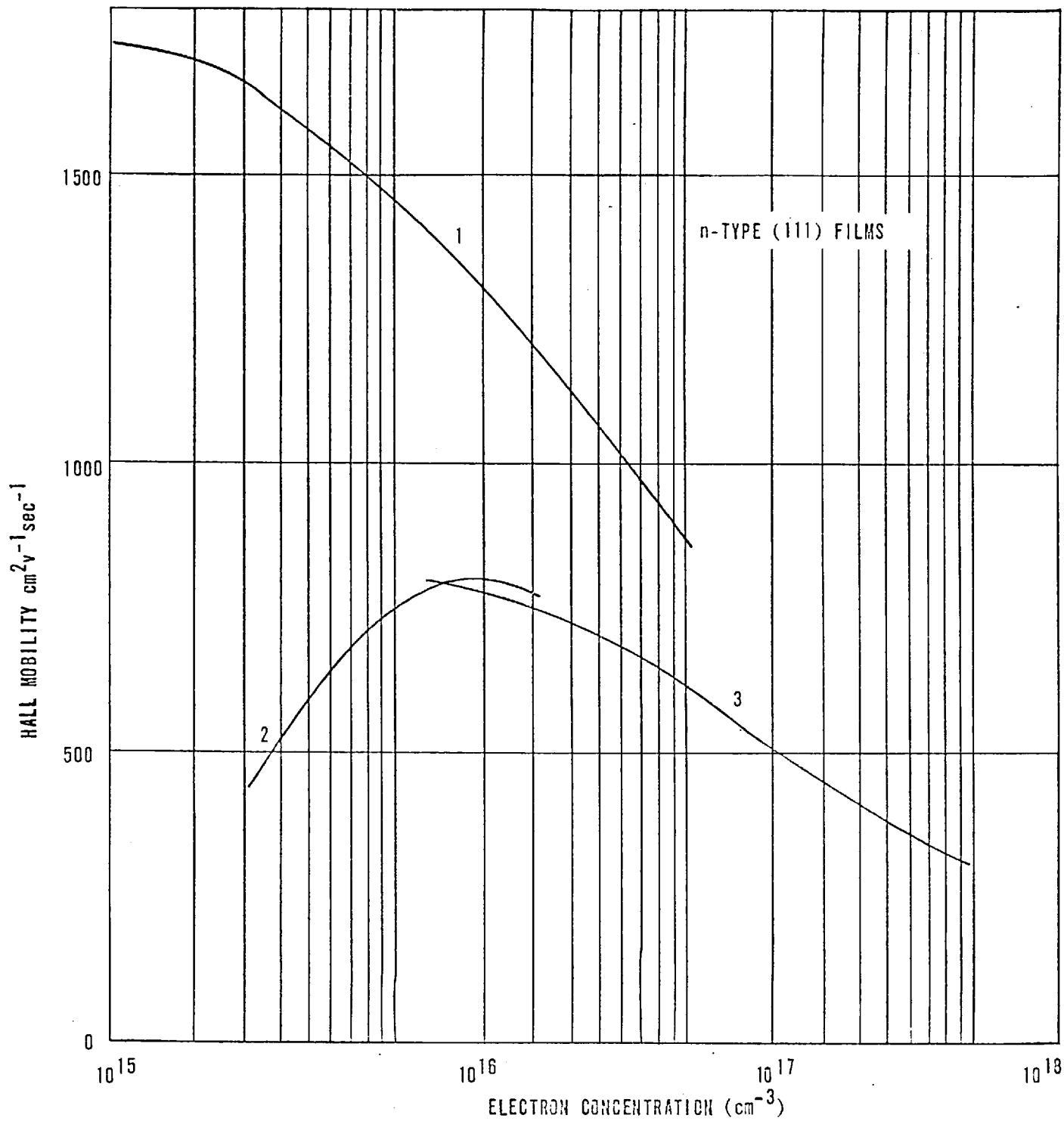
HALL MOBILITY AS A FUNCTION OF ELECTRON CONCENTRATION FOR n-TYPE
FILMS OF (111) ORIENTATION.

1. AVERAGE EXPERIMENTAL BULK SILICON MOBILITY AS REPORTED BY RUNYAN (1965)
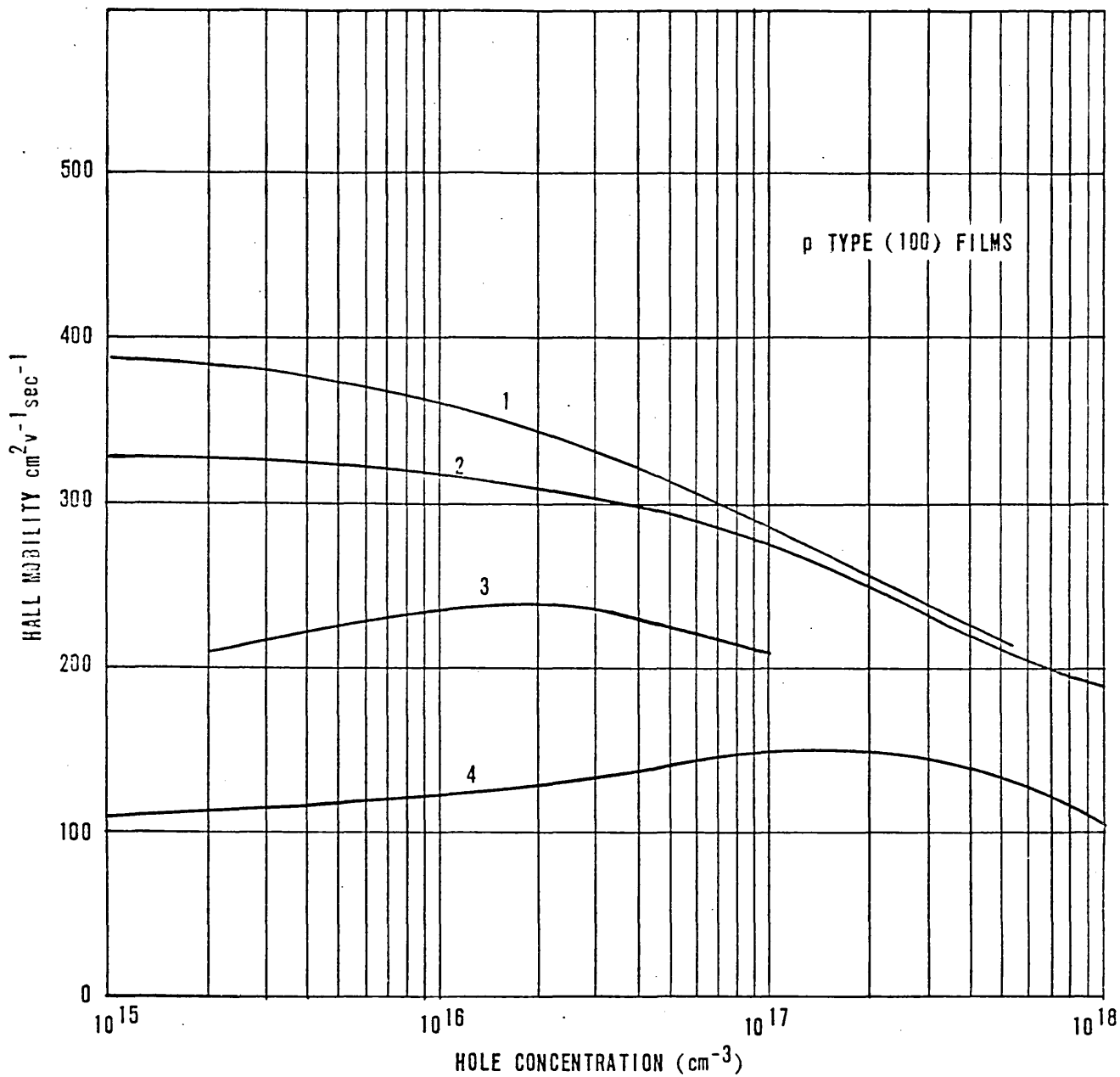
2. EXPERIMENTAL VALUES FOR $1 \cdot 7 \mu m$ THICK FILMS OF SILICON ON SPINEL,
   (GOTTLIEB ET AL 1970)

3. EXPERIMENTAL VALUES FOR $9 \mu m$ THICK FILMS OF SILICON ON SPINEL,
   (TION ET AL 1969)

FIG. 2.6.1.
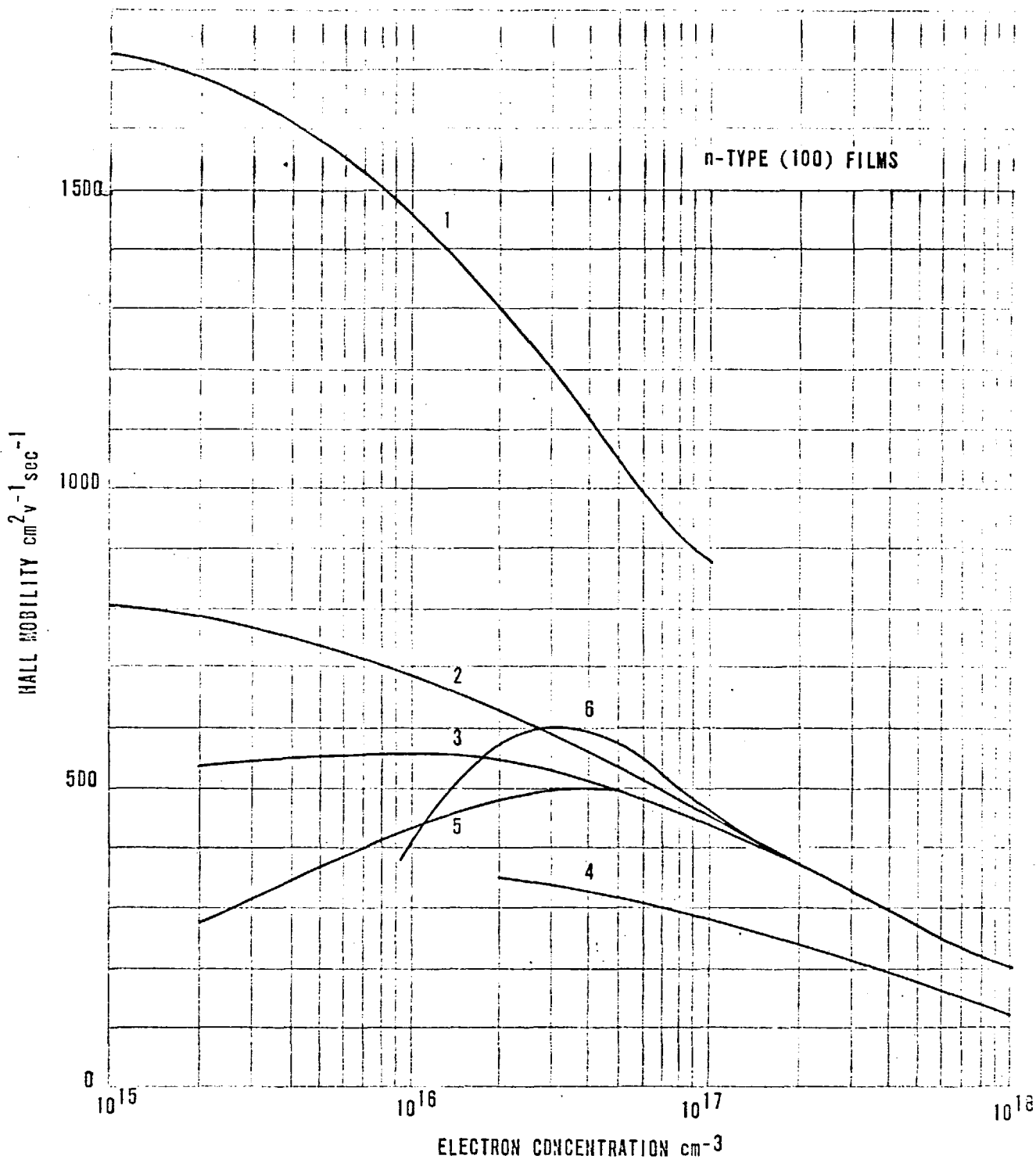
HALL MOBILITY AS A FUNCTION OF ELECTRON CONCENTRATION FOR
n-TYPE FILMS OF (100) ORIENTATION

1. AVERAGE EXPERIMENTAL BULK SILICON MOBILITY AS REPORTED BY RUNYAN (1965)

.2. THEORETICAL VALUES FOR BULK MATERIAL WITH A STRESS OF 8×10^9 DYNES cm^2
     (SCHLOTTERER 1968)

3. EXPERIMENTAL VALUES FOR 2μm THICK SILICON FILMS ON SAPPHIRE AFTER
    THERMAL OXIDATION (DUMIN AND ROBINSON 1968)

4. EXPERIMENTAL VALUES FOR 2μm THICK SILICON FILMS ON SAPPHIRE BEFORE
    THERMAL OXIDATION. (DUMIN AND ROBINSON 1968)

5. EXPERIMENTAL VALUES FOR 20μm THICK SILICON FILMS ON SPINEL,
    (SCHLOTTERER 1968)

6. EXPERIMENTAL VALUES FOR 1 7μm THICK SILICON FILMS ON SPINEL,
    (GOTTLIEB ET AL 1970)

FIG. 2.6.2.

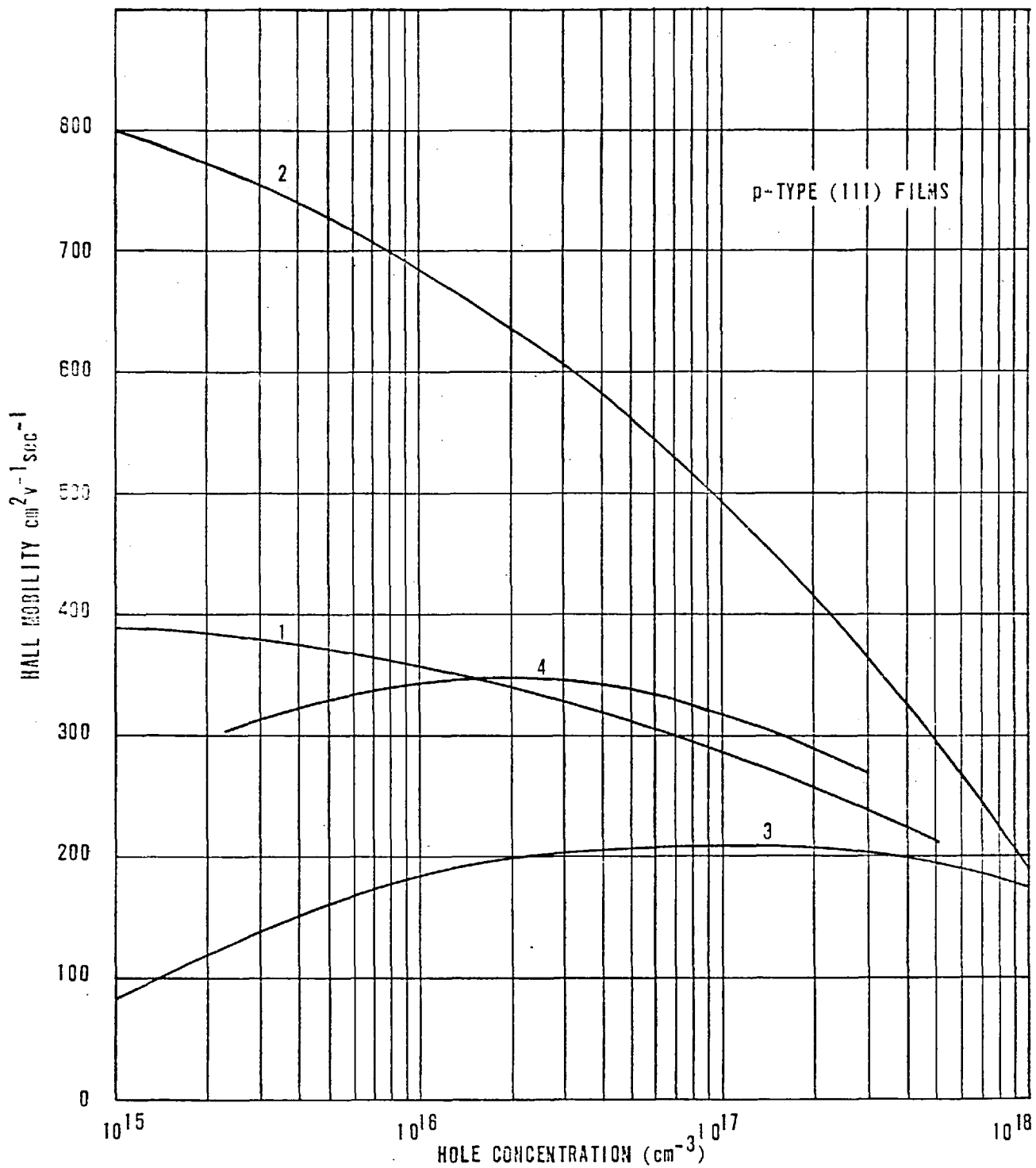HALL MOBILITY AS A FUNCTION OF HOLE CONCENTRATION FOR p-TYPE
FILMS OF (100) ORIENTATION.

1. AVERAGE EXPERIMENTAL BULK SILICON MOBILITY AS REPORTED BY RUNYAN.(1965)

2. EXPERIMENTAL VALUES FOR $6\mu m$ THICK SILICON FILMS ON SAPPHIRE,
   (ALLISON ET. AL. 1969)

3. EXPERIMENTAL VALUES FOR $1\cdot 7\mu m$ THICK SILICON FILMS ON SPINEL,
   (CULLEN ET. AL. 1970)

4. EXPERIMENTAL VALUES FOR $2\mu m$ THICK SILICON FILMS ON SAPPHIRE,
   (DUMIN AND ROBINSON 1968)

FIG. 2.6.3.

HALL MOBILITY AS A FUNCTION OF HOLE CONCENTRATION FOR p-TYPE
FILMS OF (111) ORIENTATION

1. AVERAGE EXPERIMENTAL BULK SILICON VALUES AS REPORTED BY RUNYAN.(1965)

2. THEORETICAL VALUES FOR BULK MATERIAL WITH A COMPRESSIVE STRESS OF
   8×10^9 dynes/cm^2. (SCHLOTTERER 1968)

3. EXPERIMENTAL VALUES FOR 20μm THICK SILICON FILMS ON SPINEL.
   (SCHLOTTERER 1968)

4 EXPERIMENTAL VALUES FOR 1·5μm THICK SILICON ON SPINEL,(CULLEN ET AL 1969

FIG. 2.6.4.

39

from the resistivity/impurity concentration curves of Irvin (1962), they are conductivity mobilities not Hall mobilities as suggested by Schlotterer (1968). The experimental results indicate that the piezoresistance effect is significant. The best (111) p-type results lie between the theoretical strained and unstrained values at the higher carrier concentrations, (ie $> 10^{16}$ cm^{-3}) and the (100) n-type results lie close to the theoretical strained curve above the mid 10^{16} carrier concentrations.

All the Figures 2.6.1 to 2.6.4 show one marked difference in behaviour between bulk silicon, (whether strained or unstrained), and the silicon films. Namely the reduction of mobility with reducing carrier concentration at the lower doping levels. Therefore there must be at least one other carrier scattering mechanism involved in addition to the charged impurity and phonon scattering mechanisms which occur in bulk silicon. A similar phenomena in III-V and II-VI semiconductors has been analysed by Weisberg (1962) in terms of scattering by large space charge regions surrounding local inhomogeneities. Hasegawa et al (1969) have explained the variation of mobility with carrier concentration in vacuum evaporated n-type (111) films, (curve 3 of Figure 2.6.1) by using a combination of three scattering terms:-

$$\frac{1}{\mu} = \frac{1}{\mu_B} + \frac{1}{\mu_D} + \frac{1}{\mu_{SC}} \quad \dots (2.6.1)$$

where μ_B is the bulk mobility, μ_D is the mobility resulting from dislocation scattering and μ_{SC} is the mobility resulting from space-charge scattering. Since these are n-type (111) films the piezoresistance effect should be small. The dislocation scattering term which is independent of carrier concentration and is normally insignificant at room temperature is given by

40

$$\mu_D = 1.0 \times 10^{10} \ \frac{(T)}{(N_{Di})} \ cm^2 V^{-1} \ s^{-1} \quad \text{for electrons} \ \ldots \ (2.6.2a)$$

$$\text{and} \ \mu_D = 6.3 \times 10^9 \ \frac{(T)}{(N_{Di})} \ cm^2 V^{-1} \ s^{-1} \quad \text{for holes} \ \ldots \ (2.6.2b)$$

where T is the absolute temperature and N_{Di} is the dislocation density, (cm^{-2}).

In the absence of any deionisation the space-charge mobility, μ_{SC}, is given by

$$\mu_{SC} = C \ N_D^{\frac{1}{3}} \ T^{-\frac{1}{2}} \quad cm^2 V^{-1} s^{-1} \quad \ldots \ (2.6.3)$$

where C is a fixed parameter for a given film and N_D is the donor density. Hence at sufficiently low carrier concentrations the $1/\mu_{SC}$ term in equation (2.6.1) will dominate and the mobility will decrease as the carrier concentration is reduced.

The highest hole and electron mobilities in the (111) films are greater than those in the (100) films. This result is to be expected as a consequence of the piezoresistance effect which increases the hole mobility in (111) films and reduces the electron mobility in (100) films. The measurement of hole mobilities in the (111) films which are greater than the bulk values is additional confirmation of the existence of a piezoresistance effect.

In general the mobility increases with film thickness up to at least the 5 or 6 μm value, (see for example curves 2 and 4 in Figure 2.6.3). One set of measurements by Cullen et al (1969) on p-type (111) silicon-on-spinel films reached a plateau level from 3 μm. Since dislocation scattering is insignificant at room temperature, the improvement in mobility with film thickness must be due to a reduction in the space-

charge scattering. This explanation is supported by the fact that the 2 μm film in Figure 2.6.3 shows a mobility maximum at 10^{17} cm^{-3} whereas down to 10^{15} cm^{-3} there is no mobility maximum for the 6 μm film.

It is apparent that the density of space-charge regions is greater near the sapphire interface but it is not clear whether they are a direct result of crystal defects (eg low angle grain boundaries), or whether they result from precipitates formed at defect sites. A study of the effect of post-deposition heat treatments may be expected to aid identification of the type of space-charge defects.

Dumin and Robinson (1968) have shown that the effect of thermal oxidation on silicon-on sapphire films is to increase the electron mobility at a given carrier concentration whereas no significant changes in the hole mobility are observed. The thermal oxidation removes aluminium from the substitutional state both by diffusion into the oxide and by causing it to precipitate on crystal defects. Therefore in the pre-oxidation state the n-type silicon is partially compensated and the number of charged scattering centres is greater than the number of free carriers. Part of the post-oxidation increase in mobility is due to a reduction in the density of charged scattering centres at a given carrier concentration. This effect alone is not sufficient to account for the fairly large shifts shown in Figure 2.6.5. The most probable explanation is that some of the defects which give rise to space-charge scattering are being annealed out. In Figure 2.6.6 the type of inhomogeneities suggested as scattering centres by Weisberg are reproduced and clearly the removal of the electron barriers present in types (a), (b) or (e) could give rise to an improved electron mobility without substantially effecting the hole mobility.
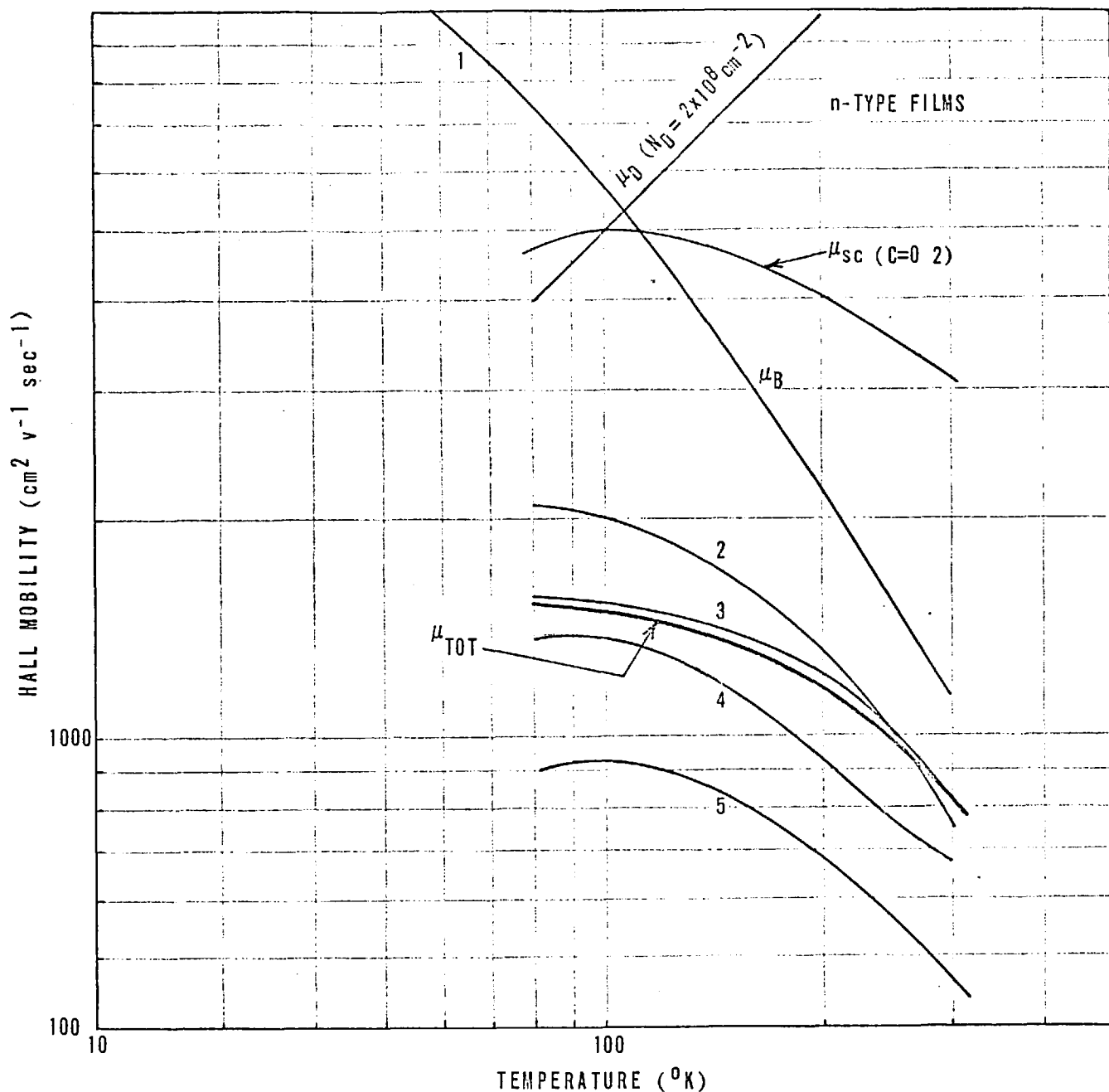
ELECTRON MOBILITY AS A FUNCTION OF MEASUREMENT TEMPERATURE

1   BULK SILICON WITH AN ARSENIC DONOR CONCENTRATION OF 1.8×10^{16} cm^{-3}.
    (MORIN AND MAITA 1954)

2   A 1.7μm THICK (100) SILICON FILM ON ($1\bar{1}02$) SAPPHIRE WITH A PHOSPHORUS
    DONOR CONCENTRATION OF 6.2×10^{15} cm^{-3}; (DUMIN AND ROSS 1970)

3   A 9.7μm THICK (111) SILICON FILM ON SPINEL WITH AN ANTIMONY DONOR
    CONCENTRATION OF 2×10^{16} cm^{-3}, (ITOH ET AL 1969)

4   A 9.1μm THICK (111) SILICON FILM ON (0001) SAPPHIRE AFTER OXIDATION
    ANTIMONY DONOR CONCENTRATION IS 3.3×10^{16} cm^{-3} (ITOH ET AL 1968)

5   A 9.1μm THICK (111) SILICON FILM ON (0001) SAPPHIRE BEFORE OXIDATION
    ANTIMONY DONOR CONCENTRATION IS 4.4×10^{16} cm^{-3} (ITOH ET AL 1968b)

FIG. 2.6.5.

43
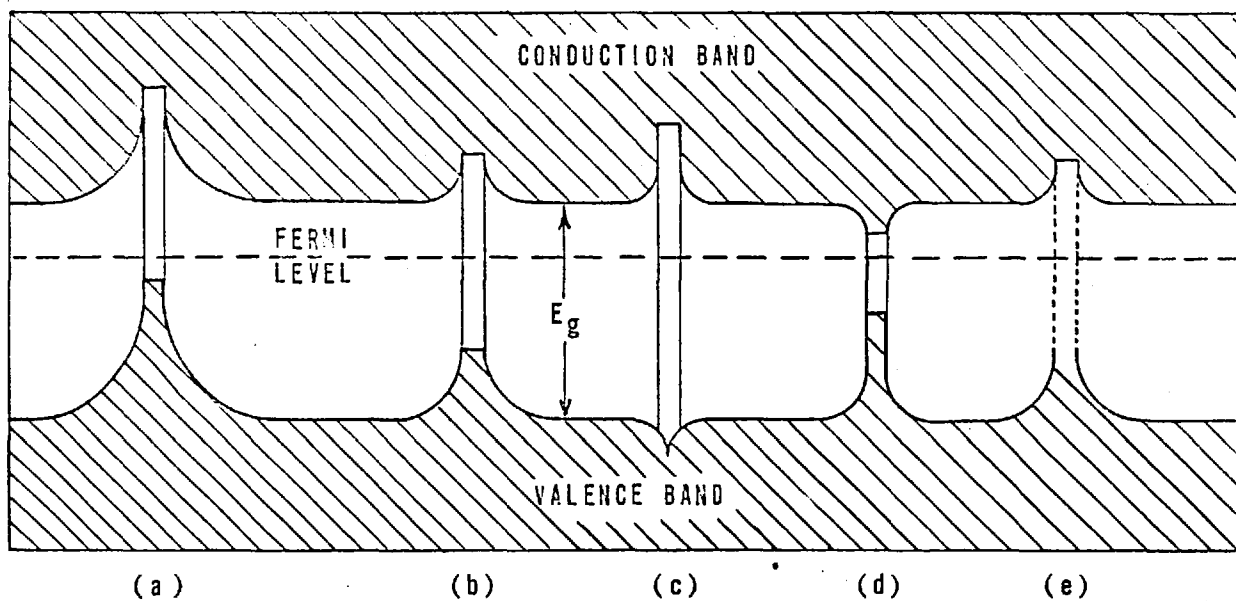
CONTOURS OF CONDUCTION AND VALENCE BANDS IN n-TYPE MATERIAL
SURROUNDING A REGION OF, (a) REVERSED CONDUCTIVITY TYPE, (b)
RESISTIVITY FLUCTUATION , (c) PRECIPITATE WITH LARGER BAND
GAP, (d) PRECIPITATE WITH SMALLER BAND GAP, (e) METALLIC
PRECIPITATE. AFTER WEISBERG (1962)

FIG. 7.6.6.

The effects of oxidation on silicon-on-spinel films as reported by Gottleib et al (1974) are slightly different. The increase in electron mobility is less and could be accounted for by the removal of the compensating aluminium acceptors although an apparent decrease in electron concentration in some of the measurements would appear to invalidate this explanation. The hole mobilities show a small but significant post-oxidation decrease. Therefore additional hole scattering centres must be generated by the oxidation. It is possible that dielectric precipitates of type c. in Figure 2.6.6 are being formed (possibly $Si\ O_2$) (Booker and Tunstall 1966). If this were accompanied by the removal of type b. defects, as appeared to be the case with silicon-on-sapphire then the net effect on the electron mobility could be the small increase observed. The speculative nature of the comments above clearly shows the need for a detailed study to correlate the electrical and structural properties of films before and after thermal oxidation processes. The highest room temperature electron mobility reported is 810 $cm^2 V^{-1} s^{-1}$ at 2×10^{16} cm $^{-3}$ on (111) spinel by Itoh et al (1969) and the highest hole mobility is 360 $cm^2\ V^{-1} s^{-1}$ at 3×10^{16} cm^{-3} also on (111) spinel by Gottlieb et al (1974).

Further information on the predominant carrier scattering mechanisms is obtained from the dependence of the Hall mobility on the measurement temperature and typical results from a number of authors are shown in Figures 2.6.5 and 2.6.7. In both conductivity types the difference between the film mobility and the bulk mobility increases as the temperature decreases and if the film temperature is reduced sufficiently the mobility goes through a maximum and then decreases rapidly. Therefore the additional scattering mechanisms present in the films must
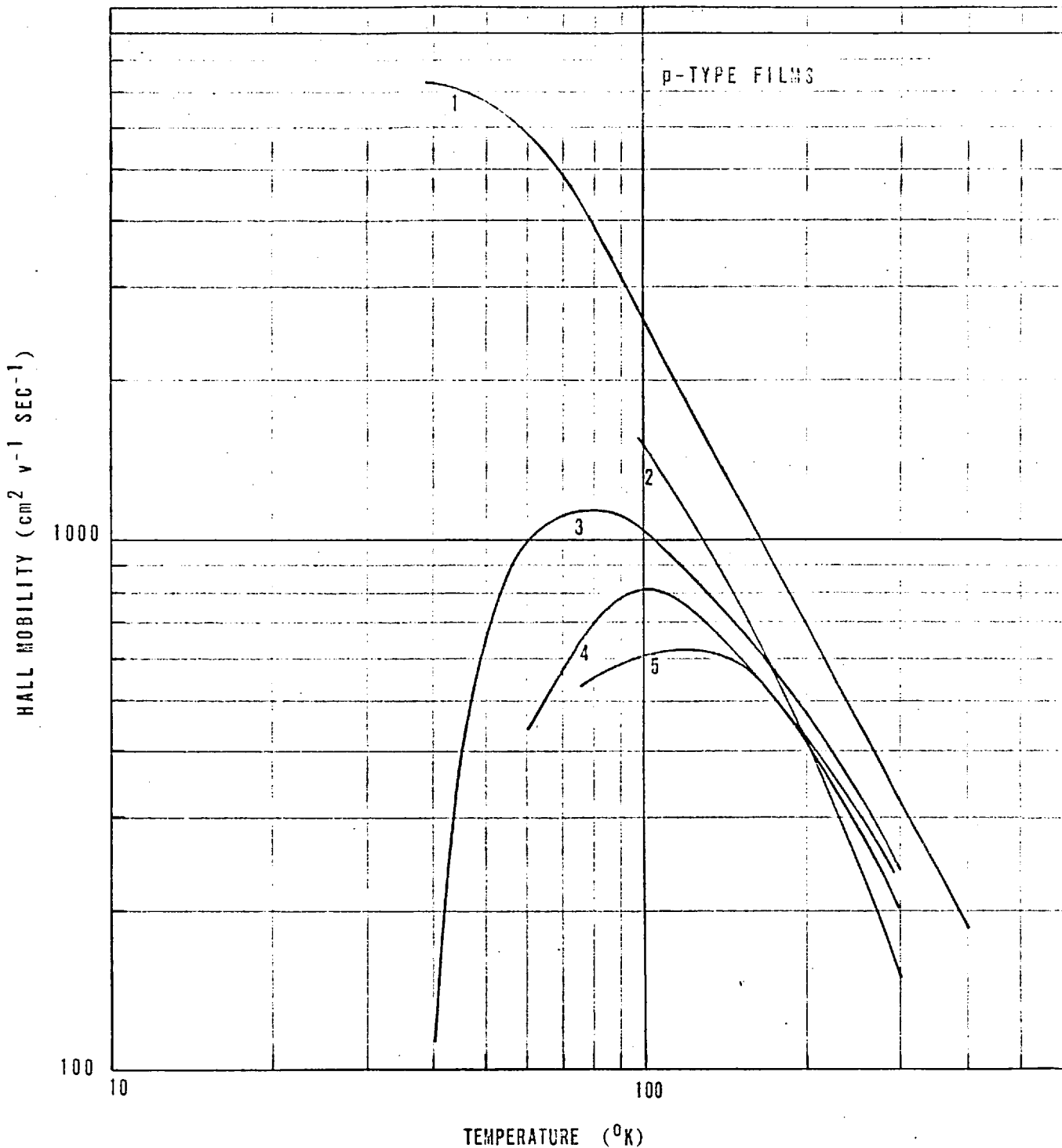
HOLE MOBILITY AS FUNCTION OF MEASUREMENT TEMPERATURE

1   BULK SILICON WITH A BORON ACCEPTOR CONCENTRATION OF $2 \cdot 4 \times 10^{16}$ cm^{-3}
    (MORIN AND MAITA 1954)

2   A $1 \cdot 7 \mu m$ THICK (100) SILICON FILM ON (1T02) SAPPHIRE WITH AN ALUMINIUM
    ACCEPTOR CONCENTRATION OF $2 \cdot 2 \times 10^{16}$ cm^{-3} (DUMIN AND ROSS 1970)

3   A $7 \mu m$ THICK (111) SILICON FILM ON (0001) SAPPHIRE WITH AN ALUMINIUM
    ACCEPTOR CONCENTRATION OF $3 \cdot 4 \times 10^{16}$ cm^{-3} (PRESENT WORK)

4   A $5 \cdot 1 \mu m$ THICK (111) SILICON FILM ON (0001) SAPPHIRE WITH AN ALUMINIUM
    ACCEPTOR CONCENTRATION OF 4×10^{16} cm^{-3} (YASUDA AND OHMURA 1969)

5   A $6 \cdot 4 \mu m$ THICK (111) SILICON FILM ON (0001) SAPPHIRE WITH AN ALUMINIUM
    ACCEPTOR CONCENTRATION OF $2 \cdot 7 \times 10^{16}$ cm^{-3} (ITOH ET AL 1968A)

FIG. 2.6.7.

become increasingly significant at the lower temperatures and their related mobilities must have a positive temperature coefficient. It can be seen from equation (2.6.2) that the mobility due to dislocation scattering has the required properties and although the direct temperature coefficient of the space charge mobility given by equation (2.6.3) is negative this is overwhelmed at lower temperatures by the indirect exponential temperature dependence inherent in the $N^{\frac{1}{3}}$ term. Typical space-charge and dislocation mobilities calculated by Hasegawa et al (1969) are shown in Figure 2.6.5 and these are combined with the bulk mobility to give the theoretical μ_{TOT} plot.

The theoretical plot was in good agreement with experimental results on films on (111) spinel and in general is of the same form as all the plots in Figure 2.6.6. Note that the point at which μ_{SC} goes through a maximum is determined primarily by de-ionisation and therefore is essentially independent of the density of space-charge regions. Therefore the movement of the maximum in the mobility to higher temperatures for plots 3, 4 and 5 in Figure 2.6.5 is presumably due to the differing dislocation densities, (the interception of μ_D with μ_B moves to higher temperatures as N_D increases) An analysis of the results suggest approximate dislocation densities of 1.3×10^8, 2.4×10^8 and 4.2×10^8 cm^{-2} for plots 3, 4 and 5 respectively.

The effect of thermal oxidation on the temperature dependence of mobility in n-type (111) films on (0001) sapphire was investigated by Itoh et al (1968 b.). The films were oxidised in steam at 900°C for 25 minutes and then annealed in argon at 1000°C for 2 hours. The results of measurements before and after this process are shown in curves 5 and 4 of Figures 2.6.5 respectively. The bulk mobility cannot change significantly since the piezoresistance effect is small for this

type of film and the dislocation mobility contribution is insignificant at room temperature. Therefore the observed displacement of the mobility curve must be due to the annealing out of space-charge defects.

It is concluded from the above discussion that any degradation of room temperature mobility below the bulk value is probably due to space-charge scattering and the mobility can sometimes be improved by post-deposition high temperature anneals. At temperature below 100°K dislocation scattering is likely to be the dominant mobility limiting mechanism.

There are two reasons why the field-effect mobility at the oxide interface in MOS devices is likely to be greater than one might expect from the film measurements. Firstly it has been shown that the majority of the space-charge scattering is taking place near the sapphire interface and secondly the strong gate field will lower the potential barriers at the oxide interface due to space-charge regions, (Waxman et al 1965). One factor which may lead to a lowering of the field-effect mobility after a high temperature anneal in an oxidising atmosphere is the formation of silicon dioxide precipitates at points where crystal faults emerge from the silicon.

2.7 Lifetime

The techniques that have been employed to measure minority carrier lifetimes in silicon films on sapphire and spinel are

(a) Transient response of a p-n junction, (Lederhandler and Giacolletto 1955)

(b) Transient response of an MOS capacitor, (Heiman 1967a)

and (c) Analysis of p-n junction I-V characteristics in terms of the SNS Theory, (Sah, Noyce, and Shockley 1957).

When the material being investigated has a short lifetime, method (a) requires complex high-speed measuring equipment whereas the equipment needed for method (b) is much simpler. The rate of signal decay in method (a) is given by

$$\frac{dV}{dt} = \frac{kT}{qT_o}$$

If T_o the mean lifetime, is 10^{-10} sec this gives a rate of change of voltage equal to 2.5×10^8 volts/sec and hence an extremely wideband scope is required as a detector. In Method (b) the time constant of the transient is given by

$$T_t = \frac{2 \, T_o \, N_D}{n_i}$$

where N_D is the doping density. If N_D is say 6×10^{15} cm^{-3} and T_o is again 10^{-10} sec. This gives a time constant of 8×10^{-5} sec which can readily be detected by a standard oscilloscope.

On cooling from growth temperature the thermal expansion mismatch causes the silicon-on-sapphire specimen to deform thereby causing a convex silicon surface to form. Dumin and Silver (1968) have polished the

convex silicon surface on 25 μm thick p and n-type silicon films on (1$\bar{1}$02) sapphire and then fabricated shallow p-n junction diodes in the resultant variable thickness wedge. The forward I-V characteristics of the diodes were analysed according to SNS Theory and the calculated effective minority carrier lifetimes are shown in Figure 2.7.1. The lifetime in the bulk of the film is in the nanosecond range and is an increasing function of distance from the sapphire interface. When the diode is diffused through to the sapphire interface the effective lifetime drops by an order of magnitude due to recombination at the back surface. SNS analysis of reverse-bias diode characteristics in 5μm thick p-type (100) films on (1$\bar{1}$02) sapphire by Hart et al. (1967) gave values of 2-3 nanoseconds for the effective minority carrier lifetime. The results were confirmed by current gain measurements on bipolar transistors and by examining the transient response of a p-n junction. Values in the nanosecond range have also been obtained from reverse-bias diode measurements on (100) and (111) films on commercial M_gO: 3.5 Aℓ_2 O_3 spinel by Robinson & Dumin (1968). But values as high as 40 nanoseconds obtained by method (b) have been reported by Cullen et al (1969) for (111) films on flame-fusion spinel in the composition range M_gO: 1.5 Aℓ_2 O_3 to M_gO: 2.2 Aℓ_2 O_3.

Gettering of metallic impurities in the silicon films by a high temperature anneal in an HCℓ-O_2 ambient has produced significant improvements in minority carrier lifetimes, (Allison et al 1969). Films on both sapphire and spinel substrates have yielded minority carrier lifetimes up to 45 nanoseconds after this process. This lifetime corresponds to diffusion lengths of 6 μm and 12 μm for holes and electrons respectively, at a carrier concentration of 10^{15} cm^{-3}, and these values are sufficiently large to enable bipolar transistors with high current gain values to be made.

MINORITY CARRIER LIFETIME vs. SILICON THICKNESS.
THE p-n JUNCTION WAS 3-4 μ BELOW THE SILICON SURFACE.

AFTER DUMIN AND SILVER (1968)
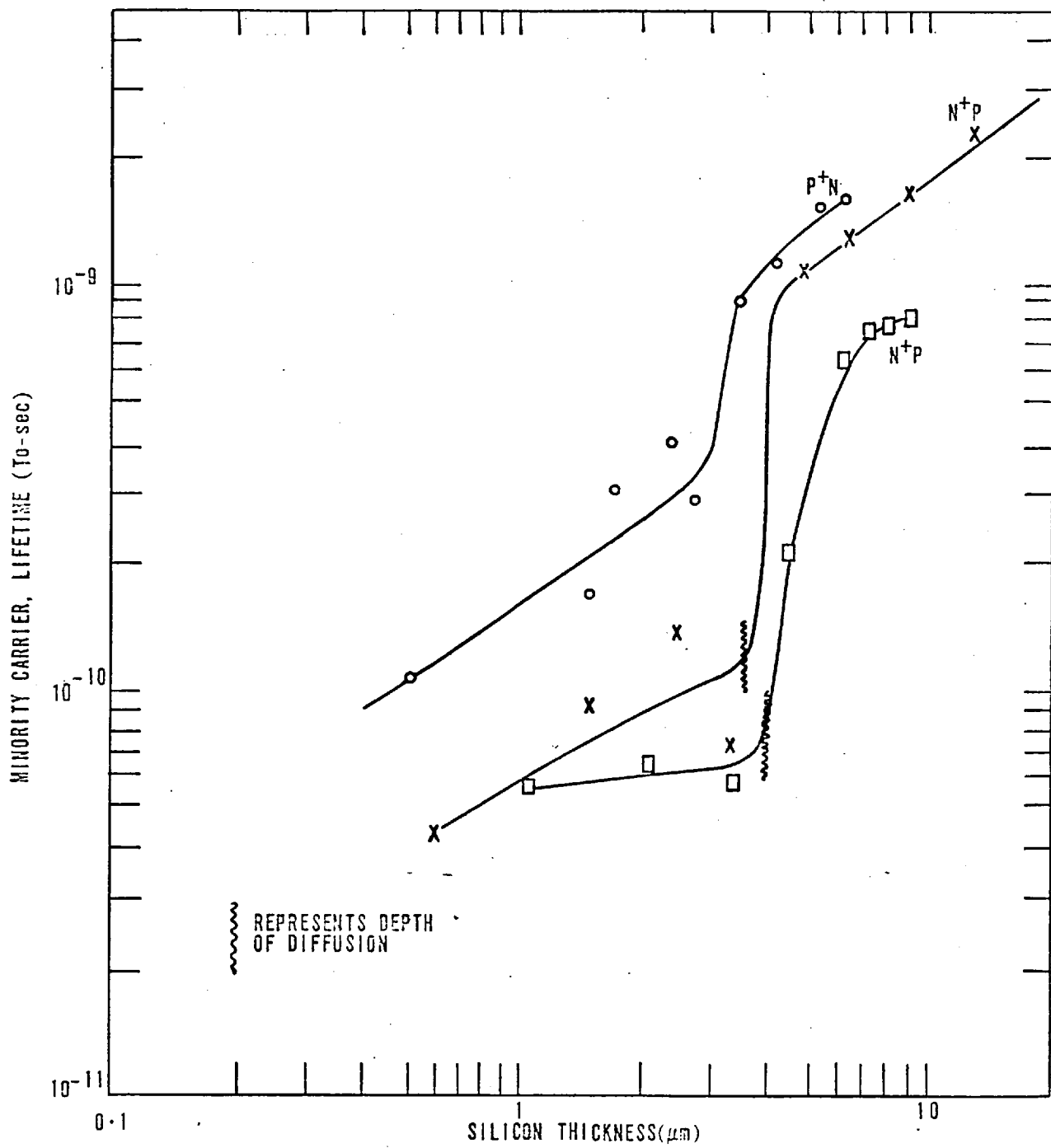
FIG. 2.7.1.

51

CHAPTER 3

THE USE OF MIS DEEP DEPLETION HALL EFFECT STRUCTURES TO STUDY THE CARRIER
TRANSPORT PROPERTIES

A technique has been developed for measuring the variations in carrier
density and mobility as a function of distance from the sapphire
interface in 1-2µm thick silicon-on-sapphire films. The conductivity
and Hall constant are measured in the conducting channel of a deep
depletion MOS transistor as the depletion region is driven into the
film and these results are combined with C-V measurements on an
adjacent capacitor to provide both the carrier concentration and mobility
profiles. Anomalous behaviour due to processing irregularities is
detected and explained.

A new method is described of measuring the energy distribution of the
fast interface states at the oxide interface on layers with non-
uniform doping and the results obtained are compared with those
reported on bulk silicon.

3.1 Experimental Techniques

The structures of the deep depletion MIS Hall effect transistor and the MIS capacitor used in this investigation are shown in Fig 3.1.1, [the two additional deep depletion transistors on the left of Fig 3.1.1 (a) were not used]. Both the channel length and width of the Hall effect transistor are 300μm, as drawn on the photolithographic mask. Sideways diffusion and silicon undercutting will tend to reduce the length and width respectively, but the errors in both will be less than 1.5 per cent. The MIS dielectric is an oxide-nitride double layer with approximately 400Å of oxide and 800Å of nitride. The nitride top layer is used to prevent the migration of ionic contamination into the oxide. Spreads of up to ± 10 percent in the capacitance/unit area of the dielectric are observed but this does not affect the accuracy of the calculations because the measured value of the dielectric capacitance in heavy accumulation is substituted in the analysis. The contacting fingers, numbered 1, 2 and 3 in Fig 3.1.1 (a) are diffused into the edge of the channel under the gate electrode to enable four terminal resistivity and Hall measurements to be made. The source, drain and contact regions are all formed by a heavy phosphorus diffusion which is driven right through to the sapphire interface. The metallisation is aluminium. A detailed description of the transistor fabrication process is given in Appendix 1.

In a deep depletion MIS device (Heiman 1966), (in contrast to the more usual surface inversion layer device), the source and drain are in ohmic contact with the bulk of the film and the conduction channel exists in the bulk of the film. This conduction channel is restricted by applying a potential to the gate of the device such as to drive a depletion region into the film, (e.g. a negative potential on an n-channel device). If the maximum depletion
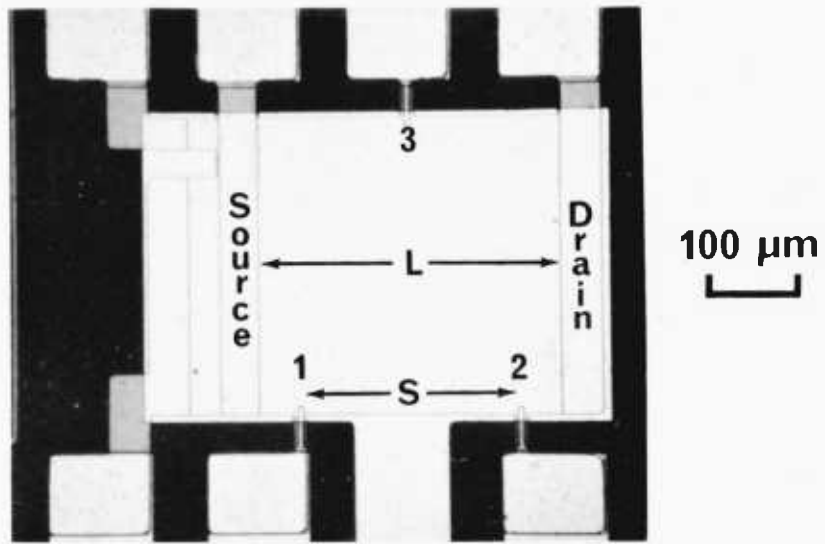
(a) Plan view of Hall effect device
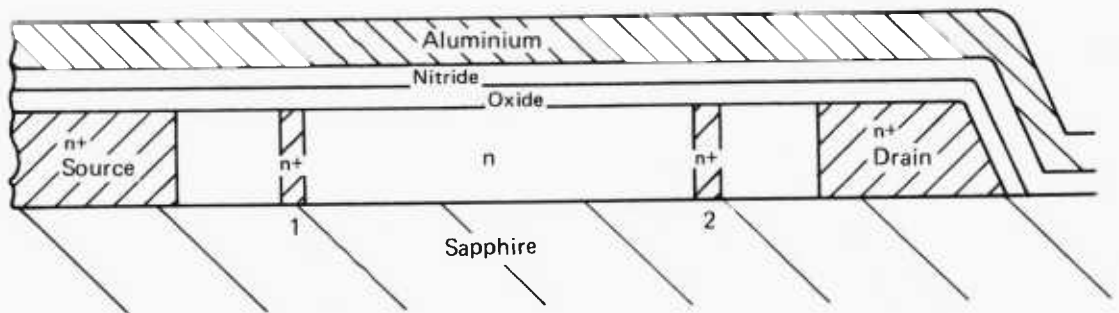
(b) Cross section of Hall effect device
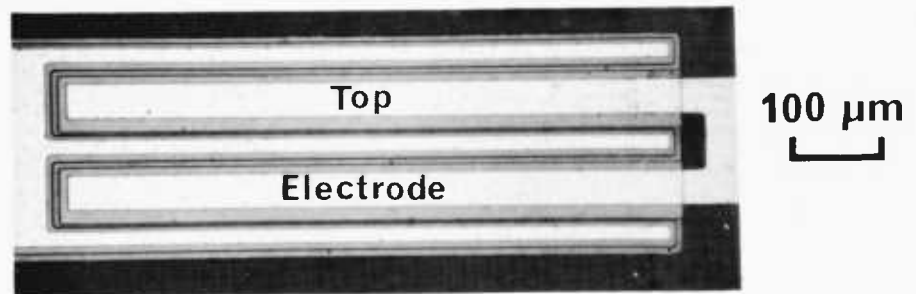
(c) Plan view of capacitor

Fig. 3.1.1. Structure of the deep depletion MIS Hall effect transistor and the MIS capacitor.
    (a) Plan view of Hall effect transistor.
    (b) Cross-section of Hall effect transistor.
    (c) Plan view of capacitor.
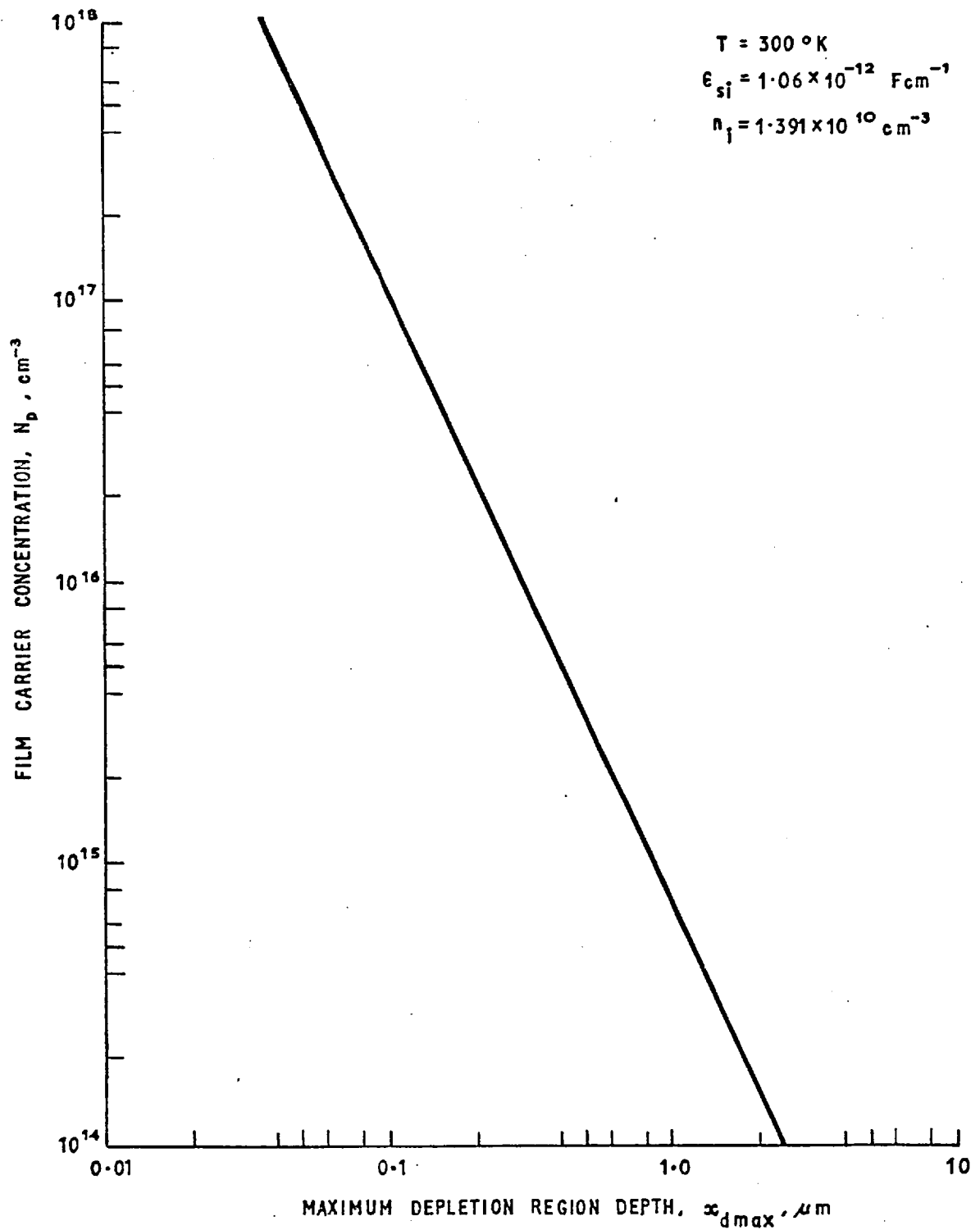
FIG 3. I. 2.   MAXIMUM DEPLETION REGION DEPTH AS A FUNCTION OF FILM CARRIER CONCENTRATION

depth at heavy inversion for the particular film carrier concentration is greater than the film thickness then the device can be completely turned off, and a plot of the maximum depletion depth as a function of the film carrier concentration is given in Figure 3.1.2.

The capacitor has the interdigitated structure shown in Fig 3.1.1 (c) in order to minimise the potential drop in the silicon electrode. When the depletion region is driven back into the silicon electrode of a conventional dot capacitor structure the spreading resistance of this electrode is usually sufficiently large to introduce significant errors in the capacitance measured at 1 MHz. The effect of the distributed resistance in the interdigitated capacitor in Fig 3.1.1 (c) was calculated by treating the structure as a simple R-C transmission line. The analysis showed that with the dielectric thickness employed in this investigation the capacitance error is less than 2.5 per cent provided that the sheet resistance of the conducting layer below the depletion region is less than $10^5 \Omega$/square, which corresponds to a residual layer thickness of $0.1\mu m$ at a film resistivity of $1 \Omega cm$. The area of the interdigitated capacitor, as defined on the photolithographic mask, is $4.912 \times 10^{-4} cm^2$ but since the areas of processed capacitors were up to 7.5 per cent less than this value the dimensions of all the completed specimens used in this study were measured to an accuracy of better than 2 per cent. The silicon films were (100) orientation grown on ($1\bar{1}02$) sapphire substrates by thermal decomposition of silane*. All the films were phosphorus doped with net donor concentrations in the range $3 \times 10^{15} cm^{-3}$ to $3 \times 10^{16} cm^{-3}$. The film thicknesses after processing were in the range $0.6-1.7\mu m$, as measured on a Talysurf to an accuracy of $\pm 0.05\mu m$.

*The films were obtained from the Allen Clark Research Centre of the Plessey Company.

The carrier transport properties of the films were derived from the following measurements made over a range of gate voltages from heavy inversion to heavy accumulation. Four-terminal resistivity measurements were made by using the source and drain for current contacts and points 1 and 2 in Fig 3.1.1(a) for voltage contacts. The potential difference between points 1 and 2 was always within 5 per cent of the value expected from the outer terminal voltage, assuming zero contact resistance at the source and drain. Evidence that the small differences that did exist were probably due to small non-uniformities in film resistivity rather than contact resistance was obtained from the fact that the difference always decreased when the device was driven into heavy accumulation where the fractional contact resistance error would be expected to be at its greatest. The Hall voltages were obtained by using a digital voltmeter with an accuracy of $\pm$ 10μV to measure the change in voltage between the source and contact 3 when the magnetic field was reversed. Typical Hall voltages were in the range 0.25-3.0mV. The linearity of the Hall voltage with magnetic field was checked up to 6000G and all subsequent measurements were made at 4460G.

Referring to Fig 3.1.1(a), the sheet conductivity of the device channel is given by

$$G_s = \frac{S}{L} \cdot \frac{I_c}{V_{12}} \quad \Omega^{-1} \qquad \qquad ...(3.1.1)$$

where I_c is the channel current and V_{12} is the potential difference between the voltage contacts. The Hall coefficient is given by

$$R = 10^8 V_H \frac{d}{I_c B} \cdot \frac{1}{\alpha_m} \quad cm^3 C^{-1} \qquad \qquad ...(3.1.2)$$

where d = thickness of conducting channel, V_H is the Hall voltage and

α_m is a magnetoconductive correction factor for the shorting of the Hall field by the source and drain electrodes (Isenberg, Russell and Greene 1948). For an infinitely long specimen $\alpha_m = 1$ and for the square geometry used in this work $\alpha_m = 0.68$. Equations (3.1.1) and (3.1.2) can be combined by using the relationships

$$R = \frac{1}{nq} \qquad\qquad ...(3.1.3)$$

and

$$\sigma = nq\langle \mu_H \rangle = \frac{G_s}{d} \qquad\qquad ...(3.1.4)$$

to give

$$\langle \mu_H \rangle = \frac{S}{L} \cdot \frac{V_H}{V_{12}} \cdot \frac{10^8}{B\alpha_m} \quad cm^2V^{-1}sec^{-1} \qquad\qquad ...(3.1.5)$$

and

$$\langle n_\square \rangle = nd = \frac{\alpha_m I_c B}{qV_H} \times 10^{-8} \; cm^{-2} \qquad (3.1.6)$$

where $\langle \mu_H \rangle$ = 'mean mobility' in the channel and $\langle n_\square \rangle$ = 'effective number of carriers per cm^2' in the channel. Note that in equations (3.1.3) and (3.1.4) it has been assumed that the conductivity mobility equals the Hall mobility.

In order to estimate the uncertainties in $\langle \mu_H \rangle$ and $\langle n_\square \rangle$ consider first the uncertainties in the individual parameters which appear in equations (3.1.5) and (3.1.6). Maximum errors in I_c and V_H are less than 2 and 4 per cent respectively and the uncertainty in the S/L ratio is less than 1 per cent. The maximum error in α_m due to the 1.5 per cent uncertainty in the square geometry is less than 1 per cent. The uncertainties in V_{12} and B are insignificant, (<0.3 per cent). Using the above uncertainties, examination of equations (3.1.5) and (3.1.6) shows that the maximum errors in $\langle \mu_H \rangle$ and $\langle n_\square \rangle$ are both less than 7.5 per cent and that the major part of this is due to the uncertainty in V_H. The relatively large uncertainties of 4 per cent in V_H only exist when the depletion region is driven deep into the film. For all measurements where $\langle \mu_H \rangle$ is greater than 100 cm^2 V^{-1} sec^{-1} the uncertainty in V_H is less than 1 per cent and the maximum possible errors in $\langle \mu_H \rangle$ and $\langle n_\square \rangle$ less than 4.5 per cent. Non-uniformities in sheet resistivity could give rise to further errors in $\langle \mu_H \rangle$ and $\langle n_\square \rangle$ but the sample used in this investigation is more than an order of magnitude smaller than typical 6 terminal Hall samples and therefore long-range non uniformities are likely to be less significant. The use of four-terminal resistivity measurements to verify the absence of large localised inhomogeneities by checking the constancy of the potential gradient along the sample has already been referred to.

In the next section a method is described of calculating the true average mobility and total carrier density per cm^2 from the values of the 'mean mobility' and 'effective carrier density' obtained from equations (3.1.5) and (3.1.6).

Hall measurements alone simply enable the mobility and carrier density to be calculated as a function of the gate voltage. C-V measurements are made at a frequency of 1 MHz on the adjacent interdigitated capacitor to determine the relationship between the gate voltage, V_G, and the depletion depth x_d. The C-V measurements on a typical device were checked at a number of points over the full range of bias voltages on a 100 kHz bridge and all the results were within 3 per cent of the 1 MHz values, confirming that the interdigitated capacitor is not significantly affected at 1 MHz by the distributed resistance in the silicon electrode. It is shown in Appendix 2 that in the range where the reduced surface potential $u_s < -3$ for an n-type specimen the value of x_d can be obtained from a high frequency C-V plot by using the equation (A 10), i.e.

$$\frac{1}{C} = \frac{1}{C_i} + \frac{x_d}{\epsilon_{si}}$$

where C and C_i are the capacitance per cm^2 of the MIS capacitor and the dielectric respectively and $\epsilon_{si} = 1.06 \times 10^{-12}$ F cm^{-1} is the dielectric constant for silicon. The uncertainty in x_d is determined mainly by the uncertainty in the area of the capacitor and is therefore of the order of 3 per cent.

The major limitations of this method is that under equilibrium conditions there is a maximum depth for the depletion region, which is given by the equation

$$x_{dmax} = 2 \left(\frac{\varepsilon_{si}}{qN_D} \right)^{1/2} \left\{ \frac{kT}{q} \ln \left(\frac{N_D}{n_i} \right) \right\}^{1/2}$$

$$= 2L_D \left\{ \ln \left(\frac{N_D}{n_i} \right) \right\}^{1/2} \qquad \qquad \dots (3.1.7)$$

where L_D is the extrinsic Debye length. Typical values for x_d, which is plotted in Figure 3.1.2, are approximately $1\mu m$ at $N_D = 10^{15}$ cm^{-3} and $0.1\mu m$ at $N_D = 10^{17}$ cm^{-3}.

All the measurements were made in a conventional long-tailed cryostat with automatic temperature control accurate to $\pm 1^\circ K$. The cryostat was used to enable the temperature dependence of mobility to be determined.

3.2 Mobility and Carrier Concentration

When the measurements described in the previous section were made on
the specimens used in this investigation the majority of the layers
exhibited room temperature variations of 'mean mobility' and 'effective
carrier density' of the form shown in Figure 3.2.1, (exceptions to this
behaviour will be discussed in sections 3.4 and 3.5).

In the silicon layers studied the maximum mobility varied from 275 to
420 cm^2 V^{-1} sec^{-1} with spreads of up to $\pm$ 7 per cent occurring across a
single layer. The value of $<n_{\square}>$ at flatband varied by up to $\pm$ 12
per cent across a layer. In deciding on the magnitude of the source to
drain voltage to be used for these measurements a compromise must be
reached between the conflicting requirements of having a large Hall
voltage to minimise measurement errors and of simultaneously ensuring
a relatively small change in surface potential along the device. A
value of 300 mV was chosen which gives a Hall voltage measurement
accuracy of better than 4 per cent under worst conditions. The effect
of the surface potential variation along the channel is most significant
when the position of the depletion layer edge is changing rapidly with
surface potential. The above drain voltage of 300 mV leads to a
maximum variation in depletion depth along the channel of 500 Å, near
the intrinsic surface condition. At the heavy accumulation and heavy
inversion limits the errors arising from surface potential variations
are small.

If the carrier concentrations and mobility vary in the direction
perpendicular to the surface of the film, then $<\mu_H>$ and $<n_{\square}>$ are not
the true average mobility and total number of carriers per cm^2
respectively. Petritz (1958) has shown that in this case the values
of $<\mu_H>$ and $<n_{\square}>$ are given by
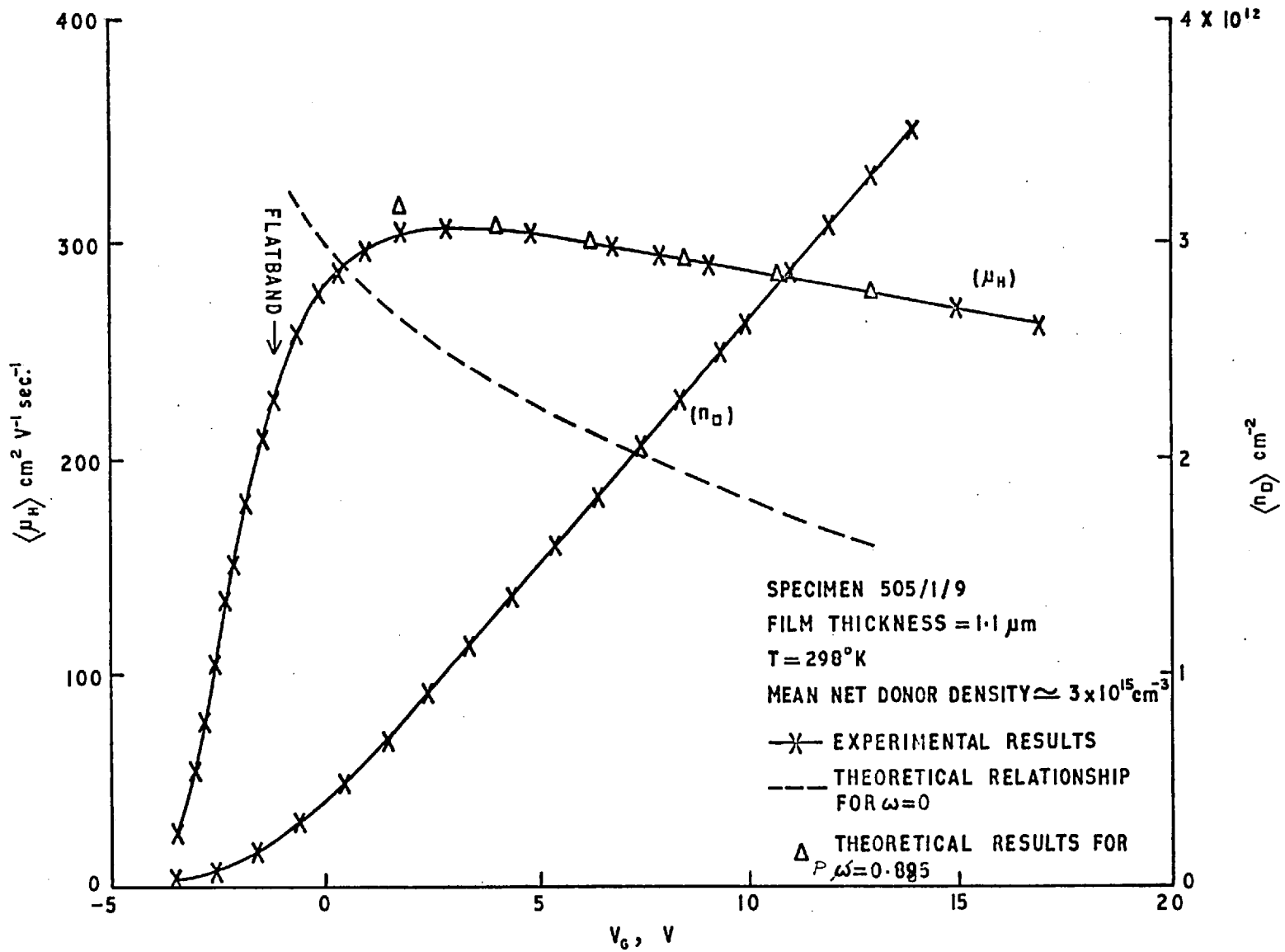
FIG. 3.2.1. 'MEAN MOBILITY' AND 'EFFECTIVE NUMBER OF CARRIERS PER cm^2 OF CHANNEL' AS
A FUNCTION OF GATE VOLTAGE FOR A TYPICAL DEVICE

$$\langle n_\square \rangle = \frac{d}{Rq} = \frac{\left(\int_o^d n(z)\mu(z)\,dz\right)^2}{\int_o^d n(z)\mu^2(z)\,dz} \qquad \ldots 3.2.1$$

and

$$\langle \mu_H \rangle = \frac{\int_o^d n(z)\mu^2(z)\,dz}{\int_o^d n(z)\mu(z)\,dz} \qquad \ldots 3.2.2$$

where $n(z)$ is the carrier concentration per cm^3 at a distance z from the sapphire interface.

It is useful to note that if the mobility is constant but the carrier concentration still retains its spatial dependence then the above equations reduce to

$$\langle \mu_H \rangle = \mu \qquad \ldots 3.2.3$$

and

$$\langle n_\square \rangle = \int_o^d n(z)\,dz \qquad \ldots 3.2.4$$

Information on the carrier transport properties of the silicon films is obtained by using equations 3.2.1 to 3.2.4 to analyse the accumulation and depletion region results in Fig 3.2.1. In the heavy accumulation region the majority of the carrier transport is in a thin region near the oxide interface where it would be reasonable to assume that the mobility is constant with respect to position and therefore equations 3.2.3 and 3.2.4 can be used. Since the gate capacitance is equal to the dielectric capacitance in this region, the total carrier density will vary linearly with gate voltage and its

expected rate of change can be calculated from a direct capacitance
measurement in heavy accumulation after removal from the cryostat,
(the stray parallel capacitance in the three terminal direct measure-
ment is less than 0.1 pF, ie less than 1 per cent of the dielectric
capacitance). A comparison between this calculated rate of change of
total charge density with gate voltage and the rate of change of free
charge density with gate voltage as obtained from Hall measurements
shows that they agree to $\pm$ 5 per cent.

A similar comparison of inversion layer charge densities in bulk
silicon by Fowler, Fang and Hochberg (1964) resulted in an equally
close agreement. Fang and Fowler (1968) concluded from this agreement
that there was little trapping of electrons in the corresponding range
of surface potentials and that r, the ratio of the Hall mobility to
the conductivity mobility is constant, but it would also appear to
signify that r is approximately equal to unity. Since heavily
accumulated surfaces are degenerate the relaxation time is expected to
be constant (Smith 1961), and Zemel (1958) has shown that for a
constant relaxation time the theoretical value of r is 0.87 for diffuse
surface scattering and unity for specular surface scattering.
Because the maximum errors in both charge densities are less than
4.5 per cent and the experimental values of r are within 5 per cent of
unity the difference between this and 0.87 is believed to be
significant and probably indicates a large proportion of specular
scattering at the oxide interface. An estimate of the probability
of specular scattering is obtained from the following analysis of
the mobility variation with gate voltage in the accumulation region.

The mobility in the accumulation region slowly decreases as the gate
voltage increases as shown in Fig 3.2.1. Similar type of behaviour

has been reported for accumulation layers in bulk silicon by Reddi (1968) and also for inversion layers by Murphy, Berz and Flinn (1969) and by Fang and Fowler (1968). The decrease is probably due to enhanced oxide-interface scattering, since at higher gate voltages there will be more interface collisions both because of the large electric field perpendicular to the surface and because a larger fraction of the carriers are within a mean free path of the oxide interface.

The results of Reddi, which were shown to be basically in agreement with a diffuse scattering mechanism at the oxide interface, exhibit a more rapid decrease of mobility with increasing surface field than the present results. In order to test whether the differences could be attributed to a greater probability of specular scattering the present results were analysed on the basis of the approximate equation for partially specular scattering derived by Many, Goldstein and Grover (1965).

$$\frac{\mu_s}{\mu_B} = \frac{1}{1 \pm (1 - p) \left(\frac{r_m \ F_s}{u_s}\right) (1 + u_s)^{1/2}} \qquad \ldots 3.2.5$$

where μ_s and μ_B are surface and bulk mobility respectively, u_s is the surface potential in units of kT, p is the probability that an electron reaching the surface will be specularly reflected, F_s is a function of surface potential and the net donor density, (see Fig 4.7 in Many, Goldstein and Grover 1965), and r_m is the ratio of the bulk mean free path to the extrinsic Debye length, given by the equation

$$r_m = \frac{\lambda}{L_D} = \mu_B \left(\frac{m^* N_D}{2\pi\varepsilon_{si}} \right)^{1/2} \qquad \text{... 3.2.6}$$

where N_D is the net donor density and m^* is the electron effective mass, $(= 0.26\ m_o)$. Equation 3.2.5 has been shown to be a good approximation to a more rigorous numerical analysis both for entirely diffuse scattering (Goldstein, Grover, Many and Greene 1961) and also for partially specular scattering at flatband (Many, Goldstein and Grover 1965).

In order to evaluate equation 3.2.5 values must be determined for μ_B and r_m. Although μ_B is not constant with respect to depth in silicon-on-sapphire films an approximate value appropriate to the material near the oxide interface can be obtained from Fig 3.2.1 in the following manner. At flatband equation 3.2.5 reduces to

$$\frac{\mu_s}{\mu_B} = 1 - (1 - p)\ r_m \qquad \text{... 3.2.7}$$

Since it shall be shown that $(1 - p)r_m < 5 \times 10^{-3}$ the theoretical value of μ_s at flatband is within 0.5 per cent of μ_B. Unfortunately at flatband it is no longer true that the majority of the carrier transport is in a layer close to the oxide interface and in fact it is seen from Fig 3.2.1 that the value of $\langle \mu_H \rangle$ has started to decrease due to significant contributions from the lower mobility regions of the silicon film closer to the sapphire interface. However an approximate

estimate for μ_s at flatband can be obtained by extrapolating the value of $\langle\mu_H\rangle$ in heavy accumulation back to the flatband voltage. The value obtained in this way from Fig 3.2.1 was 326 cm^2 V^{-1} sec^{-1} which is only 25 per cent of the good single-crystal bulk silicon value at the same impurity concentration. Although the net donor concentration also varies as a function of depth in the film, a mean value of 3×10^{15} cm^{-2} obtained from Fig 3.2.2 can be used to calculate a value of 3.4×10^{-2} for r_m from equation 3.2.6. The value of u_s at any given gate voltage is obtained from the corresponding $\langle n_\square \rangle$ value in Fig 3.2.1 by using the graphical relationship between surface charge and surface potential derived by Whelan (1965). Substitution of $p = 0$ and the above values of μ_B and r_m in equation 3.2.5 generates the theoretical mobility-gate voltage relationship for entirely diffuse surface scattering shown by the dashed line in Fig 3.2.1 which lies well below the experimental points and changes much more rapidly with gate voltage. Whereas, if p is taken as 0.885 then the excellent agreement between the experimental results and the theoretical prediction shown in Fig 3.2.1 is obtained. This confirms the high probability of specular scattering suspected from the r value of unity, and this conclusion is compatible with the work of Cheng and Sullivan (1973), who have shown from an analysis of the surface field dependence of the mobility in an MOS inversion layer that any surface asperities must be very small ($\sim$ 1Å or less).

Near the flatband condition not only are n and μ functions of z but also the function n(z) varies with the gate voltage, hence the results are very difficult to analyse in this region. As the gate voltage is reduced and the interface conditions move from accumulation into depletion the slope of the $\langle n_\square \rangle$ vs. V_G plot decreases. A qualitative explanation for this is that for a given change in gate voltage the

mean position of the modulated charge moves back from the oxide interface (this also exhibits itself as a decrease in capacitance) and furthermore because the surface potential is changing rapidly part of the change in surface charge may be a change in trapped charge which does not contribute to $\langle n_\square \rangle$.

Since inversion layer charge does not contribute to the channel current, in the range $u_s < -3$ the charge density in the channel is modulated solely by the movement of the edge of the depletion region, (this is proved in Appendix 2). Hence the functions $n(z)$ and $\mu(z)$ in equations 3.2.1 and 3.2.2 are independent of gate voltage and only the limits of the integrations are changed.

Let $\langle n_{01} \rangle$ signify that the integrations are between the limits 0 and d_1, $\langle n_{02} \rangle$ between the limits 0 and d_2 etc.

From equations 3.2.1 and 3.2.2

$$\langle n_{01} \rangle \langle \mu_{01} \rangle = \int_0^{d_1} n(z)\mu(z)\,dz$$

and

$$\langle n_{01} \rangle \langle \mu_{01} \rangle^2 = \int_0^{d_1} n(z)\mu^2(z)\,dz$$

Therefore

$$\int_{d_2}^{d_1} n(z)\mu(z)\,dz = \langle n_{01} \rangle \langle \mu_{01} \rangle - \langle n_{02} \rangle \langle \mu_{02} \rangle \qquad \ldots 3.2.8$$

and

$$\int_{d_2}^{d_1} n(z)\mu^2(z)\,dz = \langle n_{01} \rangle \langle \mu_{01} \rangle^2 - \langle n_{02} \rangle \langle \mu_{02} \rangle^2 \qquad \ldots 3.2.9$$

In the limit where $d_1 - d_2$ is small equation 3.2.2 gives

$$\mu(\overline{d_{12}}) = \frac{<n_{01}><\mu_{01}>^2 - <n_{02}><\mu_{02}>^2}{<n_{01}><\mu_{01}> - <n_{02}><\mu_{02}>} \qquad \ldots 3.2.10$$

where $\overline{d_{12}}$ signifies the value mid-way between d_1 and d_2

From equation 3.2.1 it also follows that

$$n(\overline{d_{12}}) = \frac{(<n_{01}><\mu_{01}> - <n_{02}><\mu_{02}>)^2}{<n_{01}><\mu_{01}>^2 - <n_{02}><\mu_{02}>^2} \cdot \frac{1}{\Delta x_d} \qquad \ldots 3.2.11$$

where Δx_d = increase in the depletion depth.

Typical carrier concentration and mobility profiles calculated from the results shown in Figure 3.2.1 using equations 3.2.10 and 3.2.11 together with the relationship

$$\frac{1}{C} = \frac{1}{C_i} + \frac{x_d}{\varepsilon_{si}}$$

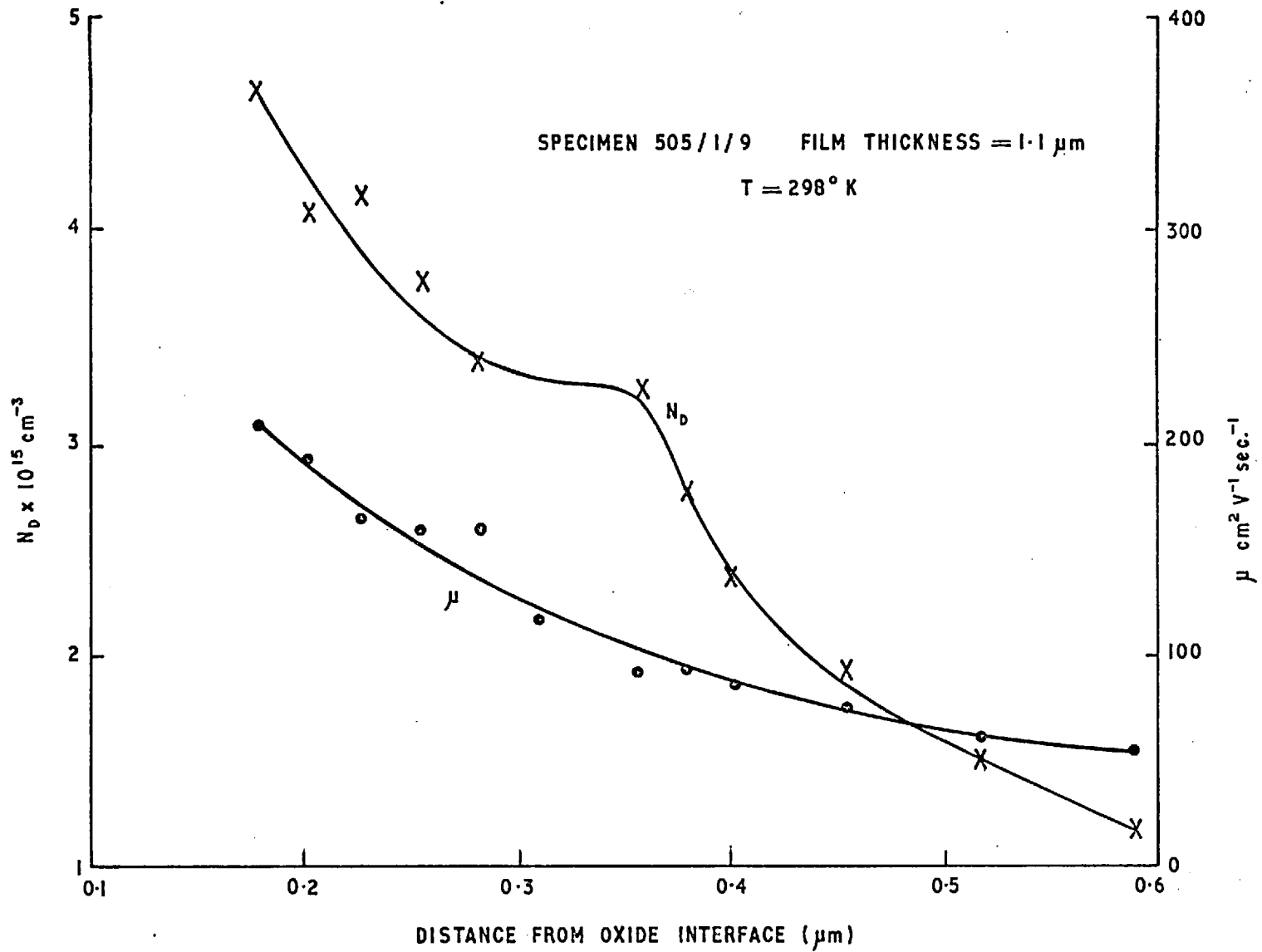as presented in Figure 3.2.2.

FIG. 3.2.2. MOBILITY AND CARRIER CONCENTRATION AS A FUNCTION OF DISTANCE FROM THE OXIDE INTERFACE

Since the difference terms in equations (3.2.10) and (3.2.11) are similar in magnitude to the maximum possible errors in the individual parameters these calculations would be virtually meaningless if the errors were random. But because the errors are largely systematic an a priori estimate of the inaccuracies in $n(\overline{d_{12}})$ and $\mu(\overline{d_{12}})$ is not possible and the validity of the analysis must be judged by comparison with other independent measurements and on the repeatability of the results. Figure 3.2.2 shows a tendency for the net donor concentration to decrease towards the sapphire interface, which is probably due to compensation by deep acceptor levels of the type reported by Heiman (1967b) and Dumin (1970). All the wafers used in this investigation had received post-growth heat treatments of 15.5 hr at 1100-1200°C in a gas stream of 25 per cent oxygen in nitrogen in order to remove aluminium by redistribution into the growing oxide film. According to the analysis of Ross and Warfield (1969) this should reduce the aluminium level below 10^{14} cm^{-3}. The mean doping level for the maximum depletion depth calculated from the C_{max}/C_{min} ratio of the interdigitated capacitor was 3.5 x 10^{15} cm^{-3} which is in good agreement with the results in Fig 3.2.2.

The mobility decreases steadily in the direction of the sapphire interface towards a limiting value of approximately 50 cm^2 V^{-1} s^{-1}. The long high-temperature post-growth heat treatment would be expected to generate aluminium-oxygen complexes, (Ross and Warfield 1969) which may be the scattering centres responsible for this low interface mobility. The shallow SiO_4 + donor centres reported by Ipri and Zemel (1973) are not likely to be responsible for the observed behaviour because they would give rise to an increase in the carrier concentration as the sapphire interface is approached.

3.3 Temperature Dependence of Mobility

The temperature dependence of the Hall mobility in the heavy accumulation region was investigated at a gate voltage of + 10 V and a drain voltage of 300 mV. The 'mean mobility' was measured as the specimens were cooled down from room temperature to below 100°K and as they were allowed to warm back up to room temperature, with the temperature controlled to $\pm 1^{\circ}$K during each measurement. The results for a typical specimen all lie on the curve labelled (A) in Figure 3.3.1.

In the vicinity of 300°K the slope of the curve approaches conformity with a $T^{-1.5}$ law, as has also been reported for electron inversion layers in bulk silicon (Fang and Fowler 1968), which probably indicates a predominance of phonon scattering. This phonon scattering must be a bulk mechanism since Greene (1964) has shown that surface electron-phonon interactions produce no diffusivity when the electron gas is degenerate, as is the case in heavy accumulation. This conclusion is also consistent with the high specular reflection coefficient of 0.885 measured in the previous section. As the temperature is lowered the mobility increases less rapidly than for pure phonon scattering and below 130°K it begins to decrease with decreasing temperature, suggesting a predominance of charged impurity scattering.

This type of behaviour has been observed in bulk silicon by Morin and Maita (1954) and in thin silicon-on-sapphire films by Dumin and Ross (1970) and it simply arises from the different temperature dependence of the two scattering mechanisms. The transition from phonon scattering to impurity scattering occurs at a lower mobility and a higher temperature than one would expect for a carrier concentration of 3×10^{15} cm^{-3}, which may indicate the presence of a much higher density of charged scattering centres due to compensation by

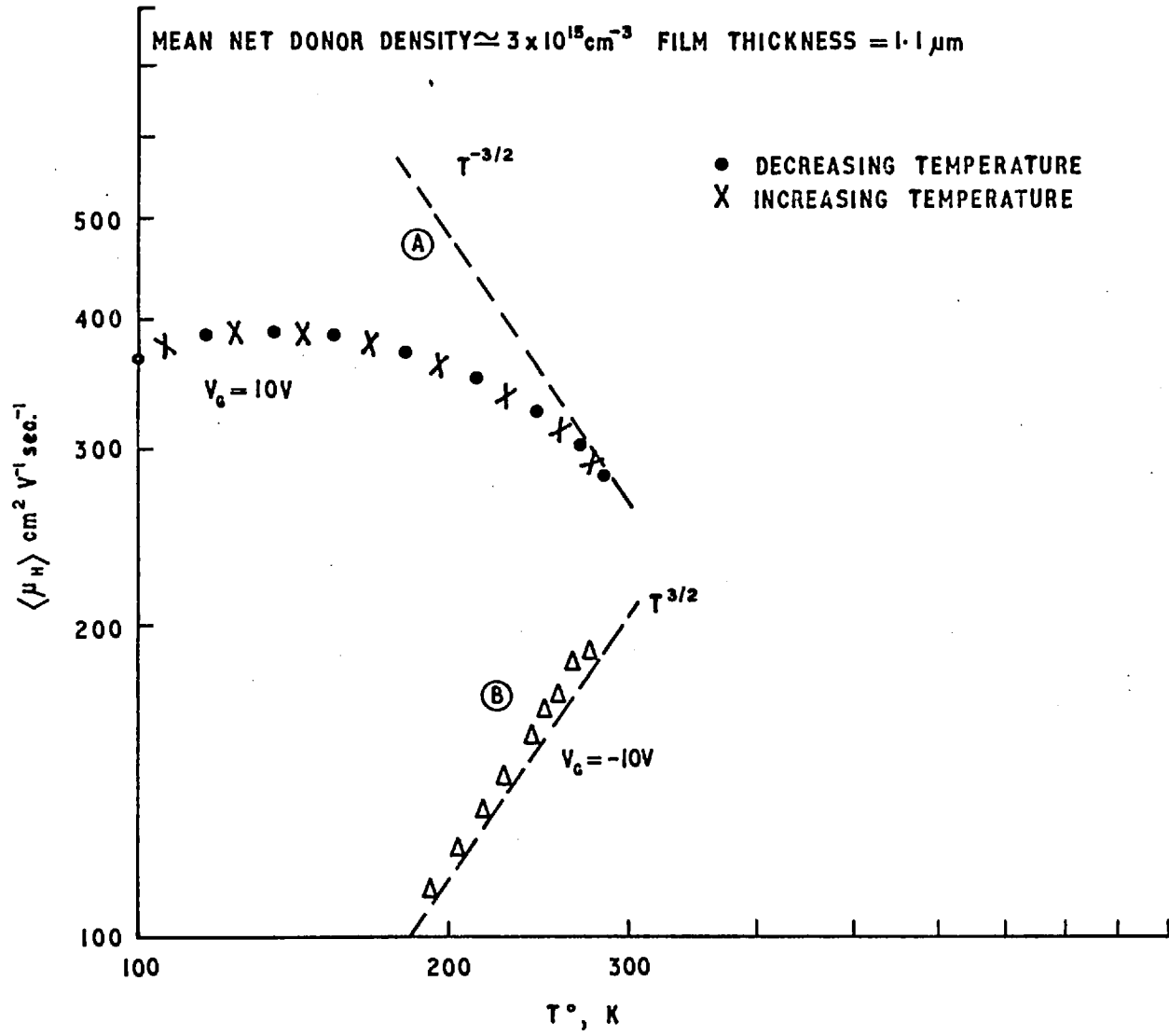MEAN NET DONOR DENSITY $\simeq 3 \times 10^{15} cm^{-3}$ FILM THICKNESS $= 1\cdot 1\,\mu m$

FIG. 3.3.1. TEMPERATURE DEPENDENCE OF MOBILITY

deep acceptors (Dumin and Ross 1970). Additional neutral scattering centres
of the type reported by Ipri and Zemel (1973) would lower the value of the
mobility maximum but since their mobility contribution is temperature
independent they would not influence the temperature at which it occurs.

Another factor in the case of an accumulated surface is the temperature
dependence of the penetration depth of the accumulation layer. At a given
surface potential the penetration depth is proportional to the extrinsic Debye
length, which is in turn proportional to $T^{1/2}$ and inversely proportional to the
carrier density in the bulk to the half power. Below 130°K the carrier density
decreases exponentially with temperature, outweighing the $T^{1/2}$ term and giving
rise to a rapid increase in the penetration depth of the accumulation layer.
This penetration into a more heavily compensated region will also increase the
ratio of charged impurity scattering events to phonon scattering events.

Curve (B) in Figure 3.3.1 shows the temperature dependence of the mobility in
the region of the film near the sapphire interface. This was measured at
$V_G = -10$ volts, where the surface is inverted and conduction is limited to a
$0.5\,\mu$m region adjacent to sapphire. The $T^{1.5}$ dependence of the curve is
evidence of the predominance of charged impurity scattering in this region.
This measurement is complicated by the long time constants required to reach
equilibrium at low temperatures, (Goetzberger 1967), but the problem
can be overcome by heavily inverting the surface at room temperature and taking
all the measurements as the temperature is lowered. Then the attainment
of an equilibrium state is not limited by the speed of the generation
process. The measurements were restricted to the temperature range 200–300°K
because below 200°K there is appreciable deionisation of the donors and the
depletion region would be driven back towards the sapphire interface. Hence the
$\langle \mu_H \rangle$ value would be an 'effective mean' over a different depth of channel. The
$T^{1.5}$ temperature dependence for carrier transport near the sapphire interface

75

indicates that although space charge scattering from neutral aluminium-oxygen complexes (Ross and Warfield 1969) probably exists it is not the major scattering mechanism in this region because a $T^{-0.5}$ dependence is to be expected for space charge scattering (Weisberg 1962) in a temperature range where the carrier concentration is not changing significantly.

3.4 Detection of the Phosphorus Transfer Anomaly

When the first few slices were processed the nitric acid boils specified in steps 10 and 16 of the processing procedure in Appendix 1 were not included. The growth of high temperature oxides on unprotected surfaces resulted in the transfer of phosphorus from the source and drain regions to the channel (Edwards 1969).

The first indication that problems due to phosphorus transfer existed came from the small C_{max}/C_{min} ratio of the capacitor C-V curve, which on a typical example indicated a doping level of 2.5×10^{17} cm^{-3}. Although the calculated total phosphorus diffusion length for all processing subsequent to the transfer is only of the order of 2500Å this is still more than three times the x_{dmax} at 2.5×10^{17} cm^{-3}, (see Figure 3.1.2). Hence this value should be a good estimate of the surface concentration.

Typical plots of the variation of 'mean mobility' and 'effective carrier density per sq cm' with gate voltage for these specimens with anomalous C-V curves are given in Figure 3.4.1. The room temperature 'mean mobility' plot does not have the same marked dependence on negative gate voltages as was observed on the uncontaminated specimens in section 3.2. Presumably this is because most of the carrier transport in the present specimens is in the heavily doped surface skin. The increase in 'mean mobility' with more negative gate voltages in the 110°K characteristic is evidence that the heavily doped region is confined to a surface skin. At high donor concentrations in the 10^{17} cm^{-3} range the degree of ionisation will drop fairly rapidly with decreasing temperature (Morin and Maita 1954). Therefore at a given gate voltage the depletion depth will be greater. As the gate voltage is made more negative the depletion range is driven through the heavily doped surface skin where the electron mobility is
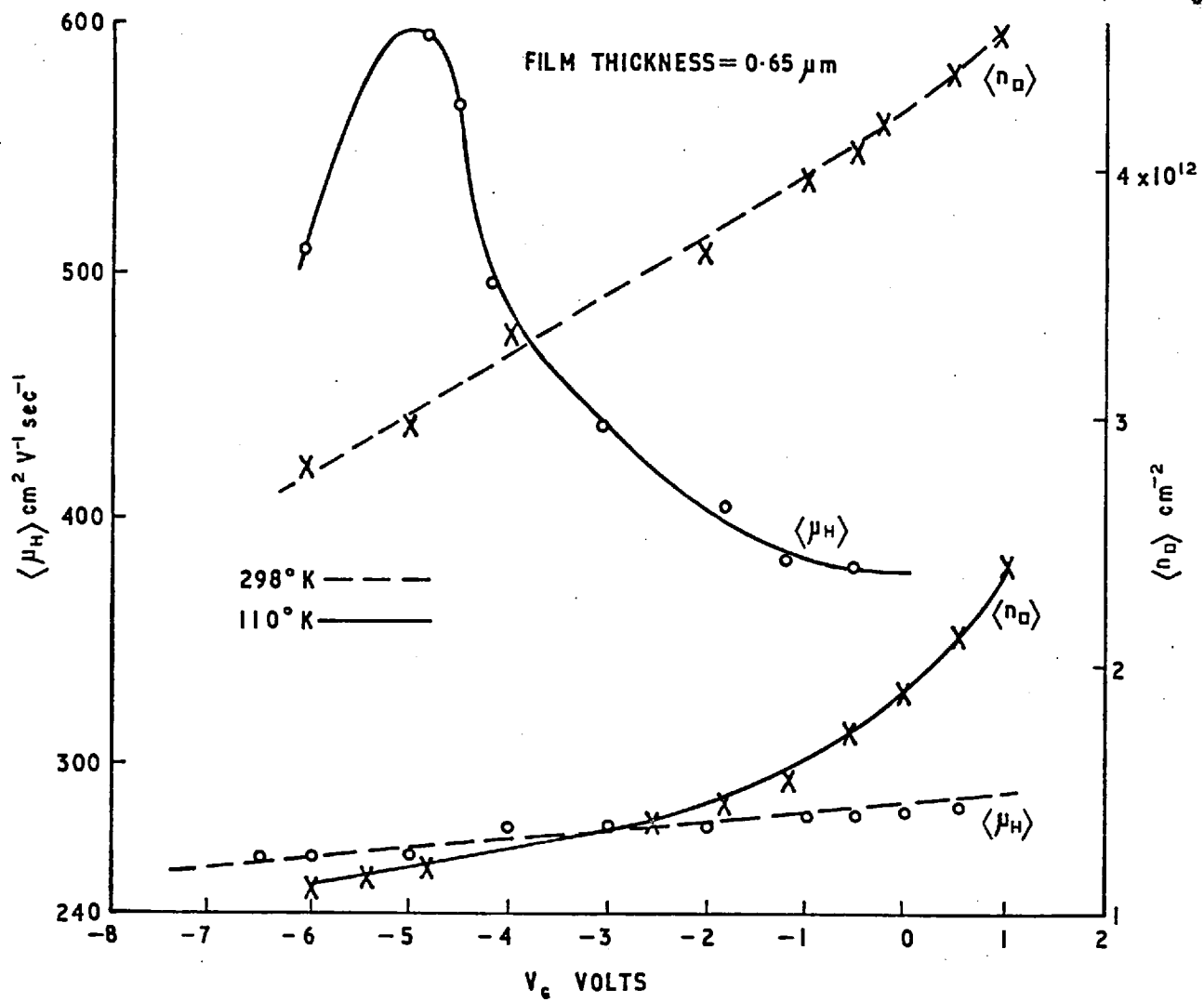
FIG. 3.4.1. 'MEAN MOBILITY' AND 'EFFECTIVE NUMBER OF CARRIERS PER cm² OF CHANNEL' AS A FUNCTION OF GATE VOLTAGE FOR A DEVICE ON WHICH PHOSPHORUS TRANSFER HAS OCCURRED

limited by scattering at the high density of positively charged ionised donors. The majority of the carriers are then in the less heavily doped bulk region of the film which has a higher mobility. If the mobility in the bulk region decreases as the sapphire interface is approached a mobility maximum will occur at the point where the depletion region is driven through the heavily doped skin. The capacitance at the mobility maximum corresponds to a depletion depth of slightly less than 2000Å, which is in fairly good agreement with the predicted diffusion length for the transferred phosphorus.

3.5 Electronic Interface Layers

Deviations from the general behaviour described in section 3.2 were also observed on another set of specimens. At room temperature they exhibited $\langle \mu_H \rangle$ plots which changed more rapidly with V_G and $\langle n_\square \rangle$ plots which tended to level off below the flatband value, some specimens even showed a gradual increase in $\langle n_\square \rangle$ as the gate voltage was made more negative. A typical set of results is shown in Fig 3.5.1. At lower temperatures the minimum in the $\langle n_\square \rangle$ plot is even more marked.

As the negative voltage on the gate is increased, the total negative charge in the channel must decrease. Therefore the minimum in the $\langle n_\square \rangle$ plot must be an anomaly of the averaging effect. If a high density of low mobility electrons is present at the sapphire interface the denominator in equation (3.2.1), which contains a μ^2 term, will decrease rapidly as the depletion region is driven back through the higher mobility regions near the oxide interface. It may decrease more rapidly than the numerator causing the $\langle n_\square \rangle$ value to increase with decreasing V_G. All the wafers had received a 30 min heat treatment at $460^\circ C$ in a moist atmosphere and it is possible that shallow donor states generated at the sapphire interface were responsible for the high density of low mobility electrons since Heiman (1967b) has reported that 15 min heating at $500^\circ C$ in a moist atmosphere is sufficient to generate interface donor states on a 'contaminated' wafer, although the contamination responsible has not yet been identified. Wrigley and Kroko (1969) have pointed out that the experimental evidence could equally well be explained by the generation of positive interface charge.

Simple mobility and carrier concentration profiles of the type shown in Fig 3.5.2 were assumed for the specimen in an attempt to explain the results in Fig 3.5.1. Values of the parameters in Fig 3.5.2 must be determined before an
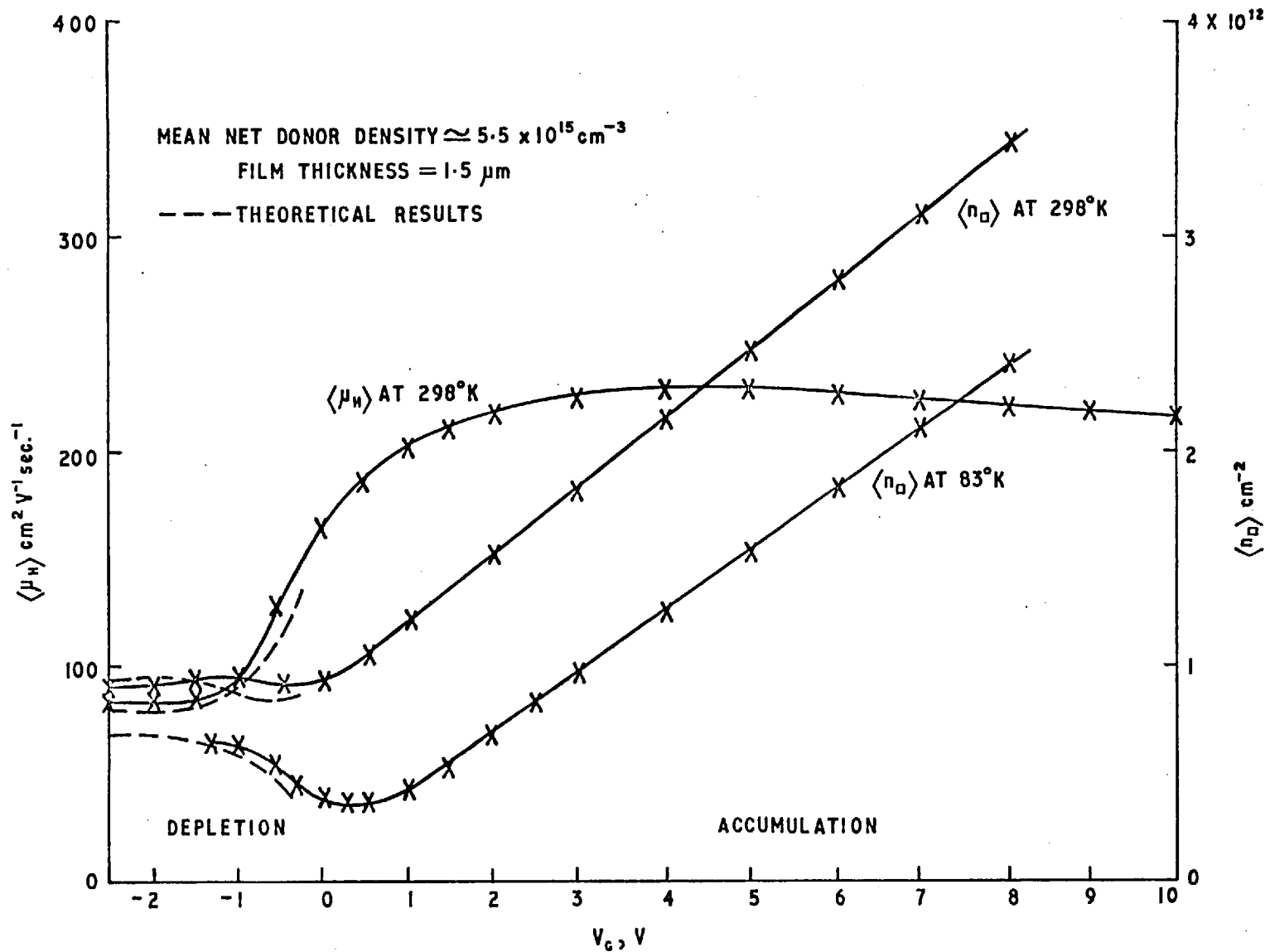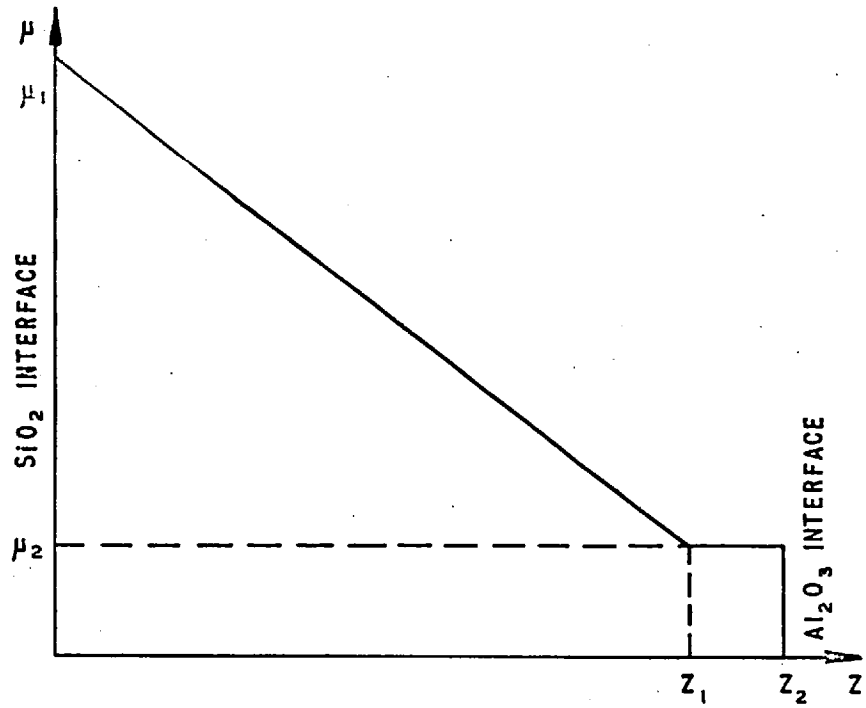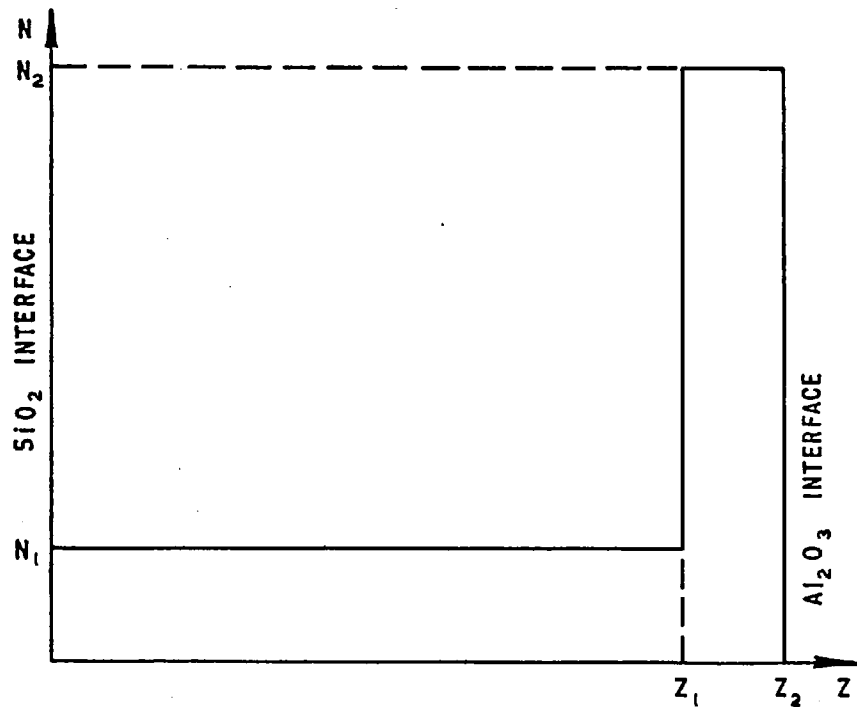
FIG. 3.5.1. EXPERIMENTAL AND THEORETICAL 'MEAN MOBILITY' AND 'EFFECTIVE NUMBER OF CARRIERS PER cm² ' AS A FUNCTION OF GATE VOLTAGE FOR A DEVICE WITH A HIGH DENSITY OF DONOR INTERFACE STATES

a. MOBILITY PROFILE

b. CARRIER CONCENTRATION PROFILE

FIG. 3.5.2. MODEL FOR DONOR INTERFACE
STATES CALCULATION

analysis can be undertaken. The value of N_1 was taken as the mean donor concentration in the depletion range as calculated from the C_{max}/C_{min} ratio of the capacitor C-V plot. The thickness of the film was measured at $1.5 \mu m$ using a Talysurf. The value of the mobility at the oxide interface was taken as the extrapolation of the 'mean mobility' in the accumulation region to flatband, which is a valid approximation since $r_m \ll 1$. Since the interface region between $Z = Z_1$ and $Z = Z_2$ is likely to be small compared with the film thickness, its exact dimensions are not important but the total charge in it and the mobility of the carriers are. These are two unknown quantities, although it is clear that the mobility in this region will be less than the minimum value of the 'mean mobility'.

A range of values was tried for the two unknowns to see if a reasonable fit with the experimental results could be obtained by analysing the model in Figure 3.5.2 using equations (3.2.1) and (3.2.2).

The parameters in the table below give theoretical results, (shown by the dashed lines), which are in good agreement with the experimental results, (shown by the solid lines), in Figure 3.5.1 even to the extent of duplicating the shallow dip in the $\langle n_{\square} \rangle$ plot.

Mobility $cm^2 \ V^{-1} \ S^{-1}$		Distance cm		Carrier Density cm^{-3}	
μ_1	μ_2	Z_1	Z_2	N_1	N_2
230	30	1.49×10^{-4}	1.50×10^{-4}	5.5×10^{15}	9.0×10^{17}

The $\langle n_\square \rangle$ points agree to better than 8 per cent and the $\langle \mu_H \rangle$ points to better than 11 per cent. The above results give an interface donor state density of 9×10^{11} cm^{-2} and the value was always within 7×10^{11} cm^{-2} to 2.5×10^{12} cm^{-2} which lies within the 10^{11} to 10^{13} range reported by Heiman (1967). The sapphire interface mobility of 30 cm^2 V^{-1} sec^{-1} is less than the 100 cm^2 V^{-1} sec^{-1} reported by Wrigley and Kroko (1969), who measured electron field-effect mobilities in inversion layers adjacent to the sapphire interface by using 75 μm thick sapphire substrates as the dielectric in MIS transistors. Since the films used by Wrigley and Kroko had an average electron mobility of 400 cm^2 V^{-1} sec^{-1} they may have had a lower overall density of crystal defects than those used in this investigation. The difference between the 30 cm^2 V^{-1} sec^{-1} interface mobility for films with electronic layers at the sapphire interface and the 50 cm^2 V^{-1} sec^{-1} value for films without interface inversion layers is not really significant because of the approximate nature of the above analysis, but a reduction is to be expected due to enhanced scattering at ionised donor centres on the Heiman model or due to enhanced interface scattering on the Wrigley and Kroko model if the sapphire interface is not entirely specular.

As the temperature is reduced the carrier density in the bulk of the film reduces but not at the sapphire interface because the donor density in the Heiman model is approaching degeneracy and the electron density in the Wrigley and Kroko model is determined by the positive interface charge. This increase in the ratio of interface carrier density to the bulk carrier density explains the larger dip in the $\langle n_\square \rangle$ plot at lower temperatures and when the appropriate $83^\circ K$ values are substituted in the profile model the results obtained are again in good agreement with the experimental observations, as shown in Fig 3.5.1.

Even with this simple model it has been possible to explain all the features of the unusual behaviour of this type of specimen. Although an accurate measurement of the density of interface donor states has not been obtained, an estimate of the approximate magnitude has been possible.

3.6 Fast State Density at the Oxide Interface

The conventional high frequency (Terman 1962 and Zaininger and Warfield 1965), low frequency (Berglund 1966) and quasi-static techniques (Kuhn 1970) for determining the surface state density as a function of the position in the band gap all require the comparison of an experimental curve with an ideal curve.

The ideal curves can readily be computed if the substrate doping is known and uniform but these techniques cannot be applied to substrates with non-uniform and unknown doping profiles such as exist with silicon-on-sapphire films. Hall measurements can be used in the following manner to provide the additional information necessary to analyse the high-frequency C-V curves for the case of non-uniform substrates.

The charge neutrality condition for an MIS capacitor gives the relationship

$$Q_M + Q_{SS} + Q_S + Q_i = 0 \qquad \dots\dots\dots\dots(3.6.1)$$

where Q_M = charge in the gate metal, Q_{SS} = charge in surface states, Q_S = charge in surface states, Q_S = charge in the semiconductor and Q_i = charge in the insulator.

When the gate voltage is changed the first three terms in equation (3.6.1) may change but Q_i is independent of V_G. The value of ΔQ_M for an incremental change ΔV_G in the gate voltage can be determined from the MIS capacitance since

$$C = \frac{dQ_M}{dV_G} \text{ therefore } \Delta Q_M = C\Delta V_G.$$

Hence if ΔQ_S can be evaluated, the value of ΔQ_{SS} follows simply from

$$\Delta Q_{SS} = - (\Delta Q_M + \Delta Q_S). \qquad \dots\dots\dots\dots\dots(3.6.2)$$

Since ΔQ_S for a non-uniform substrate cannot be evaluated theoretically it must be measured. In the range of the 'depletion approximation', where $-3 > u_S > 2u_F + 3$, the ΔQ_S is entirely due to ionised donors in the depletion region and it follows from equation (3.2.11) that

$$\Delta Q_S = q \frac{\left[\langle n_{01}\rangle \langle \mu_{01}\rangle - \langle n_{02}\rangle \langle \mu_{02}\rangle\right]^2}{\langle n_{01}\rangle \langle \mu_{01}\rangle^2 - \langle n_{02}\rangle \langle \mu_{02}\rangle^2} \quad \dots\dots\dots\dots\dots(3.6.3)$$

Hence in the region of the 'depletion approximation' by substitution of ΔQ_S and ΔQ_M in equation (3.6.2) the value of ΔQ_{SS} for any ΔV_G can be calculated. The relationship between V_G and ϕ_S is

$$V_G - \phi_{MS} - \phi_S = - \frac{(Q_{SS} + Q_S)}{C_i} \quad \dots\dots\dots\dots\dots(3.6.4)$$

and from equation (3.6.2)

$$\Delta\phi_S = \Delta V_G - \frac{\Delta Q_M}{C_i}$$

$$= \Delta Q_M \left(\frac{1}{C} - \frac{1}{C_i}\right) \quad \dots\dots\dots\dots\dots(3.6.5)$$

A plot of $N_{SS} = \Delta Q_{SS}/q\,\Delta\phi_S$ as a function of ϕ_S, with an arbitrary zero for ϕ_S, can now be determined. The approximate position of the true ϕ_S zero, (ie flatband), is readily determined from the C-V plot by the method of Lehovec (1968), although strictly speaking that analysis also assumes a uniform doping level in the substrate.

A typical set of N_{SS} results is shown in Fig (3.6.1). These results are very similar in form to those obtained by Kuhn (1970) on bulk silicon and the values are approximately 50 per cent lower.

SPECIMEN 505/2/7

FILM THICKNESS = 1·1 μm

T = 298°K

MEAN NET DONOR DENSITY ≃ 1·5 x 10^16 cm^-3
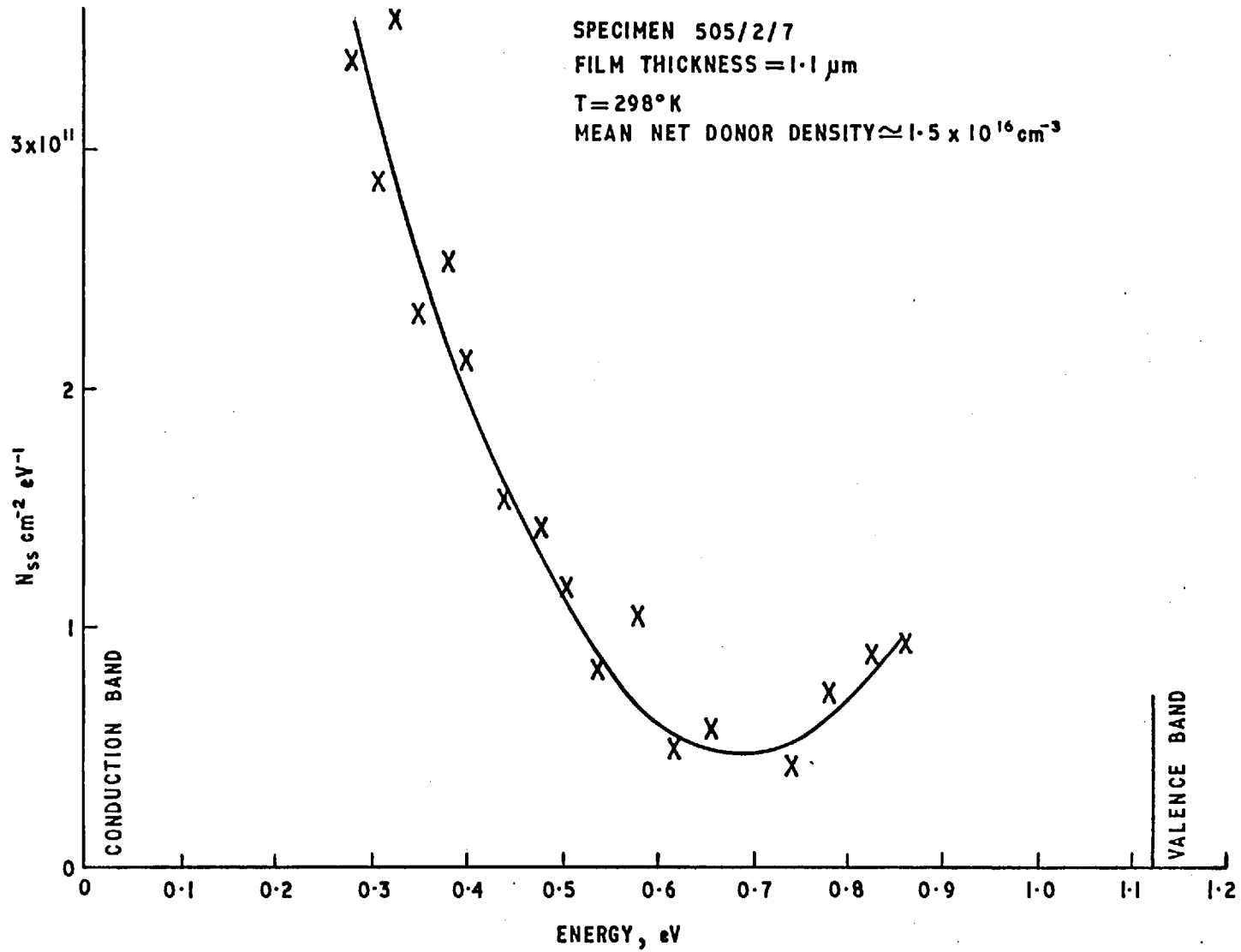
FIG.3.6.1. ENERGY DISTRIBUTION OF FAST STATES AT THE OXIDE INTERFACE

By using the Brown and Gray (1968) technique, C-V measurements on the capacitor at $300^{\circ}K$ and $80^{\circ}K$ enable a value for the total number of interface states between the $300^{\circ}K$ Fermi level and the $80^{\circ}K$ Fermi level to be calculated. These Fermi levels for the specimen in Fig (3.6.1), ($N_D = 1.5 \times 10^{16}$ cm^{-3}) are at 0.21 eV and 0.05 eV respectively from the conduction band edge. Hence the $\Delta Q_{SS}/q$ value of 2.5×10^{11} cm^{-2}, measured by the Brown and Gray experiment corresponds to a mean N_{SS} value of 1.55×10^{12} cm^{-2} eV^{-1} which is well above the mid-gap values obtained by the above technique. This rapid increase of the interface state density near the band edge is a feature of all the reported work on the $Si-SiO_2$ interface. Clearly neither the high compressive stress nor the high density of crystal faults in silicon-on-sapphire films substantially effects the density or energy distribution of silicon-oxide interface states.

CHAPTER 4

CARRIER TRANSPORT STUDIES USING JUNCTION FIELD-EFFECT STRUCTURES

It was shown in the analysis of the MIS deep depletion profiling measurements

described in the previous chapter that with that particular technique there

is a limit to the thickness of layer which can be measured for a given

resistivity of silicon, (See Figure 3.1.2). However if the MIS structure

is replaced by a reverse-biased p^+-n junction a wider depletion layer can

be developed and therefore, at a given resistivity of silicon, thicker

layers can be measured.

4.1 Experimental Techniques

Two forms of junction field-effect structures are used in these investigations and they are shown in Figures 4.1.1 (a) and 4.1.1 (b). In some of the first devices to be measured, which were fabricated according to the design in Fig 4.1.1 (a), substantial currents flowed to the gate when it was reverse-biased. This gave rise to significant errors in the conductivity measurements, particularly those low current measurements which relate to the region close to the sapphire interface. The gate current at a given reverse-bias voltage slowly increased after the application of the reverse-bias voltage until it reached a limiting characteristic of the form shown in Figure 4.1.2 after a period of 2-3 hours. It is believed that the increase in the reverse-bias current is due to the gradual formation of an interface inversion layer induced by ions migrating on top surface of the oxide in the fringing field of the junction, (Atalla, Bray and Lindner 1960). Some substantiation of the above hypothesis was obtained by noting that the rate of change of current increased when the humidity level of the atmosphere surrounding unencapsulated devices was raised.

The design shown in Figure 4.1.1 (b), which overcomes the surface ion migration problem by including an insulated gate electrode over the edges of the gate junction in order to control the interface potential, is used in all the later investigations. All the devices have a channel width of 600 μm and the lengths are 40 μm and 52 μm for the designs in Figures 4.1.1 (a) and 4.1.1 (b) respectively. The gate is separated from both the source and the drain by 10 μm of the relatively high resistivity starting material. The additional series resistance due to this material is accounted for in the analysis of the results. The gate junction was kept as shallow as possible and was always less than 0.6 μm. The processing procedure is described in Appendix 3.

91
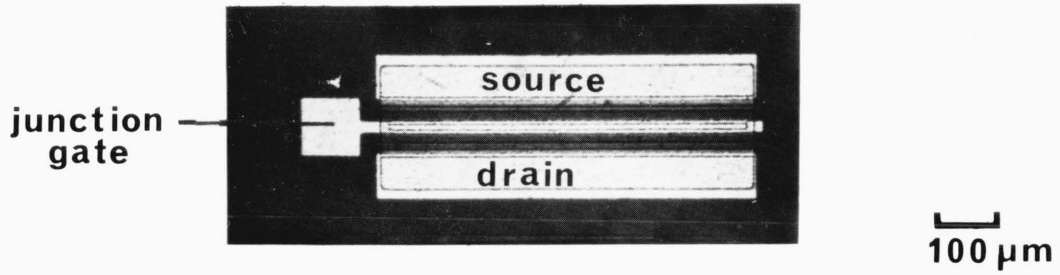
(a)   Junction field — effect structure without insulated gate
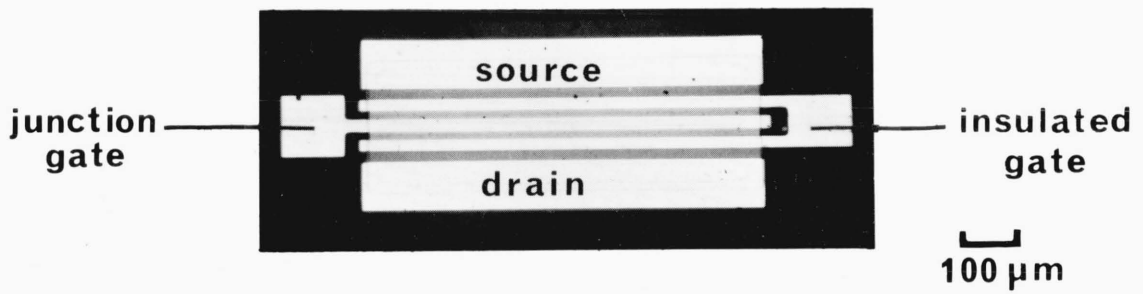
(b)   Plan view of junction field effect structure with
insulated gate over the junction gate perimeter
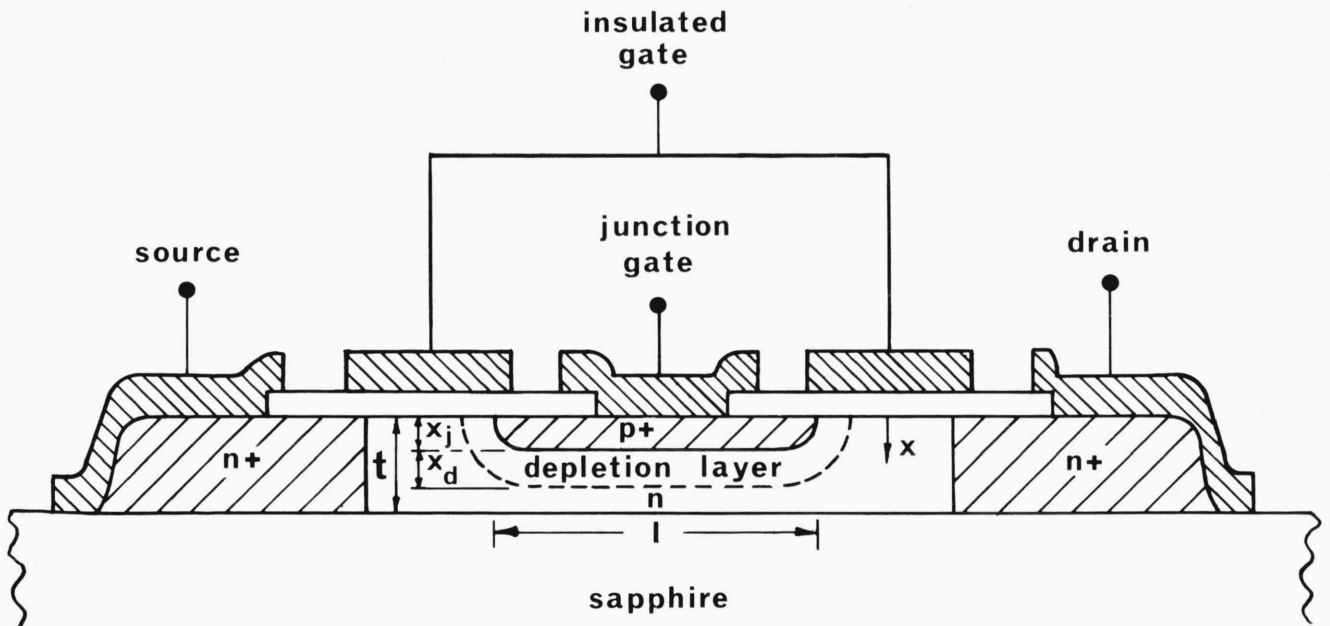
(c) cross section of junction field effect structure
with insulated gate over the junction gate perimeter

FIGURE 4. 1. 1. Junction Field Effect Test Structures
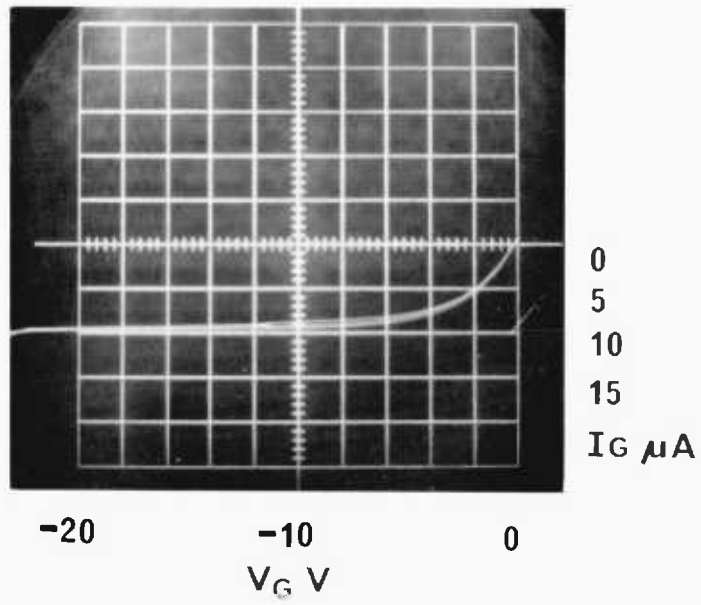
Figure 4.1.2. Gate Junction Leakage Current

In order to convert the information on conductivity as a function of the reverse-bias voltage on the gate into a conductivity versus depth characteristic the depletion layer depth must be determined as a function of the reverse-bias voltage from small signal differential capacitance measurements. These measurements can also be used to obtain the profile of electrically active impurities in the film, (see for example Thomas, Kahng and Manz 1962 or Gupta 1968). It has been shown by Kennedy and O'Brien (1969) and by Grimshawe and Osborne (1971) that under conditions where charge neutrality does not exist, (ie the majority carrier density is not equal to the net density of ionised donors and acceptors), then the analysis of the high frequency differential capacitance measurements determines the majority carrier density rather than the net ionised donor and acceptor density. Hence, since the conductivity is the product of the majority carrier density and the mobility, the latter quantity can also be determined as a function of depth from the conductivity and differential capacitance measurements.

The measurements were normally made by probing unencapsulated devices because any capacitance associated with package leads would be significant compared with the gate capacitance, which is only of the order of 3 pF for a depletion depth of $1\,\mu m$. However, the few measurements made with packaged devices did give the same results if an appropriate allowance was made for the package capacitance.

In the absence of any excess gate currents due to surface ion migration the current flowing through a reverse-biased gate junction arises from electron-hole pair generation both in the bulk and at the surface of the silicon. It is shown in section 4.4 that the bulk contribution is much greater than the surface contribution because silicon-on-sapphire films possess relatively high densities of bulk generation - recombination centres, (Heiman 1967 b).

94

Furthermore it is also shown that the bulk contribution comes predominantly from the depletion region. Therefore the changes in gate current resulting from incremental changes in reverse-bias voltage can be analysed together with the differential capacitance measurements to determine the effective carrier lifetime as a function of depth in the silicon film. Hence profiles of mobility, majority carrier density and carrier lifetime can all be measured on the same small area of silicon film using one simple structure.

When compared with the measurements made on the Hall effect structures in the previous chapter, the present measurements have the advantage that a much wider range of film thicknesses and resistivities can be examined. The film thickness that can be measured in the junction field effect measurements at a given resistivity is limited by the onset of avalanche breakdown in the depletion region adjacent to the gate junction. In a plane junction the onset of avalanche breakdown occurs at a depletion depth of approximately 20 μm for a doping level of 10^{15} cm^{-3} and at 0.4 μm for a doping level of 10^{17} cm^{-3}. Although the practical limits are much less than this due to the curvature of the shallow gate junction (Leistiko and Grove 1966), they are still significantly greater than the corresponding MIS values of 1 μm at 10^{15} cm^{-3} and 0.1 μm at 10^{17} cm^{-3}. However , since there is the disadvantage that no profile measurements can be made on the top 0.5 μm of the film because it is used for the gate junction, the techniques are really complementary.

4.2 Calculation of the Majority Carrier Density in the Presence of Traps

The analysis of the differential capacitance of a reverse-biased p-n junction as a function of the applied voltage has been employed extensively in the study of impurity distributions in semiconductors, (Hillibrand and Gold 1960, Thomas, Kahng and Manz 1962, Meyer and Guldbransen 1963 and Gupta 1968). However, when large numbers of deep trapping levels exist in the semiconductor, as is generally the case with silicon-on-sapphire films, particular care must be taken in the analysis of the results because the relatively slow charging and discharging processes for the deep traps cause the capacitance to be frequency dependent. The behaviour of p-n junctions in silicon-on-sapphire is now described in a manner similar to that used by Sah and Reddi (1964) for gold doped p-n junctions.

Although both Heiman (1967 b) and Dumin (1970) have reported a spread in the energy levels of the deep acceptors in silicon-on-sapphire films an analysis based on a single level model will be given first in order to simplify the theoretical calculations and to clarify the description of the charge modulation processes. Modifications in the behaviour due to the spread in energy levels is then discussed.

The energy-band diagram for the single trap level model of the p^+-n junction is given in Figure 4.2.1 (a), where E_t is the deep acceptor trap level, V_J is the applied voltage and E_{FN} and E_{FP} are the quasi Fermi levels for electrons and holes respectively. The corresponding potential distribution is given in Figure 4.2.1 (b), where ϕ_n and ϕ_p are the electron and hole quasi Fermi potentials and ϕ_i is the intrinsic potential. Figure 4.2.1 (c) shows the net charge distribution
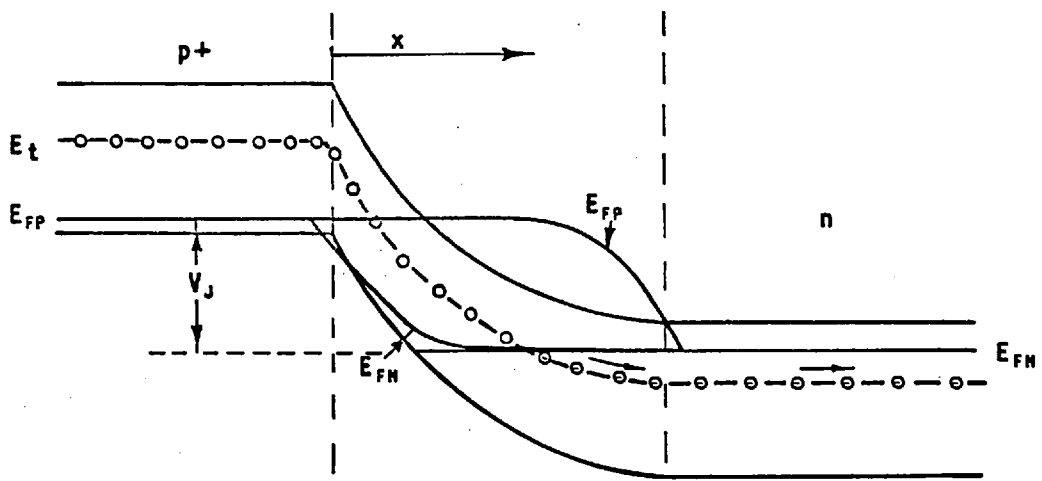
a. ENERGY DIAGRAM
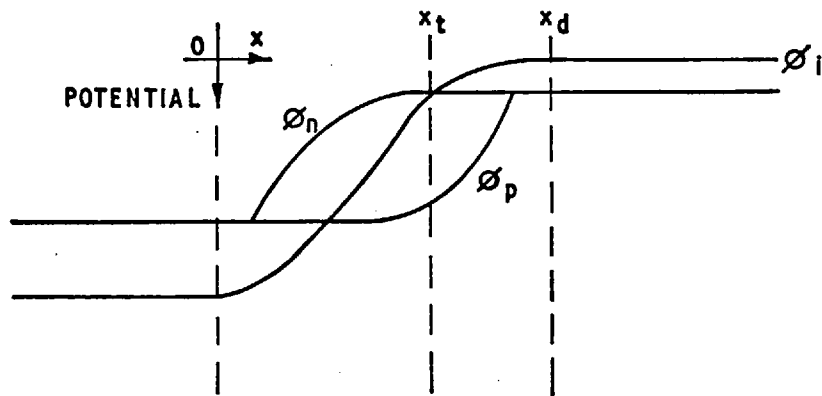
b. POTENTIAL DISTRIBUTION
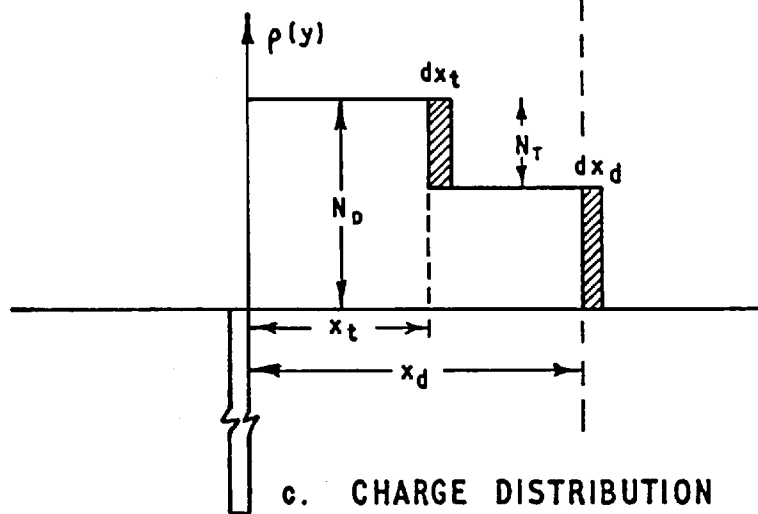
c. CHARGE DISTRIBUTION
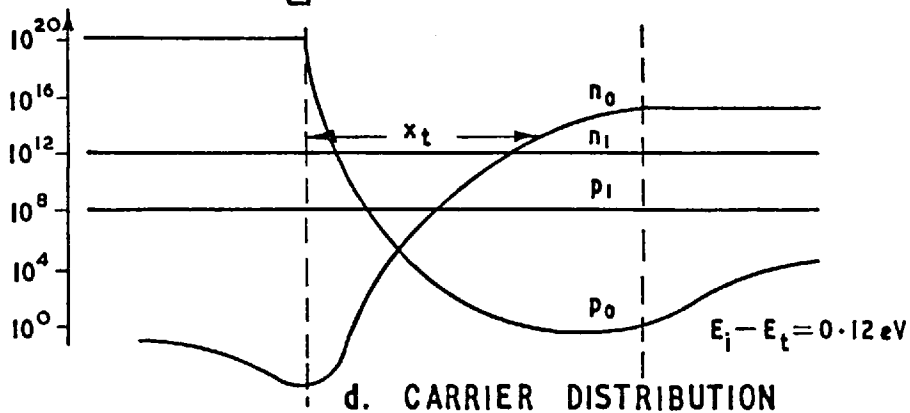
d. CARRIER DISTRIBUTION

FIG. 4.2.1. ENERGY BAND DIAGRAM, POTENTIAL DISTRIBUTION CHARGE DISTRIBUTION AND CARRIER DISTRIBUTION FOR A REVERSE BIASED P+-N JUNCTION

and the shaded regions depict the change in charge distribution resulting from an incremental change d V_J in the applied voltage. N_D is the density of shallow donors, N_T is the density of deep acceptor traps, x_d is the depth of the depletion region and x_t is the distance from the intersection of the n-type quasi Fermi level and the trap level to the metallurgical junction. The change $(N_D-N_T)\, dx_d$ is due to majority carriers being repelled from the edge of the depletion region and this process occurs very rapidly, (in the dielectric relaxation time) and the change $N_T\, dx_t$ is due to the net loss of electrons by the deep acceptors and this process has been shown by Sah and Reddi (1964) to have a time constant given by the expression

$$\tau = \left\{ v_{th}\, \sigma_p\, (p_o + p_1) + v_{th}\, \sigma_n\, (n_o + n_1) \right\}^{-1} \qquad \ldots 4.2.1$$

where $v_{th} = \left(\dfrac{3kT}{m} \right)^{\frac{1}{2}}$ is the thermal velocity of the carriers, σ_p and σ_n are hole and electron capture cross sections respectively, p_o and n_o are steady state hole and electron concentrations in the depletion region and p_1 and n_1 are given by

$$p_1 = \frac{n_i^2}{n_1} = n_i \exp \left(\frac{E_i - E_T}{kT} \right) \qquad \ldots 4.2.2$$

where n_i is the intrinsic carrier concentration.

Sah and Reddi (1964) evaluated equation 4.2.1 using the approximation $p_o = n_o = 0$ in the depletion region but it is now demonstrated that this approximation is not valid. Using the normal Boltzmann approximation (see for example p 96 of Sze 1969), p_o and n_o are given by

$$p_o = n_i \exp \left(\frac{E_i - E_{FP}}{kT} \right) \qquad \ldots 4.2.3$$

$$\text{and} \quad n_o = n_i \exp \left(\frac{E_{FN} - E_i}{kT} \right) \qquad \ldots 4.2.4$$

and if these equations are substituted into equation 4.2.1 then τ^{-1} can be separated into the following four terms

1. hole capture probability $= v_{th} \, \sigma_p \, p_o$
   $$= v_{th} \, \sigma_p \, n_i \exp \left(\frac{E_i - E_{FP}}{kT} \right) \ldots 4.2.5$$

2. hole emission probability $= v_{th} \, \sigma_p \, p_1$
   $$= v_{th} \, \sigma_p \, n_i \exp \left(\frac{E_i - E_T}{kT} \right) \quad \ldots 4.2.6$$

3. electron capture probability $= v_{th} \, \sigma_n \, n_o$
   $$= v_{th} \, \sigma_n \, n_i \exp \left(\frac{E_{FN} - E_i}{kT} \right) .. \, 4.2.7$$

4. electron emission probability $= v_{th} \, \sigma_n \, n_1$
   $$= v_{th} \, \sigma_n \, n_i \exp \left(\frac{E_T - E_i}{kT} \right) .. \, 4.2.8$$

At $x = x_t$, where the trapped charge modulation takes place, $E_{FP} \gg E_T$ and hence $p_o \ll p_1$, (as shown in the carrier distributions of Figure 4.2.1 (d)). It therefore follows that term 1 is very much smaller than term 2. In order to compare the magnitudes of terms 2 and 4 the value of $E_T - E_i$ must be known. The range of reported values is from 0.09eV (Heiman 1967) to 0.31eV (Dumin 1970), but even the smaller of these values gives a ratio of n_1/p_1 of the order of 10^3 and therefore term 2 can also be neglected, assuming that σ_p is not very much greater than σ_n. Finally terms 3 and 4 are equal at $x = x_t$ because $E_{FN} = E_T$. Therefore the reciprocal of the response time of the deep acceptor traps is given by

$$\tau^{-1} = 2 \; v_{th} \; \sigma_n \; n_i \; \exp \left(\frac{E_T - E_i}{kT} \right) \qquad \dots 4.2.9$$

This value is twice the value derived by Sah and Reddi because they neglected the electron capture contribution.

In the absence of any information on the capture cross section of the deep acceptors in silicon-on-sapphire an approximate value of 10^{-15} cm^2 is assumed. This was the value measured by Glaenzer and Jordan (1969) for edge dislocations in silicon and is within an order of magnitude of the value for a wide range of recombination centres in silicon. Substituting this capture cross section and the 0.09eV and 0.31eV limiting values of $E_T - E_i$ in equation 4.2.9 gives corresponding values for τ of 10^{-4} s and 2×10^{-8} s respectively. Since these values span the 1 MHz frequency used for the capacitance measurements it is clearly necessary to measure the frequency dependence of the junction capacitance to demonstrate that a 1 MHz

measurement is truly a 'high frequency' measurement. The frequency
dependence measurements are presented along side the other experimental
results in section 4.6.

Assuming for the present that the measurement frequency is sufficiently
high to eliminate the possibility of hole capture and emission at x_t then
all the charging and discharging takes place at the depletion layer edge
and dQ, the incremental change in charge density, for a change dV_J in the
applied reverse-bias voltage is equal to $q(N_D-N_T) dx_d$. Since dQ is also
equal to the product of the differential capacitance per unit area
of the gate junction and dV_J it follows that

$$q (N_D-N_T) dx_d = \frac{\varepsilon_{si}}{x_d} d V_J$$

Therefore

$$\tfrac{1}{2} d (x_d^2) = \frac{\varepsilon_{si}}{q (N_D-N_T)} d V_J$$

and

$$N_D-N_T = \frac{2}{q \varepsilon_{si} A_J^2} \cdot \frac{1}{d (1/c_J^2)/dV}$$

$$= \frac{c_J^3}{q \varepsilon_{si} A_J^2 (d c_J/dV)} \qquad \dots 4.2.10$$

where C_J and A_J are the differential capacitance and area of the gate junction respectively.

It is also known that

$$x_d = \frac{\varepsilon_{si} A_J}{C_J} \qquad \ldots 4.2.11$$

Therefore equations 4.2.10 and 4.2.11 can be used to determine the 'impurity profile'.

It has been assumed in the derivation of equation 4.2.10 that in the region of the semiconductor under investigation the electron density is equal to the net donor concentration. However Kennedy, Murley and Kleinfelder (1968) have pointed out that this charge neutrality assumption is not always valid and that it is particularly suspect when the net donor concentration on the less heavily doped side of the junction is less than 10^{16} cm^{-3}. In semiconductor regions which do not exhibit charge neutrality the net donor density, $(N_D - N_T)$, in equation 4.2.10 must be replaced by the majority carrier density $n(x)$ giving

$$n(x) = \frac{C_J^3}{q \, \varepsilon_{si} \, A_J^2 \, (d \, C_J / d \, V_J)} \qquad \ldots 4.2.12$$

The following analysis, originally developed by Kennedy and O'Brien (1969), shows how the net donor concentration profile can be derived from the measured majority carrier density profile. The electron current in the n-type material is the sum of diffusion

and drift components and is given by

$$J_n = q D_n \frac{dn}{dx} - q \mu_n n \frac{d\phi}{dx} \qquad \dots 4.2.13$$

where D_n is the diffusion constant for electrons and ϕ is the electrostatic potential. If it is assumed that the junction is perfectly blocking and hence that the electron current is zero it follows that

$$E(x) = - \frac{d\phi}{dx} = \frac{kT}{q} \cdot \frac{1}{n(x)} \frac{dn(x)}{dx} \qquad \dots 4.2.14$$

The perfectly blocking assumption may seem a doubtful one for silicon-on-sapphire p-n junctions where the reverse-biased leakage currents may be as high as 10^{-2} A cm^{-2}, but this is still much less than the typical diffusion current of 40 A cm^{-2} which is obtained by substituting $D_n = 2.5$ cm s^{-1} and $(dn/dx) = 10^{20}$ cm^{-4}.

Now Poissons equation gives

$$\frac{dE}{dx} = \frac{q}{\varepsilon_{si}} [N_D(x) - N_T(x) - n(x)] \qquad \dots 4.2.15$$

if the negligibly small minority carrier concentration is ignored.

Equations 4.2.14 and 4.2.15 can be combined to give the following
expression for the net donor concentration.

$$N_D(x) - N_T(x) = n(x) + \left(\frac{kT}{q}\right)\left(\frac{\varepsilon_{si}}{q}\right) \quad \frac{d}{dx}\left[\frac{1}{n(x)} \quad \frac{dn(x)}{dx}\right] \quad \ldots \text{ 4.2.16}$$

It is interesting to note that if a positive charge at the sapphire
interface induces excess majority carrier charge near the interface,
equations 4.2.14 and 4.2.15 are still valid and therefore equation
4.2.16 can still be used to evaluate $N_D(x) - N_T(x)$. The increase in
$n(x)$ is exactly compensated by a negative contribution from the
second term in equation 4.2.16. However the validity of equation 4.2.16
is destroyed if a significant minority carrier density is induced by
a negative interface charge.

Although it has been shown that the net donor concentration can
usually be determined from equation 4.2.16 it should be noted that
equation 4.2.16 contains a derivative of $n(x)$ which itself is
obtained from a differentiation of the C-V curve, (equation 4.2.12.)
Since the C-V curve is effectively differentiated twice, small
linearity errors in the capacitance measurements can give rise to
fairly large errors in the second term in equation 4.2.16 and
therefore in $(N_D(x) - N_T(x))$ if the deviation from charge neutrality
is significant.
A spread in the deep acceptor energy levels will not change the general
behaviour described above. It will simply extend the frequency range
over which the transition from the low frequency capacitance to the high
frequency capacitance takes place.

4.3 Conductivity Mobility

At low drain voltages the resistance of the junction field-effect structure shown in Figure 4.1.1(c) approximates to that of a resistor of width ω, length ℓ and thickness $(t-x_d-x_j)$ where t is the thickness of the silicon film, x_j is the depth of the p^+-n junction and x_d is the depth of the depletion region. However at any finite drain voltage there is a small discrepancy in the approximation because the depletion depth is slightly greater at the drain end than at the source end. This causes the drain conductance to be slightly less than the zero drain voltage limiting value of

$$g_{do} = \frac{\omega \, q}{\ell} \int_t^{t-x_d-x_j} n(x) \, \mu(x) \, dx \qquad \ldots\ldots 4.3.1$$

where $n(x)$ and $\mu(x)$ are the majority carrier density and mobility respectively at position x.

The magnitude of the errors arising from the need to use a finite drain voltage are now discussed for a typical specimen. Consider a structure with a channel depth, $(t-x_j)$, of 1 micron, a constant net donor density of 5×10^{15} cm^{-3} and a constant mobility of 200 cm^2 V^{-1} s^{-1}. Sze (1969 p 347) has shown that the drain current, I_D, in a junction field-effect transistor with constant doping and constant mobility is given by
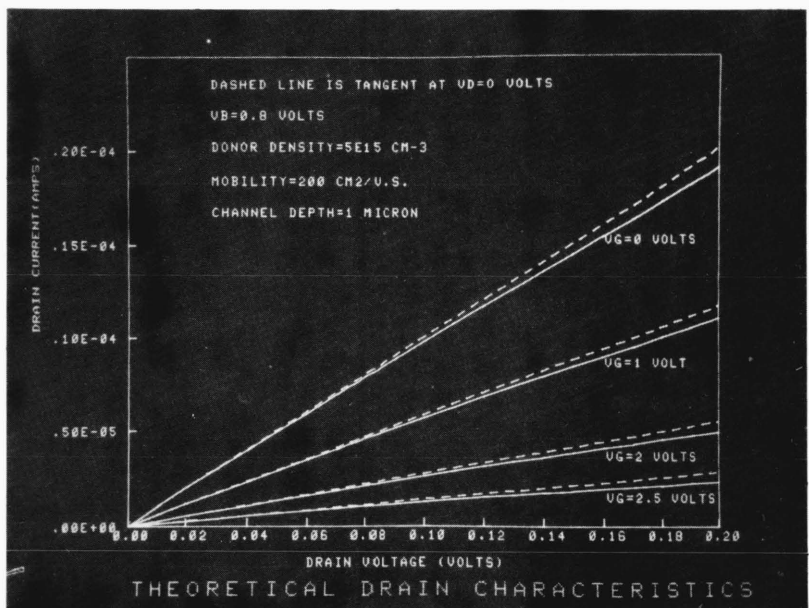
FIGURE 4. 3. 1. Junction Field Effect Transistor drain characteristics
at low values of drain voltage.
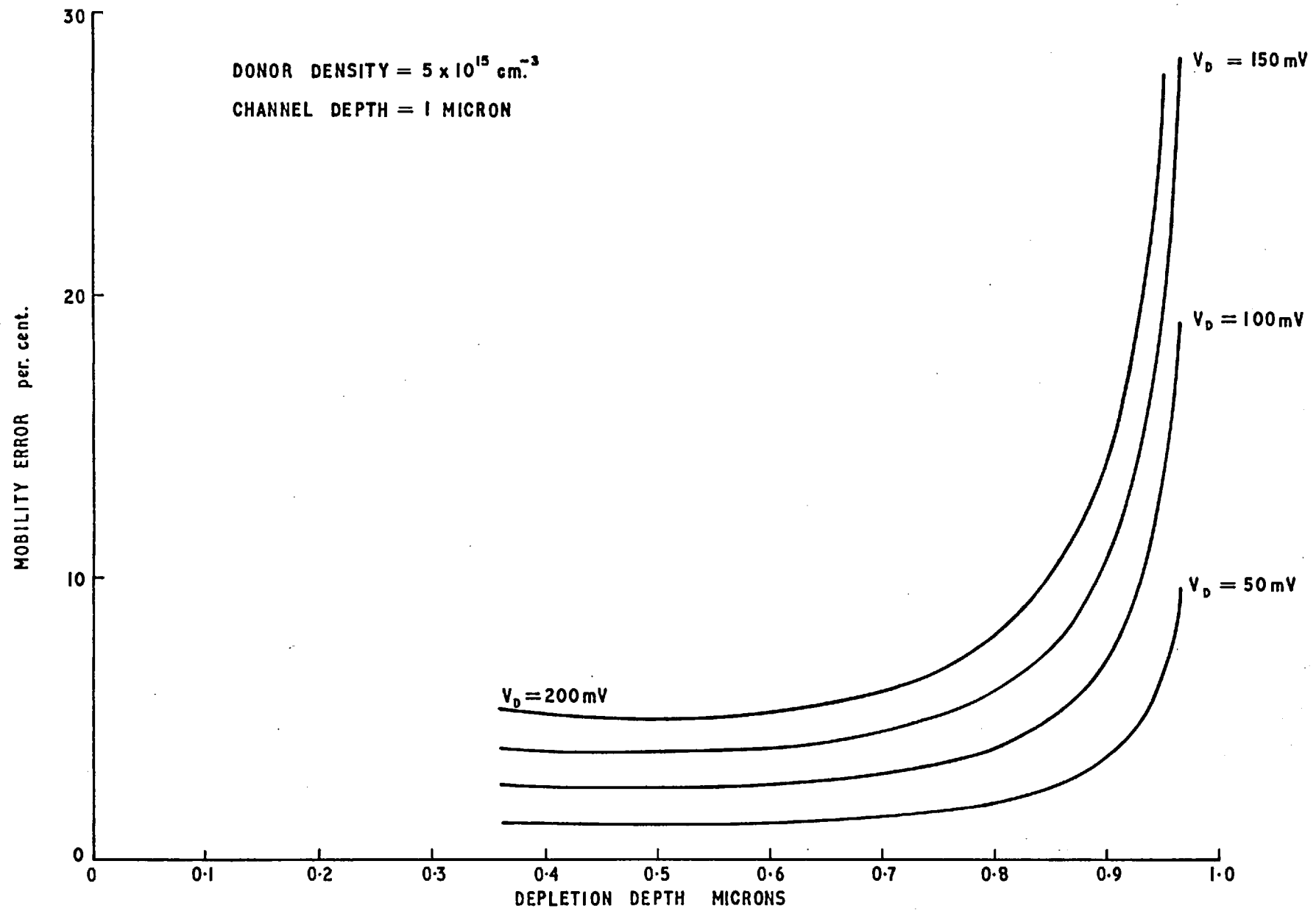
FIG. 4.3.2. MOBILITY ERRORS DUE TO DRAIN PINCHING EFFECT

$$I_D = g_{max} \left\{ V_D - \frac{2}{3(t-x_j)} \left(\frac{2\varepsilon_{si}}{qN_D} \right)^{\frac{1}{2}} \left[(V_D + V_G + V_b)^{\frac{3}{2}} \right. \right.$$

$$\left. \left. - (V_b + V_G)^{\frac{3}{2}} \right] \right\} \qquad \ldots\ldots 4.3.2$$

where $g_{max} = \omega \mu q N_D (t-x_j)/\ell$ is the limiting drain conductance as

$V_D \rightarrow 0$ and $V_G + V_b \rightarrow 0$ and V_b is the built-in potential across the gate

junction. If a typical value for V_b of 0.8V is substituted into

equation 4.3.2 together with the transistor parameters given above the

calculated drain characteristics at low values of drain voltage are as

shown in Figure 4.3.1. The dashed line in this Figure is $I_D = g_{do} V_D$

and the difference between the dashed line and the full line is a

measure of the error in the mobility calculation which results from the

use of a finite drain voltage. This error is plotted as a function of

the depth of the depletion region for four different values of drain

voltage in Figure 4.3.2. At the drain voltage of 100mV chosen for the

subsequent measurements the error is below 5 per cent in this typical

specimen until the depletion region encroaches within 0.15 μm of the

silicon-sapphire interface. Although a constant doping and constant

mobility model is used for this calculation it gives a good indication

of the behaviour to be expected in non-uniform films provided that the

doping level and mobility at the depletion layer edge are not very

different from their corresponding mean values in the bulk of depletion

layer.

It has been demonstrated that the errors due to the finite drain voltage can normally be neglected but there is another complication in the channel conductance measurements which cannot be neglected. In series with the channel are two 10 μm lengths of unmodulated material whose resistance R_t can be calculated from their geometry using a subsidiary sheet resistance measurement and the total resistance between the source and the drain is given by

$$R = 2R_t + \frac{\ell}{\omega G_s}$$

$$= 2R_t + \frac{\ell}{\omega q} \frac{1}{\displaystyle\int_0^{t-x_j-x_d} n(x)\,\mu(x)\,dx} \qquad \ldots\ldots 4.3.3$$

where G_s is the sheet conductivity in the channel.

If there is an incremental change in the reverse bias voltage across the gate junction then the new total resistance is

$$R + \Delta R = 2R_t + \frac{\ell}{\omega q} \frac{1}{\displaystyle\int_0^{t-x_j-(x_d+\Delta x_d)} n(x)\,\mu(x)\,dx} \qquad \ldots 4.3.4$$

Where Δx_d is the incremental change in the depletion depth, and the change in the channel conductance is given by

$$\frac{\omega \Delta G_s}{\ell} = \frac{\omega \, q}{\ell} \int_{t-x_j-(x_d+\Delta x_d)}^{t-x_j-x_d} n(x) \, \mu(x) \, dx$$

$$= \left(\frac{1}{R - 2 R_t} - \frac{1}{R + \Delta R - 2 R_t} \right) \qquad \ldots \ldots \ 4.3.5$$

If the carrier density and mobility are continuous functions of position, then for a sufficiently small change in the position of the depletion layer edge equation 4.3.5 gives

$$\mu(x) = \frac{\Delta G_s}{q \, n(x) \, \Delta x_d}$$

$$= \left(\frac{\ell}{\omega \, q} \right) \left(\frac{1}{n(x) \, \Delta x_d} \right) \left(\frac{1}{R - 2 R_t} - \frac{1}{R + \Delta R - 2 R_t} \right) \ \ldots \ 4.3.6$$

where $n(x)$ and $\mu(x)$ are average values of the carrier density and mobility respectively in the incremental thickness Δx_d.

The values of $n(x)$ and Δx_d could be obtained from C–V measurements but the equations can be rearranged to enable the mobility to be evaluated directly from the conductance and capacitance measurements (Pals 1970).

Equation 4.3.6 can be written as

$$\mu(x) = {}^{\Delta G}s/\Delta Q_s \qquad\qquad \dots\dots 4.3.7$$

where $\Delta Q_s = q\, n(x)\, \Delta x_d$ is the incremental change in charge per unit area in the channel and is therefore also equal to $C_d\, \Delta V$ where C_d is the depletion capacitance per unit area.

Hence

$$\mu(x) = {}^{\Delta G}s/C_d\, \Delta V$$

$$= \left(\frac{\ell}{\omega\, C_d\, \Delta V}\right)\left(\frac{1}{R - 2\, R_t} - \frac{1}{R + \Delta R - 2\, R_t}\right) \qquad \dots\dots 4.3.8$$

and this equation can be used to determine the mobility profile from the conductance and capacitance measurements.

It is shown below that large errors can arise if the terminal resistances are neglected although they have been kept to the minimum possible technological value by having only 10 μm gaps between the gate diffusion and the source and drain diffusions.

If the terminal resistances are neglected the maximum error in the mobility occurs when the channel resistance is at its minimum value. Let R_{CHM} = minimum value of channel resistance; then $R_M = R_{CHM} + 2R_t$ is the minimum value of the total source to drain resistance and also let $f = R_M/R_{CHM}$. The true value of the mobility, μ_t, is given by

$$\mu_t = \left(\frac{\ell}{\omega\,q}\right)\left(\frac{1}{n(x)\,\Delta x_d}\right)\left(\frac{1}{R_{CHM}} - \frac{1}{R_{CHM} + \Delta R_{CHM}}\right)$$

$$\approx \left(\frac{\ell}{\omega\,q}\right)\left(\frac{1}{n(x)\,\Delta x_d}\right)\left(\frac{\Delta R_{CHM}}{R^2_{CHM}}\right) \qquad \ldots\ldots 4.3.9$$

if $\Delta R_{CHM} \ll R_{CHM}$

However the measured value of the mobility, μ_m, when the terminal resistances are neglected is given by

$$\mu_m = \left(\frac{\ell}{\omega\,q}\right)\left(\frac{1}{n(x)\,\Delta x_d}\right)\left(\frac{1}{f\,R_{CHM}} - \frac{1}{f\,R_{CHM} + \Delta R_{CHM}}\right)$$

$$\approx \left(\frac{\ell}{\omega\,q}\right)\left(\frac{1}{n(x)\,\Delta x_d}\right)\left(\frac{\Delta R_{CHM}}{f^2\,R^2_{CHM}}\right) \qquad \ldots\ldots 4.3.10$$

if $\Delta R_{CHM} \ll f \cdot R_{CHM}$

Hence $\mu_m/\mu_t = 1/f^2$ $\qquad\qquad\qquad\qquad$ 4.3.11

For a 2 µm thick film with a gate junction depth of 0.5 µm the maximum value

of f for the device geometry shown in Figure 4.1.1(c), (52 µm of channel

length and 72 µm total source to drain length), is 1.29. From equation 4.3.11

it follows that the true mobility is nearly 70 per cent greater than the

value calculated neglecting terminal resistance. Even when the depletion

region has been driven to within 0.5 µm of the silicon-sapphire interface

the true mobility is nearly 20% greater than the calculated value. From

the above analysis it is clear that the corrections for terminal resistances

should always be made and it should also be noted that any misalignment

of the gate diffusion with respect to the source and drain diffusions

which gives rise to a redistribution of the terminal resistances does not

affect the correction procedure.

4.4 Lifetime Measurements from an Analysis of P-N Junction Reverse-
Bias Currents

In this section it is shown that the current flowing in a reverse-
biased p-n junction arises predominantly from electron-hole pair
generation in the depletion region and that an analysis of the
voltage dependence of the current enables a lifetime profile to be
determined.

The four basic emission and capture processes which control the
generation and recombination of charge in all regions of the
semiconductor are shown in Figure 4.4.1 and the transition
probabilities associated with these processes were given in
equations 4.2.5, 4.2.6, 4.2.7 and 4.2.8.

The transition rates for the electron emission and hole capture
processes are the product of the probability and the density of
occupied centres whereas the rates for the electron capture and
hole emission processes are the product of the probability and
the density of unoccupied centres. Therefore if f is the
occupancy factor of the recombination centres it follows that

a.   the rate of electron capture from the conduction band
is

$$r_a = v_{th} \, \sigma_n \, n_o \, N_T \, (1-f) \qquad \ldots\ldots 4.4.1$$

b.   the rate of electron emission into the conduction band
is

$$r_b = v_{th} \, \sigma_n \, n_1 \, N_T \, f \qquad \ldots\ldots 4.4.2$$
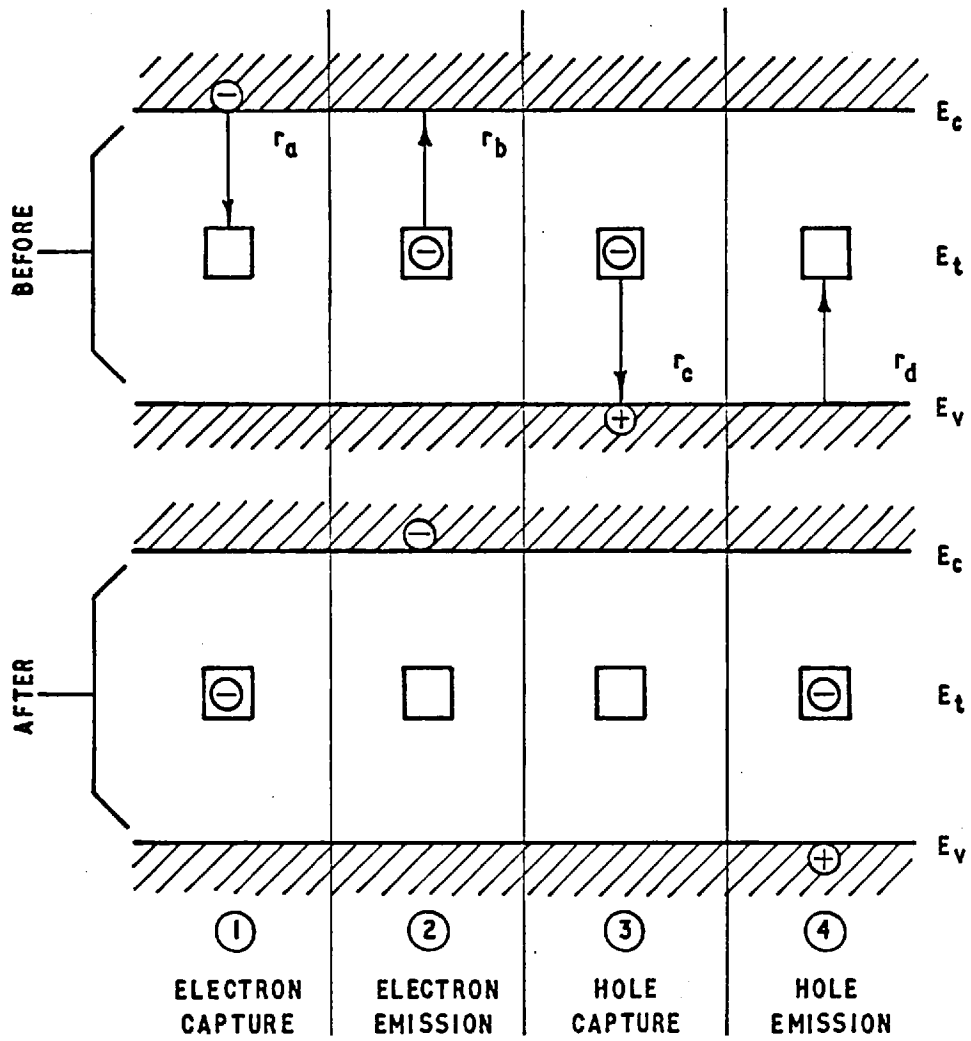
FIG. 4.4.1. TRANSITION PROCESSES INVOLVING A DEEP ACCEPTOR LEVEL

c.   the rate of hole capture from the valence band is

$$r_c = v_{th} \sigma_p p_o N_T f \qquad \ldots\ldots 4.4.3$$

and d.   the rate of hole emission into the valence band is

$$r_d = v_{th} \sigma_p p_1 N_T (1-f) \qquad \ldots\ldots 4.4.4$$

In equilibrium conditions f is given by the familiar Fermi-Dirac function

$$f = \frac{1}{1 + \exp\left(\frac{E_t - E_F}{kT}\right)}$$

but in the non-equilibrium steady-state conditions which exist in a reverse-biased p-n junction the value of f must be derived from the equation

$$r_b + r_c = r_a + r_d \qquad \ldots\ldots 4.4.5$$

which expresses the time independence of the charge in the generation-recombination centres.   This gives

$$f = \frac{\sigma_n n_o + \sigma_p p_1}{\sigma_n (n_o + n_1) + \sigma_p (p_o + p_1)} \qquad \ldots\ldots 4.4.6$$

and consequently the net recombination rate is given by

$$U = r_a - r_b = r_c - r_d$$

$$= \frac{v_{th} \, N_T \, \sigma_n \, \sigma_p \, (n_o \, p_o - n_i^2)}{\sigma_n \, (n_o + n_1) + \sigma_p \, (p_o + p_1)} \qquad \ldots\ldots 4.4.7$$

This equation can now be used to calculate the three components which contribute to the current flowing through a reverse-biased p-n junction at relatively low values of applied bias. Each component originates from a different part of the device structure and they are

1. 'Generation Current', which results from generation in the depletion region,

2. 'Diffusion Current', which results from generation in the neutral region within a minority carrier diffusion length of the depletion region and

3. 'Surface Current', which results from generation at the silicon-silicon dioxide interface.

Each contribution is analysed separately.

GENERATION CURRENT

When calculating the generation current it has been customary in the literature, (eg Sah, Noyce and Shockley 1957 and p 102 of Sze 1969), to assume that the density of free carriers, (n_o and p_o), is sufficiently small at all points within the depletion region to

be neglected and hence that the generation rate is constant across the depletion region at a value given by

$$-U_{max} = \frac{v_{th} \, N_T \, \sigma_n \, \sigma_p \, n_i}{\sigma_n \exp\left(\dfrac{E_t - E_i}{kT}\right) + \sigma_p \exp\left(\dfrac{E_i - E_t}{kT}\right)} \qquad \ldots\ldots \, 4.4.8$$

The neglecting of the $n_o \, p_o$ term in the numerator is certainly valid for applied reverse-bias voltages in excess of 0.1 volts since $n_o \, p_o = n_i^2 \exp\left(-\dfrac{qV}{kT}\right)$ but it can be seen from Figure 4.2.1(d) that although equation 4.4.8 is valid over most of the depletion layer there is a sizeable region adjacent to the neutral n-type material where the term n_o is larger than n_1 and p_1. In this region it follows from equation 4.4.7 that the generation rate falls very rapidly. A similar region exists close to the neutral p+-type material but the boron doping level is sufficiently high in the devices used to ensure that the width of that region is not significant.

The problem of calculating the 'generation current' now reduces to one of determining the value of x_t and substituting it in the equation

$$I_{gen} = q \, U_{max} \, x_t \, A_J \qquad \ldots\ldots \, 4.4.9$$

where A_J is the cross-sectional area of the p-n junction.

Since the energy level of the generation-recombination centre is not known accurately the analysis given by Calzolari and Graffi (1972) for a recombination centre at the intrinsic level is considered first. This analysis showed that

$$x_t = \left(\frac{2 \, \varepsilon_{si} \, kT}{q^2} \right)^{\frac{1}{2}} \left[\left(\frac{1}{N_D} \ln \frac{N_D}{n_i} + \frac{q \, V}{kT \, N_D} \right)^{\frac{1}{2}} \right.$$

$$\left. - \frac{1}{N_D} \ln \frac{N_D}{n_i} \right)^{\frac{1}{2}} \right] \qquad \ldots\ldots 4.4.10$$

and substitution of $E_t = E_i$ in equation 4.4.8 gives

$$- U_{max} = \frac{v_{th} \, N_T \, \sigma_n \, \sigma_p \, n_i}{\sigma_n + \sigma_p} = \frac{n_i}{2 \, \tau_e} \qquad \ldots\ldots 4.4.11$$

where

$$\tau_e = \frac{\sigma_n + \sigma_p}{2 \, v_{th} \, N_T \, \sigma_n \, \sigma_p} \qquad \ldots\ldots 4.4.12$$

is called the 'effective lifetime' and is equal to $1/v_{th} \, N_T \, \sigma$ when $\sigma_n = \sigma_p = \sigma$. An 'effective lifetime' profile can be determined from equations 4.4.9 and 4.4.10 if the substrate doping level is known and the generation current can be measured as a function of the reverse-bias voltage.

The Calzolari and Graffi analysis assumes that the quasi-Fermi potentials are constant to a point beyond their intersection with the intrinsic potential and although this is generally true for bulk silicon junctions it requires verification for the case of silicon on sapphire junctions since they pass much larger reverse-bias currents. An upper limit can be calculated for the rate of change of the quasi-Fermi potential in that part of the space charge region where $n_o > n_i$ by using the following equation for the electron current density, (see for example p 97 of Sze 1969):-

$$J_n = -q \mu_n n_o \frac{d\phi_n}{dx} \qquad \ldots\ldots 4.4.13$$

where $\phi_n = - E_{FN}/q$ is the quasi Fermi potential for electrons and μ_n is the electron mobility.

Now when $n_o > n_i$ it follows from equation 4.4.13 that

$$-\left(\frac{d\phi_n}{dx}\right) < \frac{J_{rb}}{q \mu_n n_i} \qquad \ldots\ldots 4.4.14$$

where J_{rb} is the total junction current density which is slightly greater than J_n. Since the experimental value of J_{rb} is never greater than 5×10^{-5} Acm^{-2} and μ_n is never less than 50 cm^2 V^{-1} s^{-1} it follows from the relationship 4.4.14 that 450 V cm^{-1} is the upper limit for $-(d\phi_n/dx)$ in the region where $n_o > n_i$, and using a maximum value of 5000Å for the width of this region the maximum change in ϕ_n within the region is 22.5 mV. It then follows from

an examination of Figure 4.2.1(b) that the decrease in intrinsic

potential between the neutral n-type region and the point of

intersection of the n-type quasi-Fermi level and the intrinsic level

is changed by 22.5 meV in an amount equal to $(kT/q) \ln (n_o/n_i)$

because of the existence of a small gradient in the n-type quasi-

Fermi level. At a doping level of 5×10^{15} in the n-type region

this corresponds to an additional 7% change in the intrinsic level.

The potential distribution in the depletion region can be obtained

from a solution of Poissons equation(p 161 of Grove 1967) and it

is given by

$$\phi_i \quad = \quad \frac{- q N_D}{2 \varepsilon_s} \quad (x_d - x)^2 \qquad \qquad \ldots \ldots 4.4.15$$

$$\text{Hence} \quad \frac{d\phi_i}{\phi_i} \quad = \quad \frac{2 d (x_d - x)}{x_d - x} \qquad \qquad \ldots \ldots 4.4.16$$

Therefore the 7% difference in ϕ_i gives rise to only a 3.5% error

in $x_d - x_t$ and hence it will not significantly influence the

'effective lifetime' calculation.

The above analysis has assumed that the generation-recombination

centre is at the intrinsic level. Although experimental analysis

of the frequency response of the centres in section 4.5 will show

that they are indeed close to the intrinsic level even very small

discrepancies will give rise to large errors in the effective

lifetime calculations. If $E_t > E_i$ by more than a few kT and σ_p

is not very much larger than σ_n then equation 4.4.8 reduces to

$$U_{max} = \frac{v_{th} \, N_T \, \sigma_p \, n_i}{\exp\left(\dfrac{E_t - E_i}{kT}\right)} \qquad \qquad \ldots \ldots \, 4.4.17$$

and hence

$$\tau_e = \frac{\exp\left(\dfrac{E_t - E_i}{kT}\right)}{2 \, v_{th} \, N_T \, \sigma_p} \qquad \qquad \ldots \ldots \, 4.4.18$$

If $E_t - E_i = 0.09\,\text{eV}$, which appears to be possible from the frequency response measurements, then τ_e is equal to $16/v_{th} \, N_T \, \sigma_p$ which is much larger than the true minority carrier lifetime $\tau_p = 1/(v_{th} \, N_T \, \sigma_p)$. It appears therefore that any analysis of the reverse bias current cannot be expected to give much better than an order of magnitude measure of the minority carrier lifetime unless the generation-recombination centres have a single energy level whose position is very accurately known. However, although the absolute value of the lifetime cannot be measured accurately the 'effective lifetime' measurement is a good indicator of the rate of change of N_T through the film because the energy level and the capture cross-section of the generation-recombination centres are unlikely to vary. This information is significant because a rapid variation in the dislocation density through the films has been reported by Linnington (1970) and it is quite probable that the generation-recombination centres are associated with dislocations.

It is also assumed in the Calzolari and Graffi analysis that the donor density is constant and it is shown in section 4.6 that this is a fairly good assumption through most regions in the films although the effective lifetime measurement procedure described in that section does make some allowance for the small variations that do exist.

Having shown how an approximate lifetime profile can be calculated from measurements of the gate capacitance and the generation current it remains to be demonstrated that for silicon on sapphire junctions the diffusion current and the surface current are normally much smaller than the generation current and therefore the reverse-bias current to the gate is a measure of the generation current.

DIFFUSION CURRENT

Any minority carriers which diffuse to the edge of the depletion region from the neighbouring neutral regions will be swept across the junction by the applied field. In order to calculate this diffusion current the following steady state diffusion equations must be solved for both sides of the junction:-

$$\text{p-side:-}\quad D_n \frac{d^2 n_p}{dx^2} - \frac{n_p - n_{po}}{\tau_n} = 0 \qquad \ldots\ldots 4.4.19$$

$$\text{n-side:-}\quad D_p \frac{d^2 p_n}{dx^2} - \frac{p_n - p_{no}}{\tau_p} = 0 \qquad \ldots\ldots 4.4.20$$

where D_n and D_p are diffusion constants for electrons and holes

123

respectively. n_p and p_n are minority carrier electron and hole densities respectively and the subscript 0 denotes the equilibrium value.

When solving equations 4.4.19 and 4.4.20 different coordinate axes are chosen. In both cases the adjacent depletion layer edge is taken as $x = 0$ and the positive x-axis is taken pointing away from the depletion layer. Boundary conditions for the equations are $n_p (\infty) = n_{po}$ and $p_n (\infty) = p_{no}$, (since equilibrium conditions are valid in the bulk region), and $n_p (0) \simeq p_n (0) \simeq 0$, (since minority carriers are swept away by the field at the depletion layer edge if $V_R \gg kT/q$). The solutions to equations 4.4.19 and 4.4.20 are

$$n_p (x) = n_{po} (1 - e^{-x/L_n}) \qquad \ldots \ldots 4.4.21$$

$$\text{and} \quad p_n (x) = p_{no} (1 - e^{-x/L_p}) \qquad \ldots \ldots 4.4.22$$

where $L_n = (D_n \tau_n)^{\frac{1}{2}}$ and $L_p = (D_p \tau_p)^{\frac{1}{2}}$ are diffusion lengths for electrons and holes respectively and it follows from equations 4.4.21 and 4.4.22 that

$$I_{diff,n} = (-q) \left(- D_n \left. \frac{dn_p}{dx} \right|_{x=0} \right) A_J$$

$$= q D_n \frac{n_i^2}{N_A L_n} A_J \qquad \ldots \ldots 4.4.23$$

124

and

$$I_{diff,p} = (+q) \left(D_p \left. \frac{dp_n}{dx} \right|_{x=0} \right) A_J$$

$$= q D_p \frac{n_i^2}{N_D L_p} A_J \qquad \ldots \ldots 4.4.24$$

For a p^+-n junction $N_A \gg N_D$ and therefore $I_{diff,n}$ can be neglected compared with $I_{diff,p}$. Comparing the diffusion current in equation 4.4.24 with the generation current in equation 4.4.9 gives

$$\frac{I_{diff}}{I_{gen}} = \frac{D_p n_i^2}{N_D L_p} \left/ \frac{n_i x_t}{2 \tau_e} \right.$$

$$= 2 \frac{n_i}{N_D} \cdot \frac{(D_p \tau_p)^{\frac{1}{2}}}{x_t} \cdot \frac{\tau_e}{\tau_p} \qquad \ldots \ldots 4.4.25$$

In the silicon-on-sapphire layers investigated in this work the maximum values for D_p and τ_p were 6.5 cm^2 s^{-1} and 10^{-8} s respectively and the minimum value for N_D was 7.5 x 10^{14} cm^{-3}. The value of x_t due to the built-in potential was always greater than 5 x 10^{-6} cm and it has been shown above that τ_e/τ_p is 16

for the generation-recombination level of $E_t - E_i = 0.09eV$ calculated from the frequency response measurements. If these parameter values are substituted in equation 4.4.25 it follows that I_{diff}/I_{gen} is less than 3×10^{-2} and hence that the diffusion current does not contribute significantly to the measured reverse-bias currents. It must be remembered that this conclusion is a consequence of the proximity of the generation-recombination centre energy level to the intrinsic level.

The magnitude of the surface current is now considered.

SURFACE CURRENT

The reverse-bias current due to electron-hole pair generation at surface generation-recombination centres with energy levels at the intrinsic level is given by the equation

$$I_{surf} = \frac{1}{2} q \, n_i \, s_o \, A_s \qquad\qquad \ldots\ldots 4.4.26$$

where $s_o = \sigma_{ss} \, v_{th} \, N_{st}$ = surface recombination velocity,

σ_{ss} = capture cross-section for surface states

and A_s = area of the depletion layer which intersects the surface.

The measurements were all made with the source, drain and MOS gate shorted together as shown in Figure 4.4.2 and the positive charge which is always present in the thermal silicon dioxide will tend to reduce the width of the depletion layer at the $Si - SiO_2$ interface. If the width of the depletion layer in the bulk is taken as an upper limit for the width at the interface then the
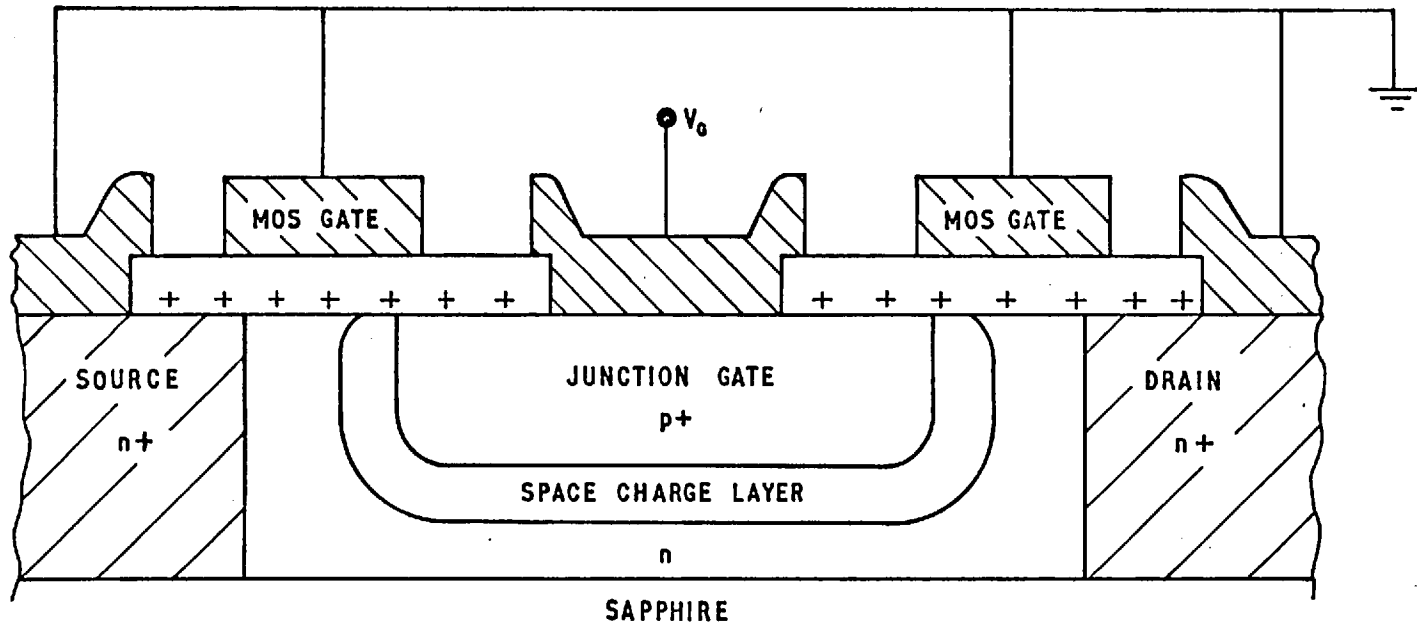
FIG. 4.4.2. SCHEMATIC DRAWING OF DEVICE DURING REVERSE-BIAS
CURRENT MEASUREMENTS

maximum value of A_s for the devices used in this work is 7.5×10^{-5} cm^2. The oxide technology used for preparing the experimental devices is capable of maintaining a surface state density of less than 2×10^{11} cm^{-2}. This taken together with the value of 2×10^{-16} cm^2 for the capture cross section of surface states reported by Grove and Fitzgerald (1966) gives a value of S_o equal to 916 cms^{-1} and consequently a surface current of 7.7×10^{-11} A, which is well below the measured reverse-bias currents. Therefore the measured reverse-bias current is predominantly generation current and can be used to determine an 'effective lifetime' profile.

4.5 Experimental Results on p Type Vacuum Evaporated Films

The first junction field-effect devices to be investigated were
fabricated from layers prepared by evaporation from a water-cooled
electron bombardment source in a stainless steel ultra-high vacuum
system (Lawson and Jefkins 1970). These 1.5 - 2.0 μm thick p-type
layers were of (100) orientation grown at $1000^{\circ}C$ on $(1\bar{1}02)$ sapphire
substrates at a rate of approximately 10^{-8} ms^{-1}. The layers were not
intentionally doped and the average acceptor concentrations due
to autodoping were in the range 1-2 x 10^{16} cm^{-3}. Average majority
carrier mobilities ranged from 100 - 150 cm^2 V^{-1} s^{-1}

The p-channel devices made from these films were of the type shown in
Figure 4.1.1(a). The processing was generally similar to that
described in Appendix 3 for the n-channel components although,
naturally, the source-drain and gate diffusants were interchanged
and the shallow phosphorus gate diffusion was for only 5 minutes at $1000^{\circ}C$
followed by a 10 minute oxidising drive-in at $950^{\circ}C$. The device
properties varied quite extensively from layer to layer and even across
a single layer, with most of the differences being attributable
to variations in the gate junction depth. However the general
form of the calculated carrier concentration and mobility profiles
were the same for all the devices although the magnitudes of the
values were somewhat variable. The results for slice 270/1 are
presented below in some detail because they are characteristic of the
range of properties which were observed on a single slice.
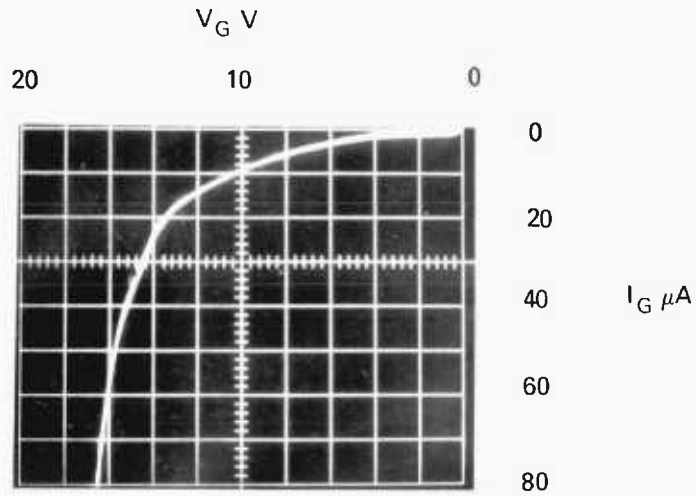
Large variations in the channel depth of the devices fabricated on slice 270/1
were indicated by the fact that the drain current measured at V_G = 0 volts
and V_D = 10 volts varied from less than 1μA to 520μA. At a carrier

concentration of 2×10^{16} cm^{-3}, which will be shown to be the value

adjacent to the sapphire interface, the width of the space charge

region adjacent to the gate at zero applied voltage is 0.22μm and the

drain current measurements give grounds for inferring that the

channel depth on some of the devices is less than this.

Now the silicon layer on slice 270/1 was 1.2μm thick at the end

of processing and the depth of the p-n junction in the bulk silicon

test-piece which accompanied it during the phosphorus gate diffusion

and subsequent drive-in was 0.5μm.

Consequently if the silicon-on-sapphire had the same diffusion

constant as the bulk silicon, the channel depth would be 0.7 μm but a

diffusion constant four times larger than bulk silicon, which is

within the range of values reported by other workers (Ross and

Warfield 1959 and Dumin and Silver 1968), would give a junction

depth of 1.0μm and account for the channel being turned off at zero

gate-bias.

Because the ungated junction type of structure was used for the 270/1

devices the surface ion migration phenomena referred to in section 4.1

resulted in relatively large currents flowing to the reverse-biased gate

and a typical characteristic is shown in Figure 4.5.1 (a). Although this

behaviour makes an analysis of the lifetime profile in the manner

described in section 4.4 impossible the carrier concentration

profile can be calculated from the capacitance results in the normal

manner and a mobility profile can be determined from the drain

conductance, at least for the range of reverse-bias voltages where

the channel current is much larger than the gate curren t . The drain

characteristic shown in Figure 4.5.1 (b) shows that on the

higher current devices there is indeed a substantial range of gate

V_G V

20   10   0

0

20

40    $I_G \, \mu A$

60

80

(a)

$I_D \, \mu A$

100

0

0   − 10   − 20

V_D V

V_G = 0 to 1.8V in 0.2V steps
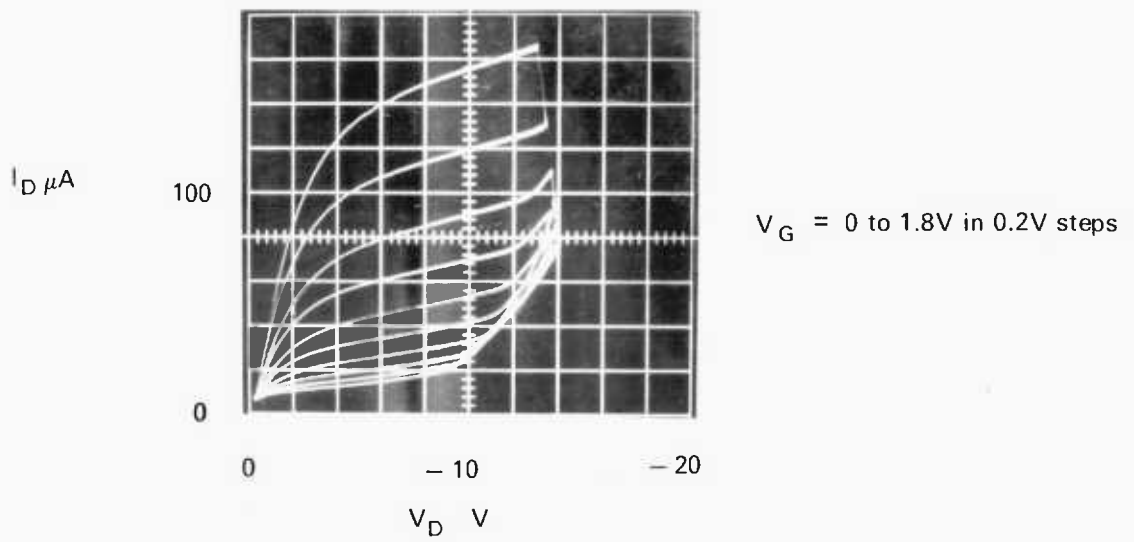
(b)

Figure 4.5.1.   (a) Reverse — biased gate I — V characteristic
for p — channel junction field-effect transistor

(b) Drain characteristics for p — channel
field-effect transistor

131

voltage over which the gate current is negligible in comparison with the channel current.

The gate capacitance measurements are made by probing the aluminium bonding pads on the device with steel needles and thereby connecting the gate to the high input terminal of a Model 71A Boonton Capacitance Meter and the source and drain to the low input terminal. The capacitance measurement with the Model 71A is not influenced by the substantial resistive component in the gate impedance because a phase sensitive detector is employed with a reference signal 90° out of phase with the signal applied to the device under test. The frequency used for the measurement is 1 MHz and the maximum amplitude of the signal appearing across the device under test is 15 mV Bias Voltage for the measurement is obtained from a calibrated voltage source with an accuracy of 0.05% and the output of the capacitance meter is measured on a digital voltmeter with an accuracy of 0.01%.

The carrier concentration profile is calculated from the capacitance measurements using equation 4.2.12 which includes a dC/dV term in the denominator. Since dC/dV is calculated using capacitance values which differ by only a few per cent and the absolute accuracy of the Model 71A is only of the order of 1 percent the first reaction may be to assume that very large errors will result from the use of equation 4.2.12, but it is non-linearities in the capacitance measurement which give rise to errors and not its absolute accuracy.

Since the non-linearity of the instrument is not known the carrier concentration profile was measured for a device fabricated in uniform substrate material and the results varied by less than 7%.

When equation (4.2.12) was applied directly to analyse the
capacitance measurements of one of the devices with the larger drain
currents the charge density profile indicated by the circles in
Figure 4.5.2 was obtained. However as the depletion region is driven
back towards the sapphire interface significant errors arise in the
capacitance measurement due to attenuation of the measurement signal
in the resistive silicon lower electrode of the capacitor. A
correction for this can be calculated by treating the capacitor as a
transmission line with no series inductance and no shunt capacitance.
This involves substituting the experimental conductance measurements
(made using the simple measurement circuit shown in Figure 4.5.3),
into the following two terminal impedance equation derived
by Hager (1959).

$$Z = \frac{R \coth \frac{1}{2} (j_\omega RC)^{\frac{1}{2}}}{2 (j_\omega RC)^{\frac{1}{2}}} \qquad \ldots\ldots \quad 4.5.1$$

where C is the true low frequency capacitance of the junction and
R is the resistance between the source and the drain. The charge
density profile calculated from the corrected capacitance
values is shown by the crosses in Figure 4.5.2

The evaluation of the net donor density from the carrier concentration
results using equation (4.2.16) involves a second order differentiation
and it is likely to be very inaccurate if done numerically from
the experim ental capacitance results. Therefore it was decided to
compute a polynomial best fit to the calculated carrier concentration
results and a second order polynomial was generally found to fit
quite well, as shown in Figure 4.5.2. The polynomial was then
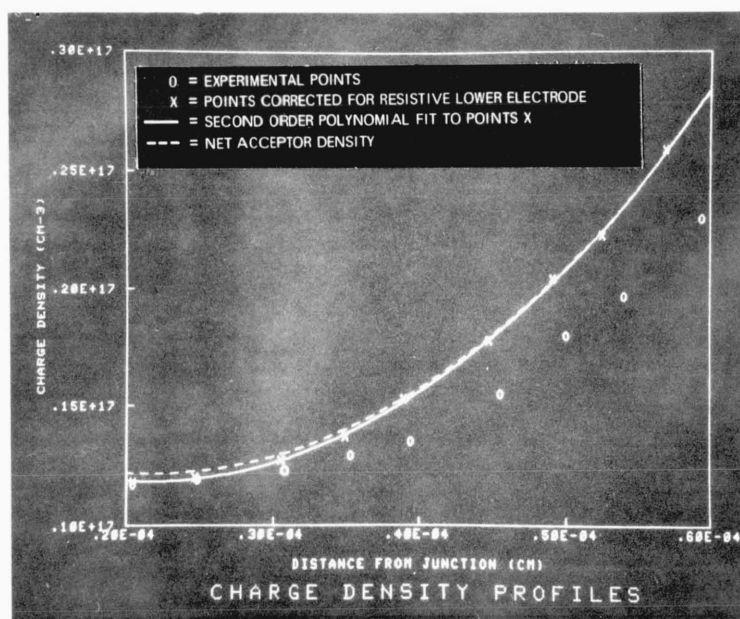differentiated in order to evaluate equation (4.2.16) and the

Figure 4.5.2. *Carrier Concentration and Net Acceptor Density Profiles for P – Type Vacuum Evaporated Films*

calculated net acceptor density which is also shown in Figure 4.5.2(a) is not significantly different from the majority carrier concentration.

The results shown in Figure 4.5.2 clearly indicate an increase in the net acceptor density as the sapphire interface is approached which is probably associated with an out diffusion of aluminium from the substrate, (Dumin and Robinson 1966). However since the density of deep donor states is also known to increase as the sapphire interface is approached, (Dumin 1970), the results shown in Figure 4.5.2 probably underestimate the magnitude of the autodoping effect and must be interpreted simply as a lower limit for the outdiffusion profile of shallow aluminium acceptors.

The derived carrier concentration profile was then used together with the field-effect conductivity results and a resistivity measurement on an adjacent resistor to calculate the mobility profile using equation (4.3.6). All the conductivity measurements were made using the circuit in Figure 4.5.3 at a drain voltage of 100 mV which ensures very little variation in depletion depth along the channel and even at a carrier concentration of 5×10^{15} cm^{-3} the calculated conductivity error is less than 5% up to within 0.15µm of the sapphire interface. Measurements were restricted to those with gate voltages in the range where the gate current, I_G, was less than 10% of the source current, I_S. This prevented results from being obtained for the region within 0.15µm of the sapphire interface. Gate currents, I_G, are only significant in the range $V_G \gg V_D$ and the channel current, I_C is given by the equation

$$I_C = I_S + I_G/2 \qquad\qquad \ldots\ldots\ldots 4.5.2$$
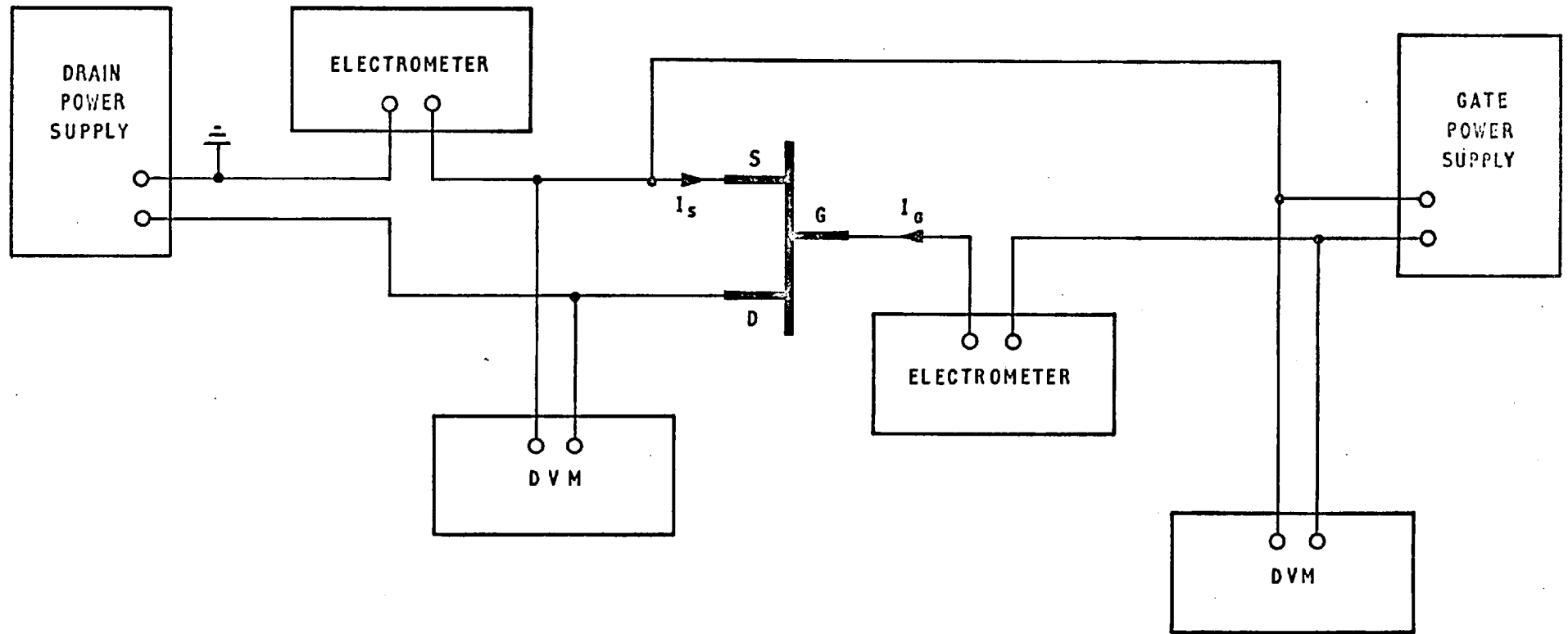
FIG. 4.5.3. MEASUREMENT CIRCUIT FOR CHANNEL CONDUCTIVITY

on the assumption that the gate current divides equally between the
source and the drain. The calculated mobility profile for the device which
gave the carrier concentration profile in Figure 4.5.2 is shown in
Figure 4.5.4. Since these mobility values are much less than
the bulk value and the piezo resistance effect is small in the (100)
orientation the reductions are probably a consequence of
additional ionic and/or space-charge scattering mechanisms associated
with the defect centres which increase in density as the interface
is approached.

The absolute value of the mobility is a little less than that observed
by Ipri (1972) for films grown by the pyrolysis of silane but the rate
of decrease as the sapphire interface is approached is very similar.

SAMPLE No. 270/1 (R 4C3)

138
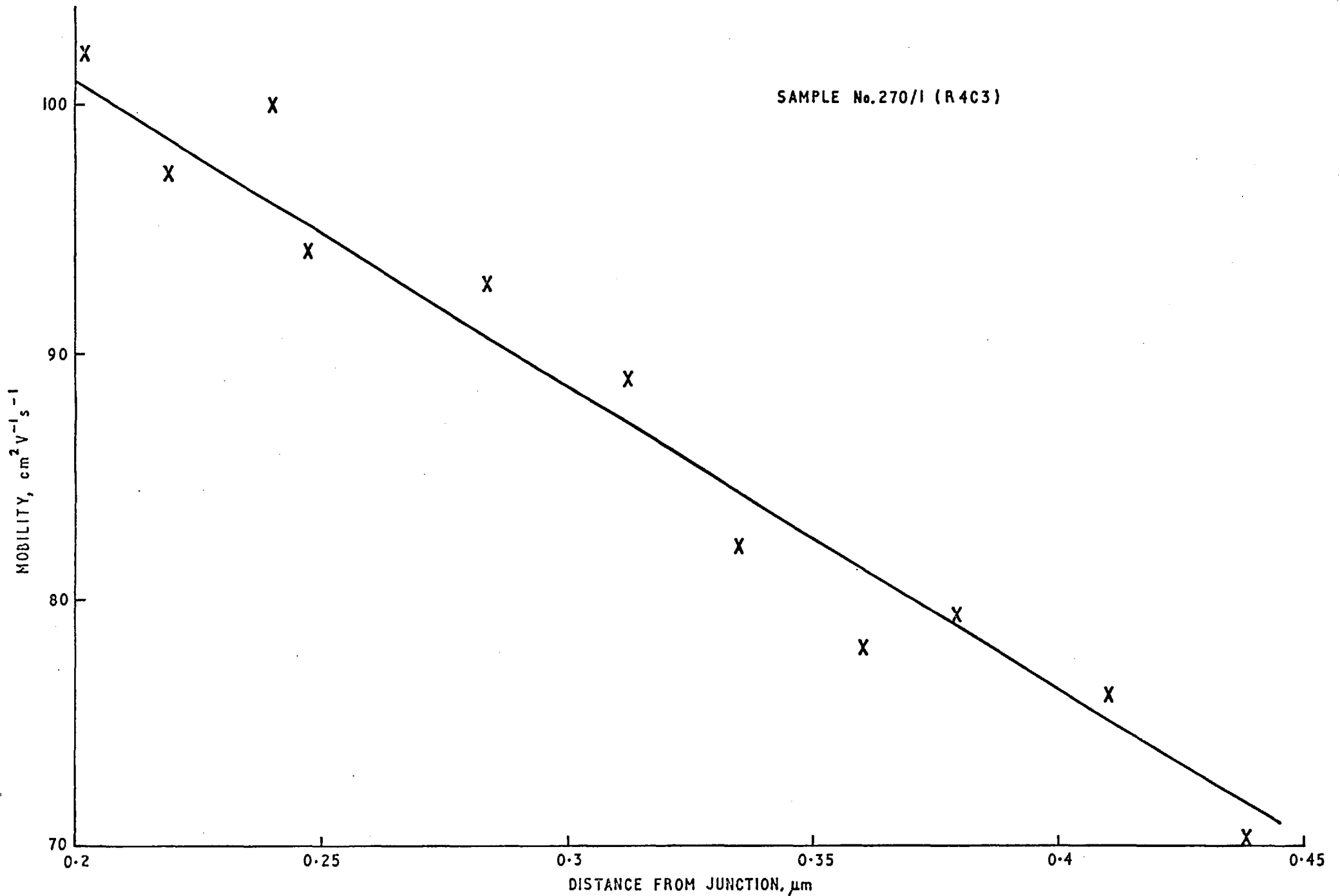
FIG 4.5.4.   MOBILITY PROFILE FOR P-TYPE VACUUM EVAPORATED FILM

4.6 Experimental Results on N-Type Films Grown by Pyrolysis of Silane

Junction field effect structures both with and without insulated gates over the gate junction perimeter (as shown in Figures 4.1.1(b) and 4.1.1(a) respectively) have been fabricated in 1.5μm to 2.0μm thick n-type films grown by the thermal decomposition of silane at 1050°C *
The films which were phosphorus doped and of (100) orientation on (1$\bar{1}$02) sapphire were annealed prior to device fabrication for 6 hours at 1100°C in an 80% nitrogen/20% oxygen mixture in order to precipitate any aluminium incorporated by autodoping from the substrate. The mean carrier concentrations after annealing were in the range 4×10^{15} cm^{-3} to 2×10^{16} cm^{-3} as measured by the van der Pauw (1958) technique and the mean Hall mobilities were in the range 110 cm^2 V^{-1}s^{-1} to 290 cm^2 V^{-1}s^{-1}.

The test devices were fabricated according to the procedure specified in Appendix 3. Appreciable surface leakage currents were again observed in many of the devices constructed according to the design in Figure 4.1.1(a). This is probably because the net positive charge in the oxide causes the surface of the n-type region between the gate and the drain to be accumulated thereby giving rise to interband tunneling across the reverse biased drain junction even at relatively low applied potentials. Since the few results obtained on the ungated test structures which did have relatively low leakage currents were broadly similar to those obtained on devices constructed according to the design in Figure 4.1.1(b)

* The films were purchased from the Allen Clark Research Centre of the Plessey Company.

a typical set of results derived from measurements on the latter structure will now be described in detail.

Since the layers used had a fairly wide range of resistivities and thicknesses the device properties again varied widely from slice to slice but the variations observed across a single slice were much less than those observed with the vacuum-evaporated p-channel slices and the following results for slice 586/1 are typical of those observed on other slices.

The thickness of layer 586/1 after processing was 1.37µm and the measured depth of the p-n junction formed in the p-type bulk silicon test piece of similar doping density which accompanied the silicon-on sapphire layer during the boron diffusion and 'drive-in' was 0.41µm, (this value is in good agreement with a theoretical junction depth calculated from the published diffusion constants). Since the diffusion constant for the silicon-on-sapphire is likely to exceed the bulk silicon value (Ross and Warfield 1969) the upper limit for the separation between the p-n junction and the silicon-on sapphire interface is 0.96µm. Measured values of the 'pinch-off' voltage, V_p', defined as the gate voltage required to reduce the drain current to 10^{-7} A at a drain voltage of 100 mV ranged from -0.91V to -1.24V and the values of drain current at V_{GS} = 0V and V_{DS} = 100 mV ranged from 1.53µA to 4.24µA.

A typical set of drain characteristics for one of the n channel devices on slice 586/1 is given in Figure 4.6.1(a)[+] and the corresponding gate

+ The first part of the characteristics were omitted simply to prevent excessive overexposure of the film in the region surrounding the origin.

I_S

V_{GS} = 0 TO 0·6V
IN 0·IV STEPS

V_{DS}

(a)

V_P^1

I_G
nA

V_G V

(b)

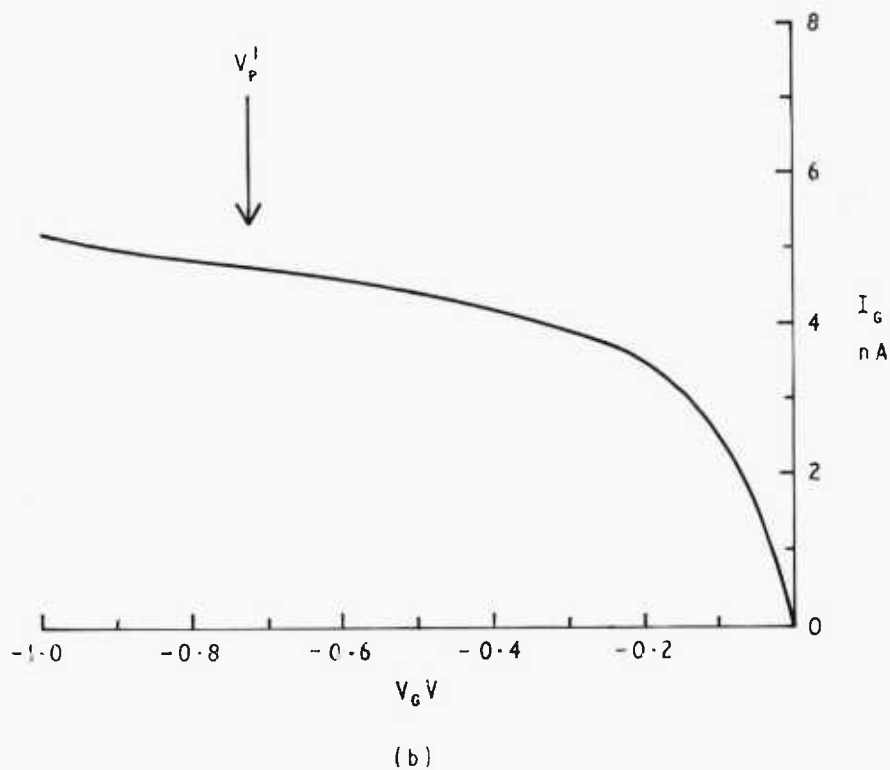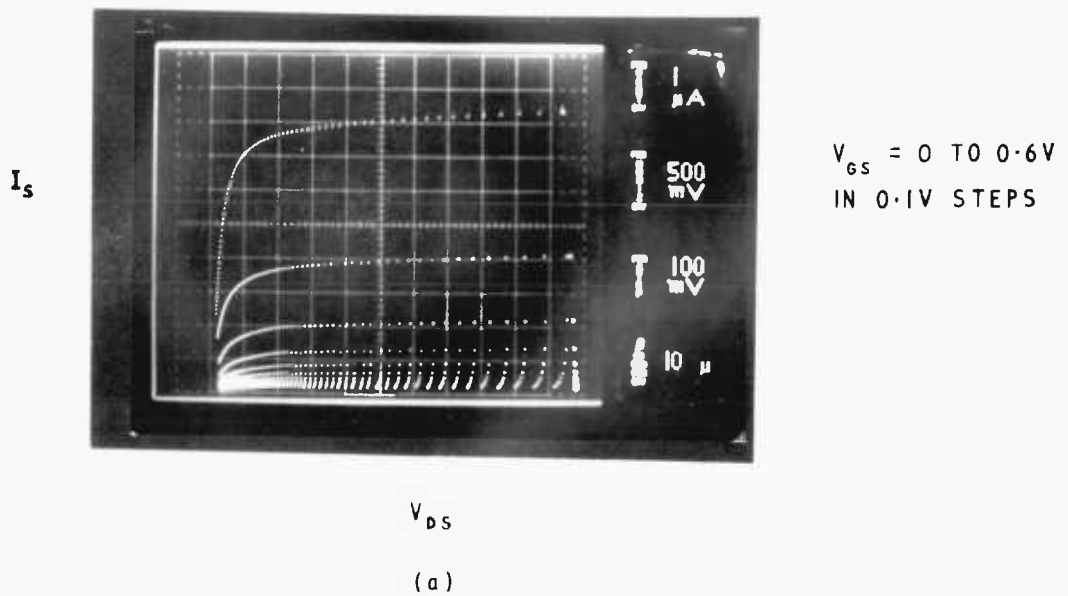FIG. 4. 6. I. (a) DRAIN CHARACTERISTICS OF N – CHANNEL
                    JUNCTION FIELD EFFECT STRUCTURE
              (b) GATE LEAKAGE CURRENT FOR N – CHANNEL
                    JUNCTION FIELD EFFECT STRUCTURE

141

junction leakage characteristic for the same device up to 'pinch-off'
is given in Figure 4.6.1(b). The early saturation in the drain
characteristics clearly shows the need to keep V_{DS} as low as possible
when making channel conductivity measurements. The lower-limit for V_{DS}
is determined by the condition that the channel current must be significally
greater than the gate leakage current. Since the gate leakage current
of all the devices used was less than 10^{-8}A at 'pinch-off' the use
of V_{DS} = 100 mV is a reasonable compromise because this ensures that
the channel current is at least an order of magnitude greater than
the gate leakage current if measurements are restricted to gate voltages
above 'pinch-off'. Using a doping density of 3.5 x 10^{-15}, (which
will be shown later to be the measured value for layer 586/1), the
'pinch-off' drain current of 10^{-7}A can be shown to correspond to the
depletion layer edge being within 0.03μm of the sapphire interface
if the interfacial mobility is 50 $cm^2 V^{-1} s^{-1}$ and within 0.15μm for an
interfacial mobility of 10 $cm^2 V^{-1} s^{-1}$. Note that as 'pinch-off'
is approached the saturation of the characteristic will cause the channel
conductivity to be underestimated.

All the conductivity measurements were taken with the insulated gate
electrode connected directly to the source. However to ensure that
surface currents were not playing a significant role in the
measurements the effect of changing the insulated gate electrode
potential over the range -10V to +_10V was examined but
changes in the drain current and the gate current were less than 1%.

An experimental transfer characteristic, measured at V_D = 100 mV is
given for one of the devices on the 586/1 layer in Figure 4.6.2. This
characteristic can be analysed using the drain current equation
(4.3.2) given in the conductivity mobility section and (4.3.2) can be
rearranged to give
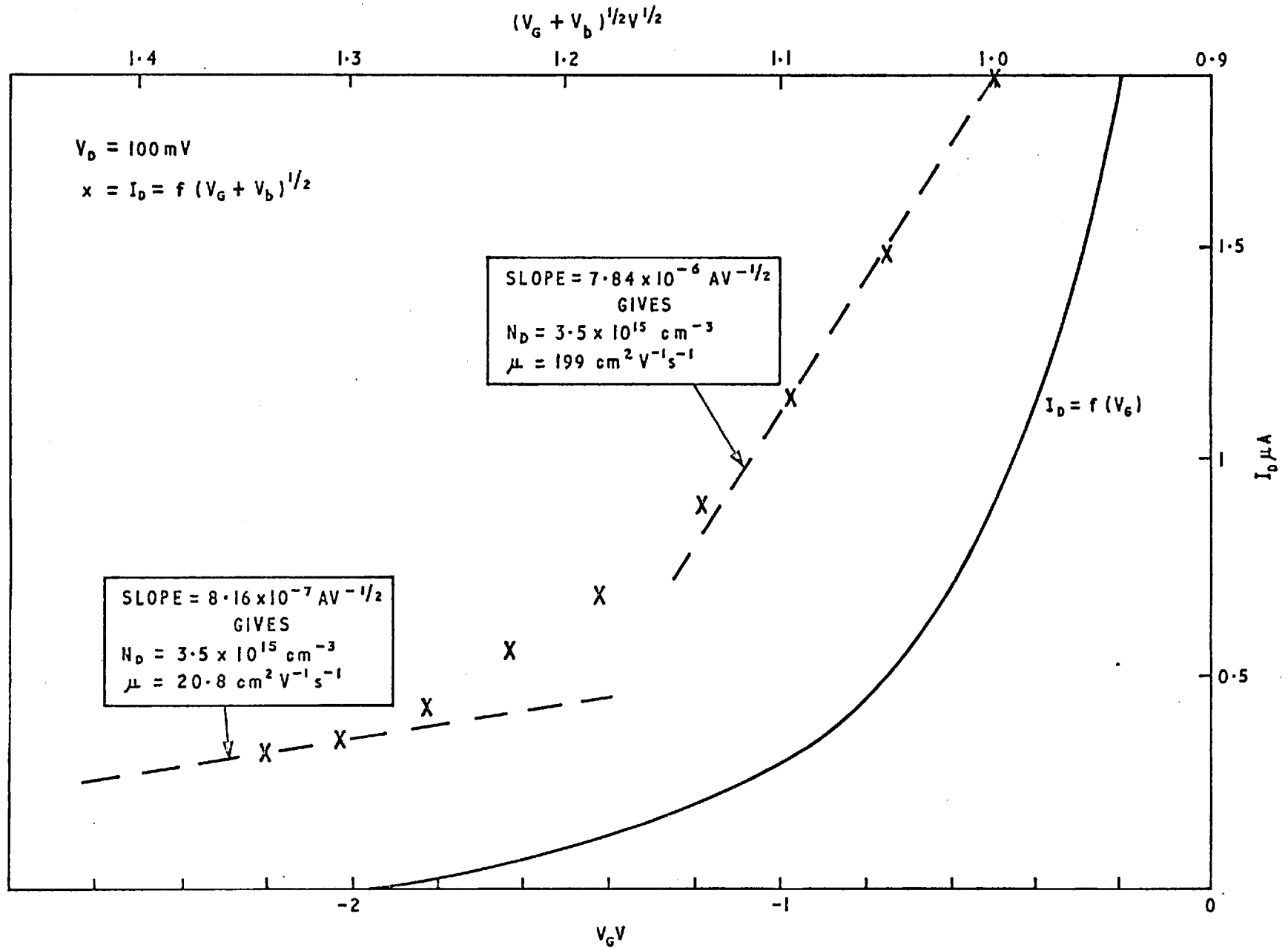
FIG. 4.6.2. JUNCTION FIELD EFFECT TRANSISTOR TRANSFER CHARACTERISTICS
AND PLOT OF I_D AS A FUNCTION OF $(V_G + V_b)^{1/2}$

143

$$I_D = g_{max} \left\{ V_D - \frac{2}{3(t-x_j)} \left(\frac{2\varepsilon_{si}}{qN_D} \right)^{\frac{1}{2}} \left(V_G + V_b \right)^{\frac{3}{2}} \left[\left(1 + \frac{V_D}{V_G + V_b} \right)^{\frac{3}{2}} -1 \right] \right\} \qquad \ldots\ldots 4.6.1$$

and by the use of a Taylor expansion series this reduces to

$$I_D = g_{max} \left\{ V_D - \left(\frac{2\varepsilon_{si}}{qN_D(t-x_j)^2} \right)^{\frac{1}{2}} \left(V_G + V_b \right)^{\frac{1}{2}} V_D \right\} \qquad \ldots\ldots\ldots 4.6.2$$

when $V_D \ll V_G + V_b - V_p$ where

$V_p = q N_D (t-x_j)^2 / 2\varepsilon_{si}$ is called the 'pinch-off' voltage of the

junction field-effect transistor and is the total gate voltage at which

the drain current reduces to zero. Equation (4.6.2) may be written as

$$I_D = \frac{w\mu qN_D(t-x_j)}{\ell} V_D \left[1 - \left(\frac{V_G + V_b}{V_p} \right)^{\frac{1}{2}} \right] \qquad \ldots\ldots 4.6.3$$

and if I_D is plotted against $(V_G + V_b)^{\frac{1}{2}}$ a straight line should be obtained

if the carrier concentration, N_D, and the mobility, μ are not functions

of position in the silicon layer. It is clearly apparent from the

plot of this type given in Figure 4.6.2. that N_D and μ are not both

independent of position. If a value of 3.5×10^{15} cm^{-3} is taken for N_D

then the slopes at the two extremes of the plot shown in Figure 4.6.2

correspond to mobilities of 20.8 cm^{2}V^{-1}s^{-1} and 199 cm^{2}V^{-1}s^{-1} and these

values give a fair indication of the steep mobility profile in

the layer which will be confirmed by the detailed measurements later

in this section.

Before proceeding to the detailed discussion of the carrier
concentration, mobility and lifetime profiles obtained from the
capacitance, conductivity and gate current results, mention must be
made of a very slow transient effect observed on these devices. The
normal procedure of probing a slice prior to making measurements involves
an illumination of the devices and it was noted that after the light
was switched off there was a slow reduction of drain current over a
period of approximately two hours. These transients, together with
similar effects observed in the dark when the device was turned
on, will be discussed in the next section. It is sufficient to
note here that the devices were always left for three hours to
stabilise after any illumination and that any transients observed after
changes in the junction-gate potential were also given sufficient time
to stabilise.

It was shown in section 4.2 how the carrier concentration can be
calculated from a measurement of the gate capacitance if the
measurement is made at a sufficiently high frequency to ensure that
the deep acceptor traps are unable to respond to the measurement
frequency. However the energy level of the traps was not known
with sufficient accuracy to enable a theoretical analysis to
demonstrate conclusively that 1 MHz is sufficiently high a frequency
to prevent the acceptors responding to the signal and therefore
experimental evidence is now presented in support of this hypothesis.
Since the frequency dispersion due to trapping is greatest at low
applied voltages the capacitance of the gate junction was measured
at zero applied voltage over the frequency range 5 kHz to 10 MHz.

145

Two capacitance bridges with overlapping frequency ranges were used, namely a Boonton Model 75 C from 5 kHz to 500 kHz and a Wayne-Kerr Model B801 from 100 kHz to 10 MHz.

A typical result is shown in Figure 4.6.3 and the absence of any frequency dependence from 25 kHz to 1 MHz indicates that there are few acceptor traps with time constants in this range. Taking an approximate value of 10^{-15} cm^{-3} for the capture cross section it can be calculated from equation (4.2.9) that this frequency range corresponds to a 'trap-free' energy range of 0.11 eV to 0.20 eV above the intrinsic energy level. For capture cross sections of 10^{-14}cm^{-3} and 10^{-16}cm^{-3} the range becomes 0.05eV to 0.14eV and 0.17eV to 0.26eV respectively. The increase in capacitance below 25kHz is believed to be due to the modulation of charge in traps fairly close to the intrinsic energy level and the decrease in capacitance above 1 MHz is believed to be associated with the rather large resistivity of the silicon bottom electrode since the theoretical plot shown in Figure 4.6.3 is obtained by substituting a value of R=3.46x10^{4}Ω in the transmission line equation (4.5.1) and the measured value of the source to drain resistance is 3.29 x 10^4 ohms. On the basis of the above information it was considered reasonable to analyse the capacitance results on the assumption that the 1 MHz measurements were not influenced by charge modulation in deep acceptor traps although it is still possible that some trapping is being masked by the transmission line effect.

In order to confirm the existence of deep acceptor traps with a frequency response in the range 5 kHz to 25 kHz the transient response of the drain current was examined when a pulse was applied to the gate of a device using the circuit shown in Figure 4.6.4.
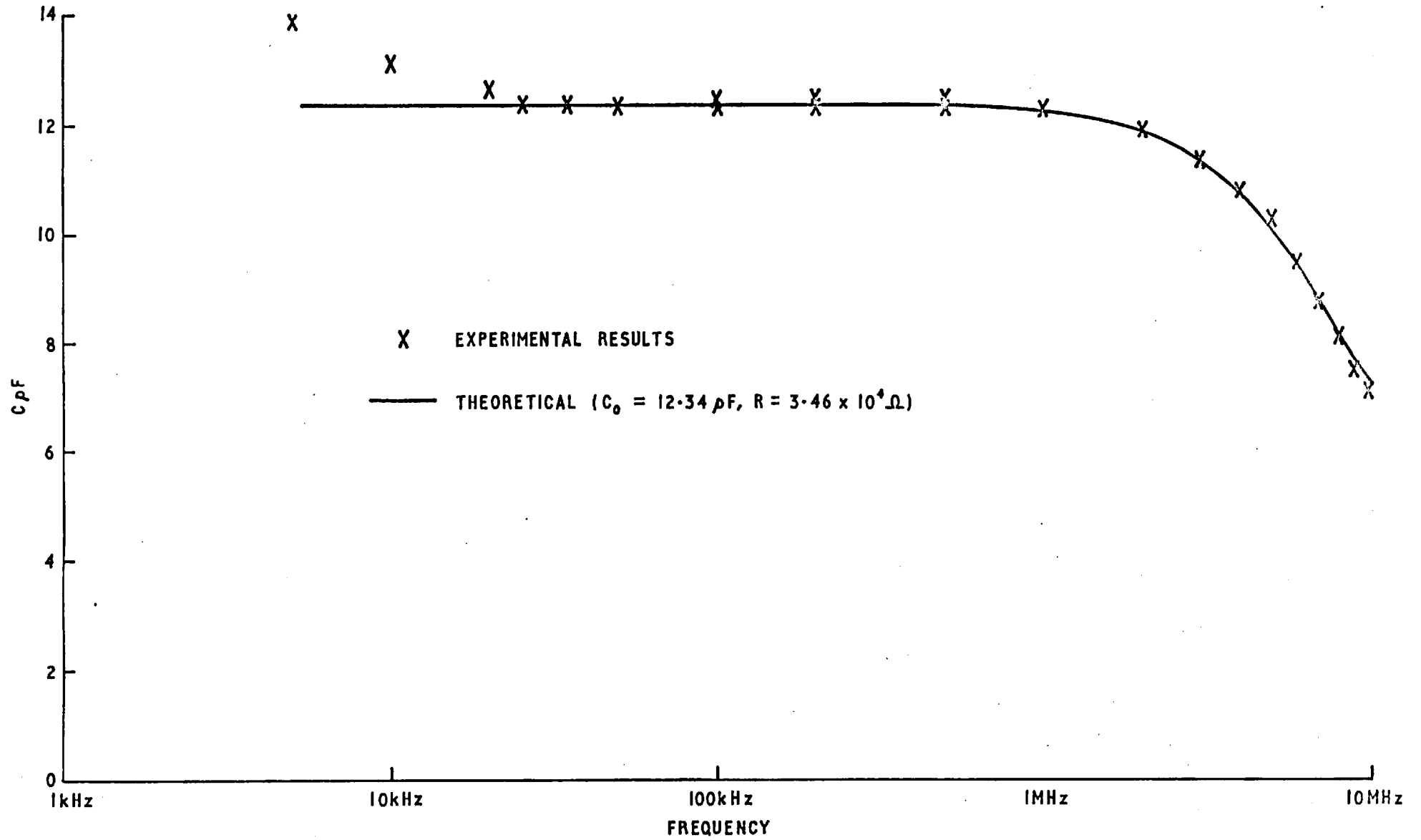
FIG. 4.6.3. FREQUENCY DEPENDENCE OF GATE CAPACITANCE AT ZERO APPLIED VOLTAGE
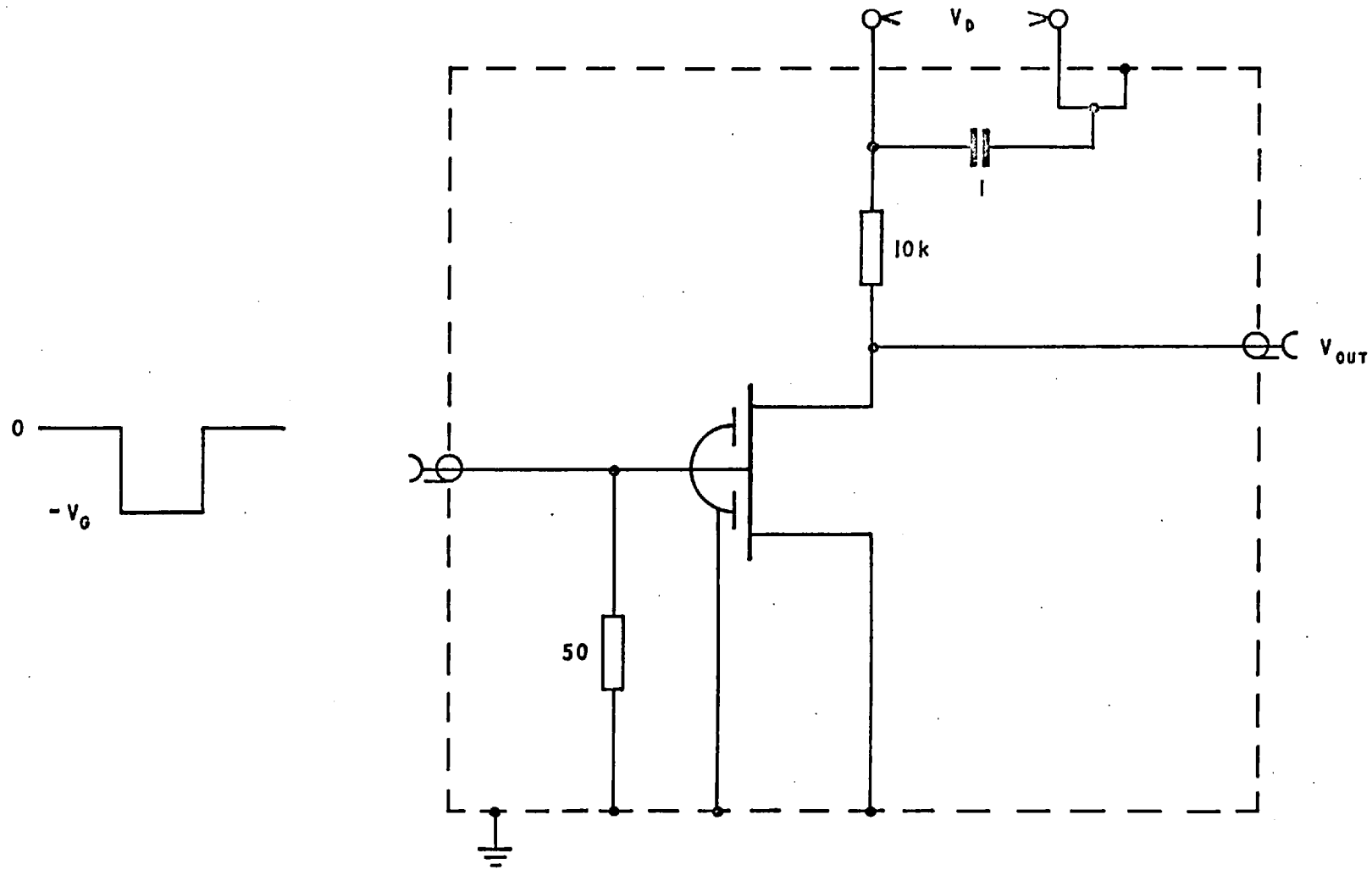
FIG. 4. 6. 4. TRANSIENT RESPONSE TEST CIRCUIT

The result shown in Figure 4.6.5 (a) clearly shows that when the drain current is switched on there is indeed an excess current which decays over a period of a few hundred microseconds. To demonstrate that this transient was a property of the device under test and not an experimental artifact a commercial junction field-effect transistor was also examined and was shown to have no detectable initial transient. A reduction of the width of the negative gate pulse below 200μs led to a decrease in the amplitude of the initial transient, presumably because there was insufficient time for many of the traps to empty during the "off-state". Figure 4.6.5 (b) shows an amplified version of the initial transient.

If this transient is a consequence of the existence of deep acceptor traps it will be proportional to the trap density and will decay with a time constant t which is related to the energy level of the traps by equation (4.2.9). Therefore

$$\Delta I_D = \Delta I_{DO} \, \exp \, (-t/\tau) \qquad\qquad \dots\dots\ 4.6.4$$

where ΔI_D is the drain current in excess of the steady state value (I_{DSS}) and ΔI_{DO} is the values of ΔI_D immediately after the device is switched on. The ratio $\Delta I_{DO}/I_{DSS}$ is a measure of the ratio of the trap density to the net donor density, (but it is not strictly equivalent to that ratio because the mobility is not uniform through the thickness of the film), and in the typical device above it has the value 0.175.

The information in Figure 4.6.5 (b) is plotted in the semi-logarithmic plot of Figure 4.6.6 and since it follows from equation (4.6.4) that

Gate voltage

0V
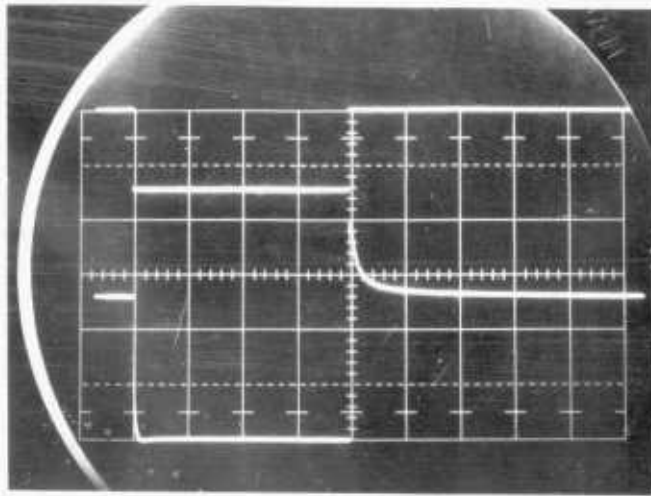
1V/div

Drain current

1μA/div

0A

Scan rate 200μ sec/div

(a)

Gate voltage

0V
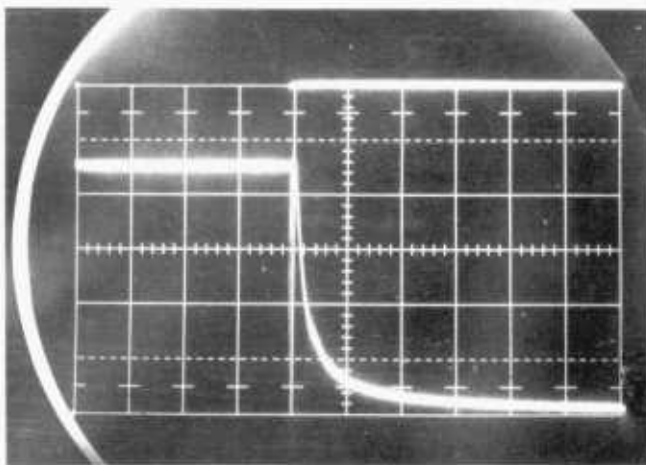
1V/div

Drain current

250nA/div

I_{DSS}

Scan rate 100μ sec/div

(b)

Figure 4.6.5.   Transient response of Drain Current

SLICE No. 586/1

V_D = 2 VOLTS

$\Delta I_D \mu A$

$\tau = 84 \mu sec$

$\tau = 310 \mu sec$

0·1

0·01

DEVICE SWITCHED ON (V_G=0)

0   50   100   150

TIME, μsec
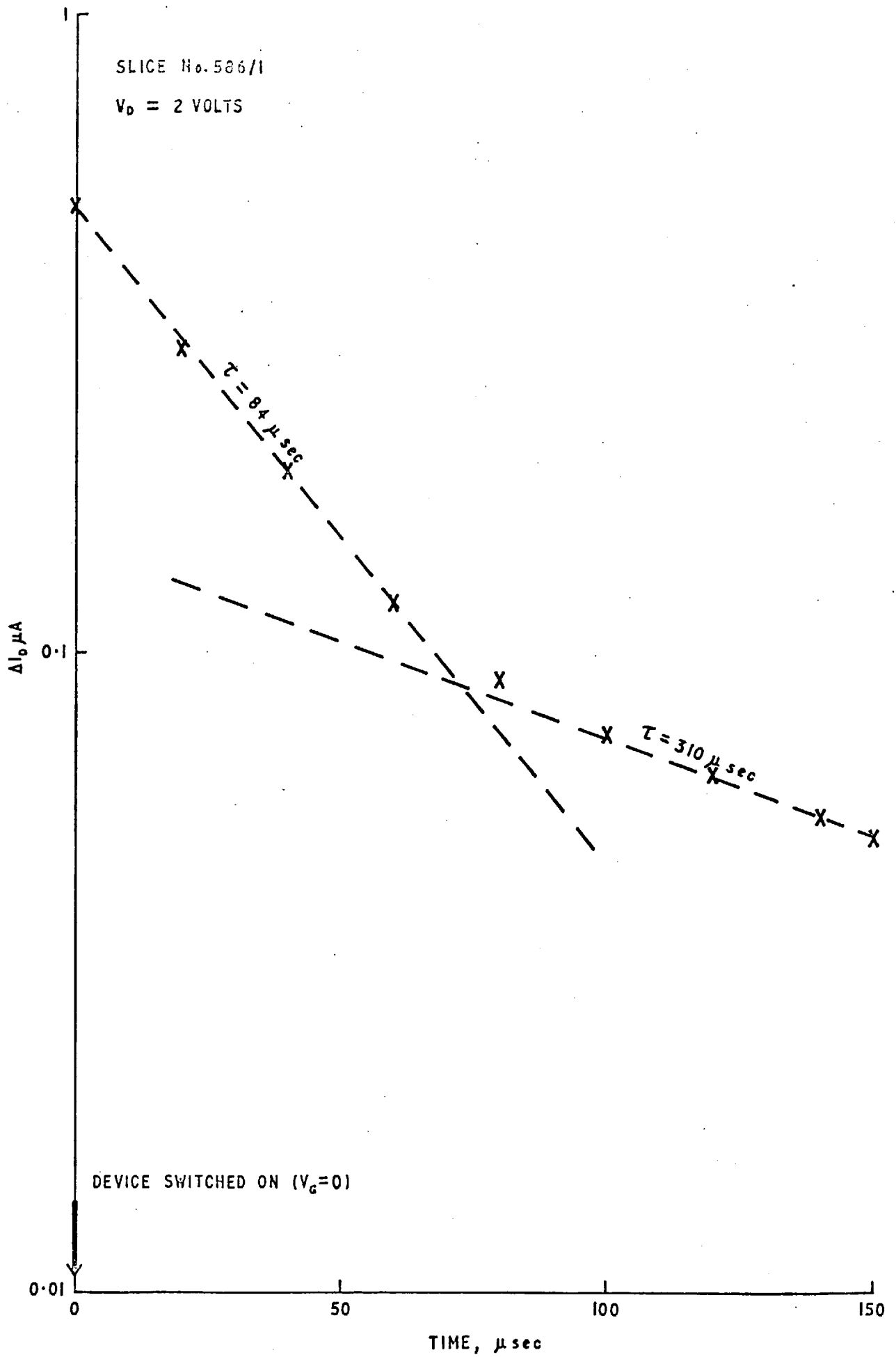
FIG. 4.6.6. TRANSIENT RESPONSE OF FIELD-EFFECT STRUCTURE

151

$dt/_{d\ln\Delta I_D} = -\tau$ all the points should lie on a straight line if the traps are at a single energy level. This is clearly not the case and they appear to have a range of time constants from 84µs to 310µs which for a capture cross section of $10^{-15}cm^2$ corresponds to energy levels of 0.09eV and 0.06eV respectively from the intrinsic level.

These results are in good agreement with the observed frequency dependence of the capacitance and indicate trapping levels close to the lower energy level limit observed by Heiman (1967b) which is much closer to the intrinsic level than the observations of Dumin (1970). The above conclusions must remain somewhat tentative however until a value has been obtained for the capture cross section of the traps since a difference of an order of magnitude in this parameter corresponds to an energy level difference of approximately 0.06eV.

The 1MHz gate capacitance measurements can now be used with some confidence together with the channel conductivity and gate leakage current measurements to determine the carrier concentration mobility and lifetime profiles by applying the procedures outlined in sections 4.2, 4.3 and 4.4 and a typical result is given in Figure 4.6.7.

The points marked by shaded symbols were calculated without making a correction for the finite resistivity of the silicon layer below the space charge layer and the points marked by open symbols were calculated using the capacitance of the space charge layer as found from an iterative computation based on equation (4.5.1). The existence in the uncorrected results of an apparent decrease in the carrier concentration and an apparent increase in the lifetime in the region of the sapphire interface are both shown to be anomolies resulting from the effect of the resistive lower electrode on the measurements,
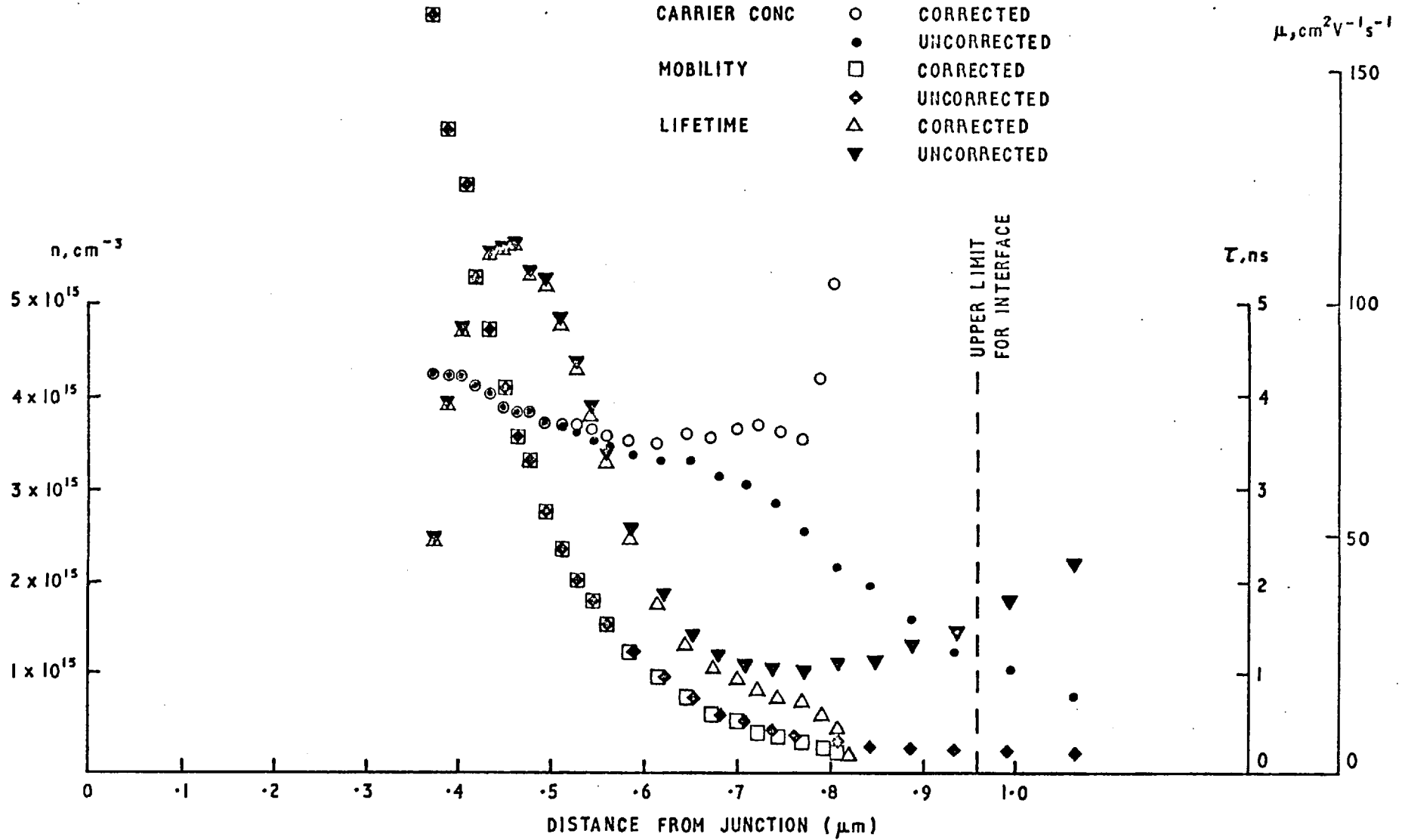
FIG. 4. 6. 7.   CARRIER CONCENTRATION, MOBILITY AND LIFETIME PROFILES

which clearly demonstrates the importance of applying this correction.

The apparent rapid increase in the carrier concentration in the corrected results and the decrease in the lifetime at approximately 0.78μm from the junction are probably caused by the depletion region reaching the sapphire interface. This would indicate a junction depth of 0.59μm for the silicon-on-sapphire layer compared with a depth of 0.41μm for the bulk silicon or in other words a diffusion constant about twice as great. This is well within the range of enhanced diffusion constants reported for silicon-on-sapphire (Dumin and Robinson 1966 and Ross and Warfield 1969).

The mobility results, which are not significantly effected by the correction for the resistive lower electrode, decrease more rapidly than those observed on the vacuum evaporated p-type layers and generally confirm the tentative conclusions derived from the analysis of the transfer characteristics. The rapid decrease of both mobility and lifetime as the sapphire interface is approached is characteristic of these films and is probably a consequence of increasing structural disorder close to the sapphire interface, (Linnington 1970).

Figure 4.6.8 shows the carrier concentration results up to 0.78μm from the junction on an expanded scale together with a second order polynomial best fit, and the second order polynomial was used in the manner described in the previous section to calculate the net donor density from equation (4.2.16) and this is shown in Figure 4.6.8 as a dashed line. The net donor density differs from the carrier concentration by less than the spread in the experimental results.
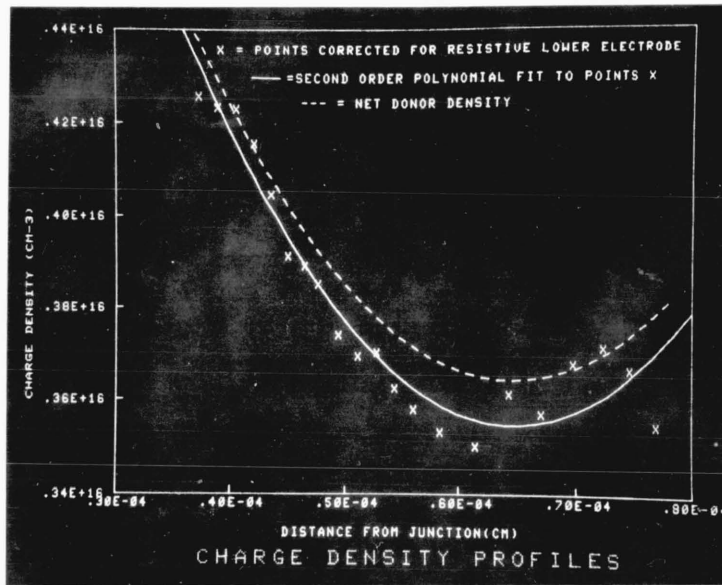
Figure 4.6.8. Carrier Concentration and Net Donor Density Profiles

The slight decrease of net donor density in the direction of the

sapphire interface which occurs over most of the layer thickness is

probably due to an increasing level of compensation by deep and

shallow acceptors and the accuracy of the measurements close to the

sapphire interface is not sufficiently good to place much credence

in the slight upturn in the profile in that region.

4.7 Slow Transients Resulting from Sapphire Interface States

The very slow decay in the drain current of a probed device that

occurs when the light on the probing machine is switched off

has been referred to in the previous section and a typical behaviour

pattern is shown in Figure 4.7.1. The rapid decrease in the current

at the instant of switching off the light is due to the recombination

of the excess electron hole pairs which were generated by the

illumination. The duration of this transition is of the same order of

magnitude as the minority carrier lifetime and in fact the analysis of this

transition is the basis of the familiar Stevenson-Keyes (1955)

method for measuring lifetimes. The much slower transient that follows

is believed to be due to slow interface states at the sapphire-silicon

boundary.

The energy band diagram at zero gate bias for a section through the

channel of a device with a distribution of interface states near the mid-

band position is shown in Figure 4.7.2 (a). Electromagnetic radiation

within the visible spectrum is sufficiently energetic to create electron

hole pairs and also to excite electrons from the interface states into

the conduction band. However if these interface states are

associated with defects within the first few nanometers of the sapphire

substrate then the emission probability may be very low and hence

the associated time constant to attain steady-state conditions

correspondingly long, (Heiman and Warfield 1965). Similarly when the light

is switched off the effective capture cross section is very small and the

time to re-establish a steady-state condition in the dark is also

very long. The slow tail of the transient in Figure 4.7.1 can be used
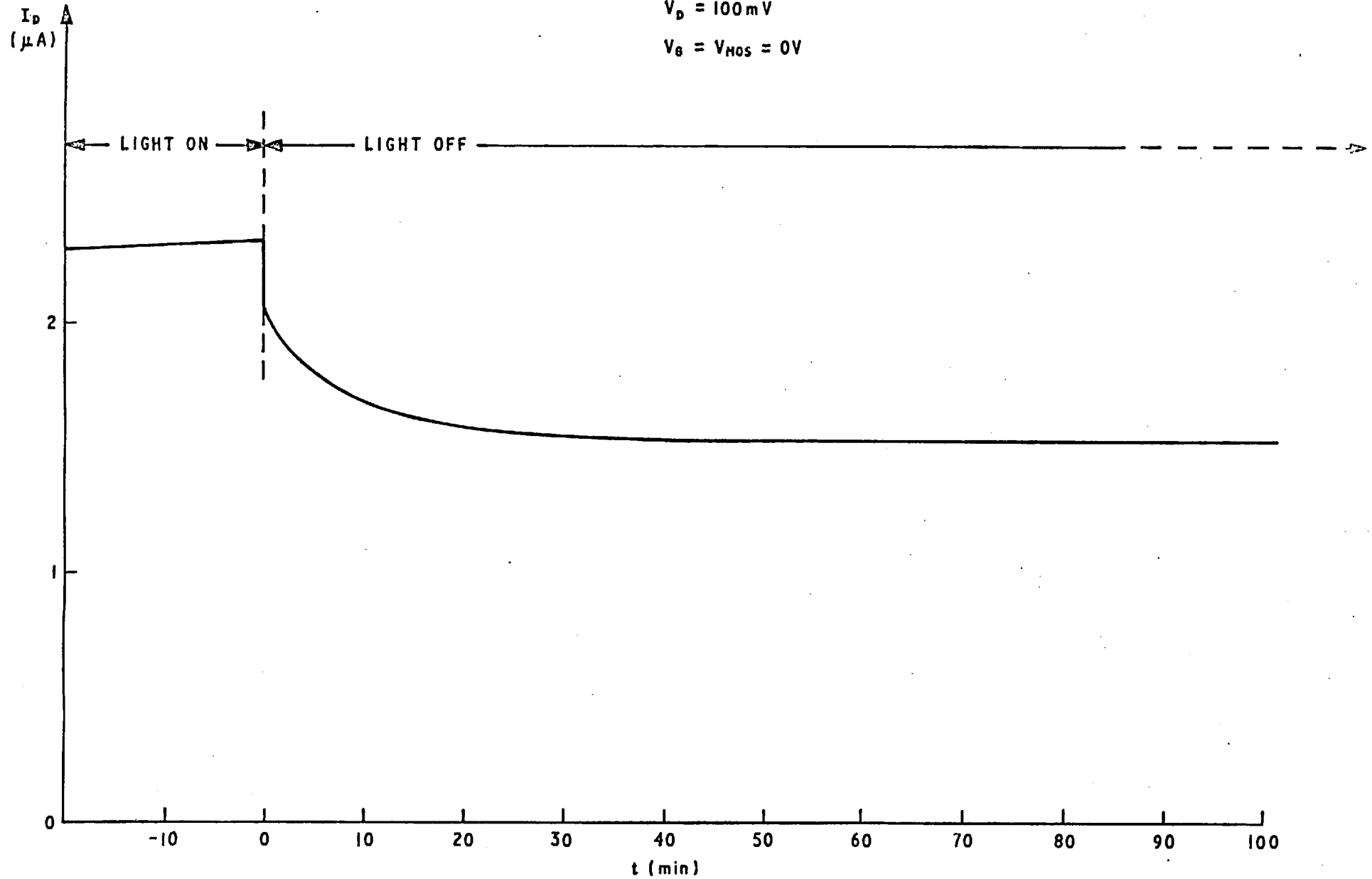
FIG. 4.7.1. TRANSIENT RESPONSE OF A JUNCTION FIELD-EFFECT STRUCTURE
WHEN ILLUMINATION IS REMOVED

SPECIMEN No. 586/1 R9C14

$V_D = 100\,mV$

$V_G = V_{MOS} = 0V$
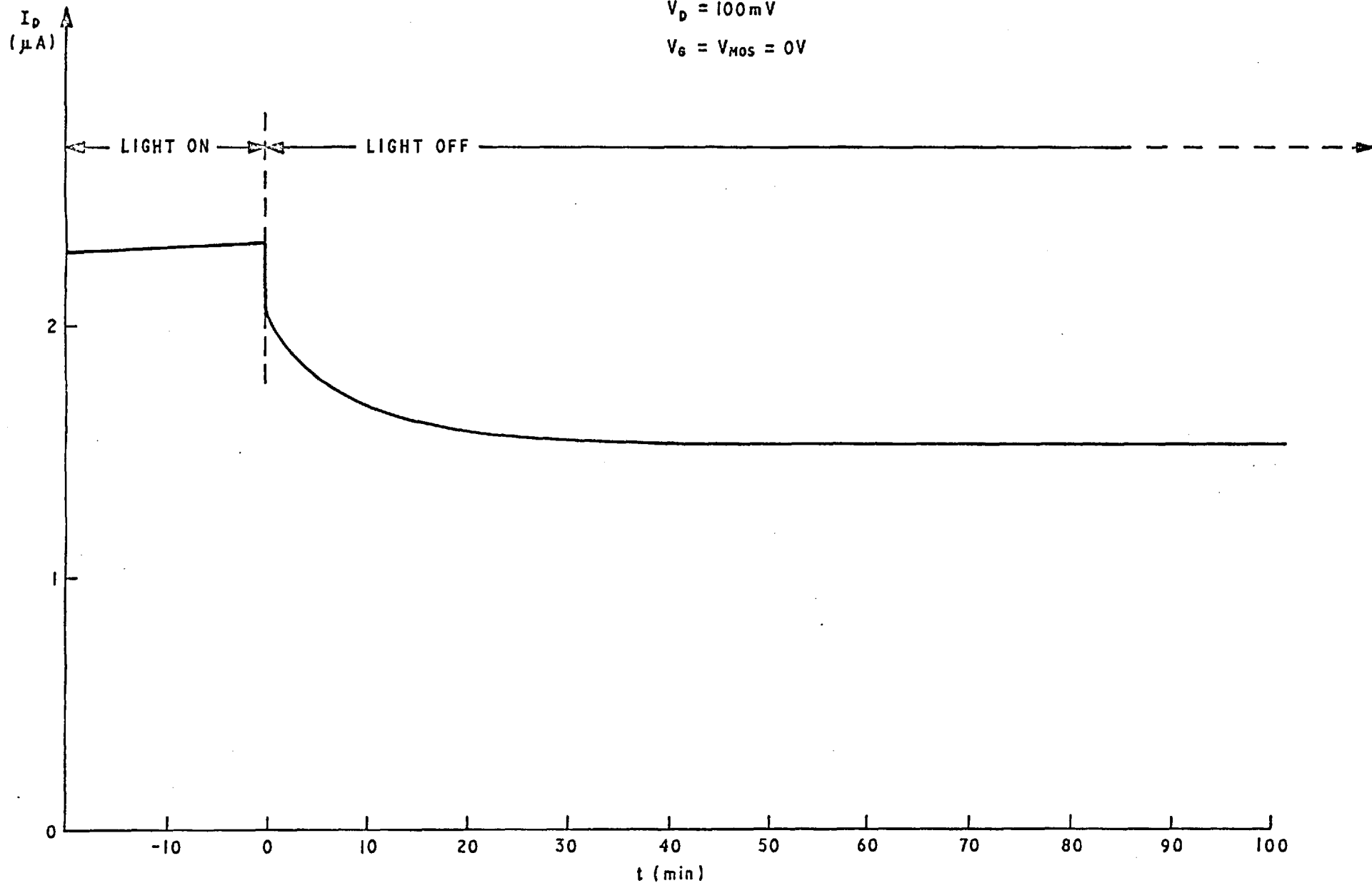
LIGHT ON — LIGHT OFF

t (min)

FIG. 4. 7. 1. TRANSIENT RESPONSE OF A JUNCTION FIELD-EFFECT STRUCTURE
WHEN ILLUMINATION IS REMOVED

to obtain an approximate estimate of 4.5×10^{11} cm^{-2} for the density

of interface states by assuming that the mean mobility for the

interfacial electrons is 10 cm^2 V^{-1} s^{-1}.

In order to justify the above assumption that the slow states are

located at the silicon-sapphire interface proof is now given that they

cannot be associated with defects in the bulk of the silicon. It was

shown in section 4.2 that the electron capture probability for deep

traps in the silicon is given by v_{th} σ_n n and therefore the

time constant associated with this electron capture process is $1/v_{th}$ σ_n n.

At a typical carrier density of 10^{15} cm^{-3} this gives a time

constant of 10^{-7} sec for a capture cross section of 10^{-15} cm^2 whereas a

time constant of 20 minutes would require a capture cross section of

10^{-25} cm^2 at this carrier density. A capture cross-section of this

size is so much smaller than atomic dimensions that it is not

really feasible. Furthermore if bulk traps were responsible for the slow

transients then the width of the depletion region would relax after the

application of a reverse-bias pulse to the junction gate (Sah and

Reddi 1964). This would give rise to a variation in gate capacitance

but no slow transients were observed in the gate capacitance if the

capacitance was measured at 1 kHz unless the pulsed gate voltage approached

the pinch-off voltage. The measured capacitance did exhibit a slow

transient if the reverse bias pulse was sufficiently large and the

measurement was made at 1 MHz but this was due to variations in the

series resistance of the silicon film beneath the depletion layer and

not to a variation in the width of the depletion layer.

Measurements were also made on encapsulated field-effect devices of the

slow transient which occurs when the channel is switched on after being

held 'hard-off' for sometime. Figure 4.7.3(a) shows the decay which

SPECIMEN No.586/2 A19
$V_D = 100mV.$ $V_{MOS} = 0V$
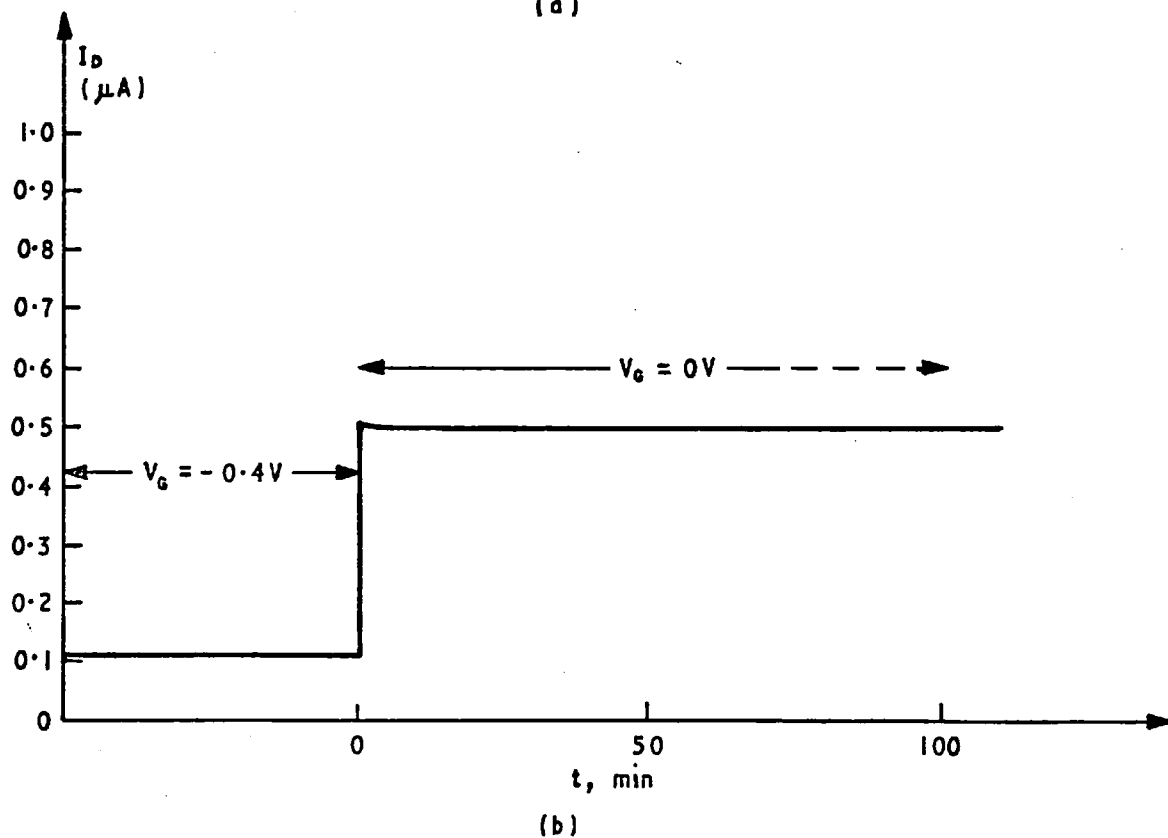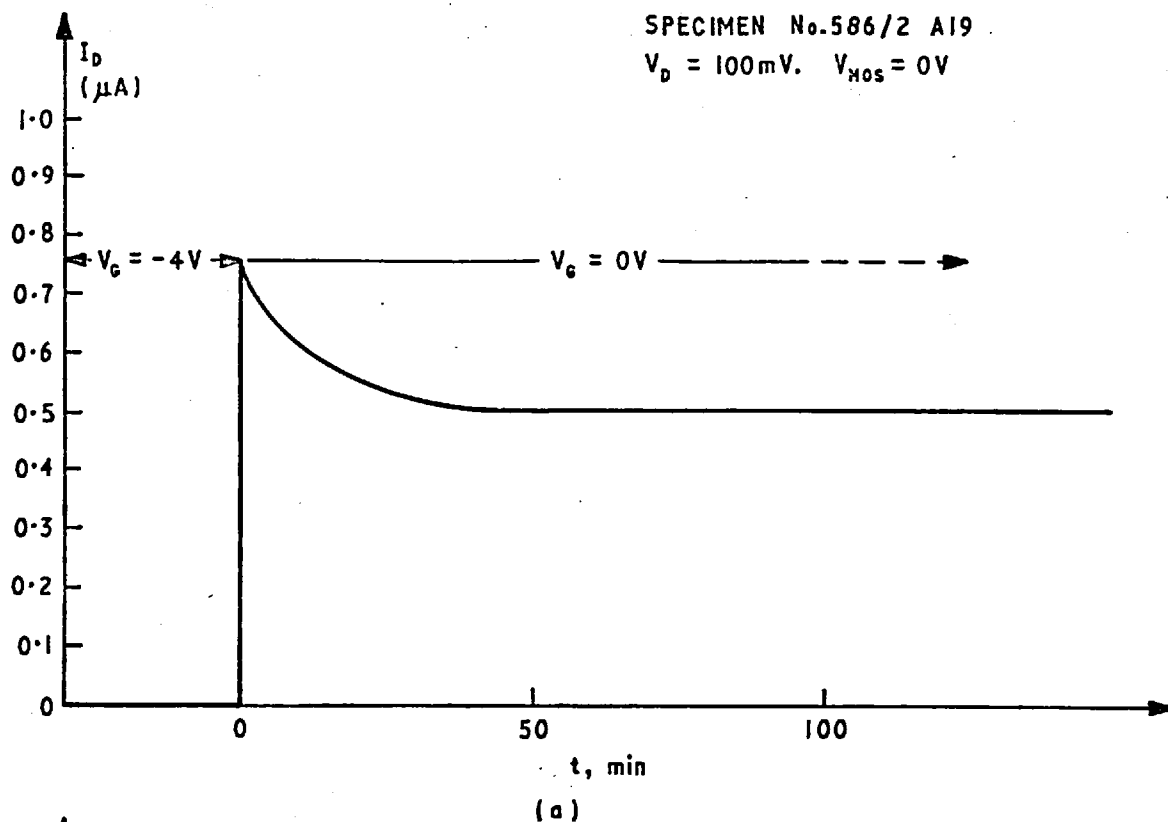
FIG. 4.7.3. TRANSIENT RESPONSE OF JUNCTION FIELD-EFFECT
STRUCTURE TO CHANGES IN GATE VOLTAGE
(a) WHEN CHANNEL IS 'OFF' IN INITIAL STATE
(b) WHEN CHANNEL IS 'CONDUCTING' IN INITIAL STATE

161

occurs in the drain current of a typical device when it is turned on after being turned 'hard-off' for 4 hours with a potential of -4 volts on the junction gate. The energy band diagram for the 'off' device is shown in Figure 4.7.2 (b) and from this it can be seen that any interface states near the mid-gap position lie above the quasi Fermi level for electrons and therefore they will slowly empty. When the applied junction gate bias is reduced to zero and the device is turned back on the energy bands are as shown in Figure 4.7.2(a) with the exception that the interface states are initially empty and therefore the drain current rises to a value above the steady-state level and then decreases slowly as the interface states are filled. The transient shown in Figure 4.7.3 (a) corresponds to an interface state density of 1.4×10^{11} cm^{-3} for a mean interface mobility of 10 cm^2 V^{-1} s^{-1}. All the devices examined exhibited some trapping at the sapphire interface with estimated interface state densities ranging from 1.2×10^{11} cm^{-2} to 6.5×10^{11} cm^{-2}.

Figure 4.7.3 (b) shows that when the junction gate bias is reduced to zero after having had a bias applied which is not sufficient to completely turn-off the device then the drain current returns quickly to its steady state value. This is simply because the quasi Fermi level at the silicon-sapphire interface only starts to shift significantly when the depletion region penetrates to this interface. U nfortunately, since it is not possible to determine the potential at the silicon-sapphire interface as a function of the applied junction gate bias, the interface state density cannot readily be determined as a function of its position in the energy gap.

CHAPTER 5

GENERAL CONCLUSIONS

Detailed interpretation and analysis of the results has been developed
in each Chapter. Consequently many of the more significant conclusions
concerning the electrical properties of the layers and their controlling
transport mechanisms have already been discussed. All that remains
to be done is to summarise those conclusions while demonstrating
that the information obtained with the 'MIS Hall Effect' and the 'Junction
Field Effect' structures is indeed compatible and complementary and to
interpret its significance in terms of the performance, yield and
reliability of the various integrated circuit structures which are likely
to be fabricated in silicon-on-sapphire layers.

The most important conclusion, at least from a technological viewpoint,
is that the mobility of the charge carriers decreases by approximately
an order of magnitude across the thickness of the film. Electron
mobilities up to 450 $cm^2V^{-1}s^{-1}$ were observed in accumulation layers
adjacent to the silicon dioxide interface using MIS Hall effect
structures and, although accurate values cannot be obtained immediately
adjacent to the sapphire interface, junction field-effect measurements
show that it certainly drops below 50 $cm^2V^{-1}s^{-1}$ in most layers. Since the
two different measurement structures were not fabricated on the same
layers a direct comparison of results is not possible even in the middle
regions of the layers where the measurement techniques overlap, however
the general trends of both sets of results show rapid decreases in
mobility with up to 50% change across a distance of only 100 nm. The Hall
effect results indicate that the surface mobility is limited by
phonon scattering but scattering at charged centres plays a major role
in determining the mobility at the sapphire interface.

The observed variations in carrier concentration were smaller and less important but still significant with changes by a factor of five occurring across the thickness of the layer in some specimens. In n-type films the net donor density decreased towards the sapphire interface and in p-type vacuum evaporated, autodoped films the net acceptor density increased towards the sapphire interface. The 'MIS Hall effect' and 'junction field-effect' results on the n-type layers were again compatible and indicated that the doping profile was steeper adjacent to the silicon dioxide interface, probably because of 'pile-up' during oxidation resulting from the relatively high segregation coefficient of phosphorus.

The "effective lifetime" profile obtained from the junction field-effect results has a similar shape to the mobility profile in the region of the sapphire interface. Both profiles are probably a consequence of the reduction of the mean free path of the carriers by the increasing density of defect centres. The smaller decrease in lifetime as the oxide interface is approached is not understood.

No special silicon-dioxide interface problems seem to exist in these layers. The density and energy distributions of interface states appears to be similar to that obtained on bulk silicon and the scattering of carriers at this interface is more specular than it is in bulk silicon.

The silicon-sapphire interface on the other hand does present additional problems. In the MIS measurements enhanced electronic conduction was observed at this interface in some of the specimens and although this effect was not present on the junction field-effect structures, for which the fabrication process was different, it remains a potential problem until the cause has been identified and reliable methods established to ensure it does not occur. The existence of a fairly high density of slow states was also observed at the sapphire interface. These states

were only observed with the junction field-effect structures but they would

not have been detected in the MIS studies because they are only made

manifest by changes in potential at the sapphire interface.

Many different types of component can be incorporated into MIS integrated

circuits manufactured in thin films of silicon-on-sapphire in order

to achieve a wide range of digital logic and memory functions. However the

special problems associated with this material will effect different circuit

designs in different ways. Components which can be used to construct

circuits include the conventional n-channel and p-channel inversion

layer transistors together with the deep-depletion transistors and small

area gated and ungated load resistors which can only be made in these

thin high resistivity films. From the wide range of design options

the decision on which technology is most suitable for a given

application will depend on a compromise between a number of factors, such

as

    i   Ease of manufacture

    ii   Yield and maximum die size, which are inter-related,

    iii   Area required to achieve a given function,

    iv   Reliability

    v   Ability to meet electrical specifications on parameters

          such as speed, power consumption, output drive capability

          and operating voltages.

When the inversion layer p-channel, n-channel and complementary MOS

designs familiar in bulk silicon technology are transferred to silicon-on-

sapphire many of the considerations are identical because, as has been

shown, the properties of inversion layers in silicon-on-sapphire

are broadly similar to those in bulk silicon. However it must be

remembered that even with similar technologies the power-speed

product achieved in silicon-on-sapphire circuits is much greater

because the horizontal p-n junctions have been eliminated and the

capacitance of the remaining vertical junctions is an order of magnitude

smaller. Some other advantages of silicon-on-sapphire such as the

elimination of parasitic MOST's, the achievement of higher packing

densities and the replacement of p-n junction isolation by dielectric

isolation in the complementary symetry technology have been discussed

in detail by other authors, (eg Wilcock 1971), the emphasis in this

discussion will therefore be placed on the significance of the electrical

properties measured in the present investigation. The electronic

interface layers reported in the above MIS studies could give rise

to low junction breakdown voltages in p-channel inversion-layer transistors

and if similar layers exist in p-type films they would give rise to large

leakage current problems by shorting out the source and drain at the

sapphire interface. However there is enough evidence in the literature

to show that circuits can be manufactured in which these problems are not

dominant but they may still remain a yield limiting factor until their

origin is better understood.

The variations in electrical properties through the thickness of the film

are most significant when the novel silicon-on-sapphire components

which involve conduction in the bulk of the film are considered. The

proposed complementary symmetry technology, (Boleky 1970), using a single

conductivity type material with one n-type inversion layer transistor

and one p-type deep-depletion transistor looks attractive in its

simplicity of fabrication but it has already been pointed out by Boleky

that this approach requires very good control of film doping to obtain

consistently complete depletion of the p-channel device at zero applied gate

voltage. If the speed-power performance of the circuits were also limited by the low mobility of the charge carriers in the conducting channel close to sapphire interface then the commercial viability of this technology would be questionable. However, although there is conduction in the bulk of the film when the transistor is turned on the total number of bulk carriers is less than 10^{11} per square centimetre and the majority of the charge carriers are in an accumulation layer at the silicon-dioxide interface. The ratio of surface conduction to bulk conduction clearly depends on the operating voltage but even at 5 volts over 80 per cent of the carriers are in the accumulation layer.

The mobility profile can even be considered as an advantage because it means that if depletion does not quite extend to the sapphire interface at zero gate voltage the quiescent power dissipation is still almost an order of magnitude less than it would be with a uniform mobility profile and therefore a relatively low power "pseudo-complementary" high-speed technology is achievable with less stringent control of the film doping levels.

As a final conclusion it can be said that although silicon on sapphire layers contain high densities of defects and large variations in electrical properties they do make large improvements in the speed-power product of MOS integrated circuits possible but further investigation of the origin and nature of defects at the silicon sapphire interface will be necessary to improve the yield and reliability of these circuits.

REFERENCES

ALLISON, J F, DUMIN, D J, HEIMAN, F P, MUELLER, C W and ROBINSON, P H, 1969
"Thin-Film Silicon: Preparation, Properties and Device Applications", Proc. IEEE, 57, pp 1490-1498.

ANG, C Y and MANASEVIT, H M, 1965
"Residual Stress in Epitaxial Silicon Films on Sapphire", Sol. St. Electron, 8, pp 994-996.

ATALLA, M M, BRAY A R and LINDER, R, 1959
"Stability of Thermally Oxidised Silicon Junctions in Wet Atmospheres", Proc. IEEE, 106, Pt. B, Suppl. 17, pp 1130-1137.

ATALLA, M M and TANNENBAUM, E, 1960
"Impurity Redistribution and Junction Formation in Silicon by Thermal Oxidation", Bell Syst. Tech. J., 39, pp 933-946.

BERGLUND, C N, 1966
"Surface States at Steam-Grown Silicon-Silicon Dioxide Interfaces", IEEE Trans. Electron Devices, ED-13, pp 701-705.

BICKNELL, R W, JOYCE, B A, NEAVE, J H and SMITH, G V, 1966
"The Epitaxy of Silicon on Alumina-Structural Effects", Phil. Mag., 14, pp 31-46.

BLANK, J M and RUSSELL, V A, 1966
"Nucleation and Crystal Growth of Silicon on Sapphire", Trans. Met. Soc. AIME, 236, pp 291-294.

BOLEKY, E J, 1970
"The Performance of Complementary MOS Transistors on Insulating Substrates", RCA Rev., 31, pp 372-395.

BOOKER, G R and STICKLER, R, 1965
"Two-Dimensional Defects in Silicon after Annealing in Wet Oxygen", Phil. Mag., 11, pp 1303-1308.

BOOKER, G R and TUNSTALL, W J, 1966
"Diffraction Contrast Analysis of Two-Dimensional Defects present in Silicon after Annealing", Phil. Mag., 13, pp 71-83.

BROWN, D M and GRAY, P V, 1968
"Si-SiO$_2$ Fast Interface State Measurements", J. Electrochem. Soc., 115, pp 760-766.

CALZOLARI, P U and GRAFFI, S, 1972
"A Theoretical Investigation on the Generation Current in Silicon p-n Junctions under Reverse Bias", Sol. St. Electron., 15, pp 1003-1011.

CHANG, C C, 1971
"Silicon-on-Sapphire Epitaxy by Vacuum Sublimation: LEED-Auger Studies and Electronic Properties of the Films", Jnl. Vac. Sci. and Technol., 8, pp 500-511.

CHENG, Y C and
SULLIVAN, E A,
1973

"Relative Importance of Phonon Scattering
to Carrier Mobility in Si Surface Layer at
Room Temperature", J. Appl. Phys., 44,
pp 3619-3625.

CHIANG, Y S and
RICHMAN, D,
1970

Paper presented at Met. Soc. AIME, New York,
Sept., 1970.

CHU, T L, FRANCOMBE, M H,
GRUBER, G A, OBERLEY, J J
and TALLMAN, R L,
1965

Westinghouse Report No. AFCRL-65-574 for
US Air Force Cambridge Research Laboratories.

CULLEN, G W, GOTTLIEB, G E,
WANG, C C and ZAININGER, K H,
1969

"Epitaxial Growth and Properties of Silicon
on Alumina-Rich Single-Crystal Spinel",
J. Electrochem. Soc., 116, pp 1444-1449.

CULLEN, G W, GOTTLIEB, G E
and WANG, C C,
1970

"The Epitaxial Growth of Silicon on Sapphire
and Spinel Substrates: Suppression of
Changes in the Film Properties during Device
Processing", RCA Rev., 31, pp 355-371.

DAS, M B,
1969

"Physical Limitations of MOS Structures",
Sol. St. Electron, 12, pp 305-336.

DUMIN, D J,
1965

"Deformation of and Stress in Epitaxial
Silicon Films on Single-Crystal Sapphire",
J. Appl. Phys., 36, pp 2700-2703.

DUMIN, D J,
1967

"Electrical Properties of Silicon Films
Grown Epitaxially on Sapphire",
J. Appl. Phys., 38, pp 1909-1914.

DUMIN, D J,
1970

"Deep Levels within the Forbidden Gap of
Silicon-on-Sapphire Films", Sol. St.
Electron., 13, pp 415-424.

DUMIN, D J and
ROBINSON, P H,
1966

"Autodoping of Silicon Films Grown
Epitaxially on Sapphire", J. Electrochem.
Soc., 113, pp 469-472.

DUMIN, D J and
ROBINSON, P H,
1968

"Carrier Transport in Thin Silicon Films",
J. Appl. Phys., 39, pp 2759-2765.

DUMIN, D J, ROBINSON, P H,
CULLEN, G W and GOTTLIEB, G E,
1970

"Heteroepitaxial Growth of Germanium and
Silicon on Insulating Substrates, RCA Rev.,
31, pp 620-634.

DUMIN, D J and ROSS, E C,
1970

"Temperature Dependence of the Hall Mobility
and Carrier Concentration in Silicon-on-
Sapphire Films", J. Appl. Phys., 41,
pp 3139-3143.

DUMIN, D J and
SILVER, R S,
1968

"Diffused Diodes in Silicon-on-Sapphire",
Sol. St. Electron., 11, pp 353-363.

EDWARDS, J R,
1969

"Evidence of Phosphorus N-skin on Silicon
from Vapor Transport", J. Electrochem. Soc.,
116, pp 866-868.

FAKTOR, M M,
FIDDYMENT, D G and
NEWNS, G R,
1967

"Preliminary Study of the Chemical
Polishing of α-Corundum Surfaces with
Vanadium Pentoxide", J. Electrochem. Soc.,
114, pp 356-359.

FANG, F and
FOWLER, A B,
1968

"Transport Properties of Electrons in
Inverted Silicon Surfaces", Phys. Rev.,
169, pp 619-631.

FILBY, J D,
1966

"The Chemical Polishing of Single Crystal
α-Alumina using Silicon", J. Electrochem.
Soc., 113, pp 1085-1087.

FOWLER, A B, FANG, F
and HOCHBERG, F,
1964

"Hall Measurements on Silicon Field Effect
Transistor Structures", IBM Jnl. Res.
Develop., 8, pp 427-429.

GOETZBERGER, A,
1967

"Behaviour of MOS Inversion Layers at
Low Temperature", IEEE Trans. Electron
Devices, ED-14, pp 787-789.

GOLDSTEIN, Y, GROVER, N B,
MANY, A and GREENE, R F,
1961

"Improved Representation of Calculated
Surface Mobilities in Semiconductors. II.
Majority Carriers", J. Appl. Phys., 32,
pp 2540-2541.

GOTTLIEB, G E, CORBOY, J F,
CULLEN, G W and SCOTT, J H,
1974

to be published in Trans. Met. Soc. AIME.

GREENE, R F,
1964

"Surface Transport", Surface Sci., 2,
pp 101-113.

GRIMSHAW, J A and
OSBORNE, D N

"Charge Neutrality in Semiconductors with
Implanted Impurity Profiles", Sol. St.
Electron., 14, pp 603-611.

GROVE, A S,
1967

"Physics and Technology of Semiconductor
Devices", published by John Wiley and Sons
Inc.

GROVE, A S and
FITZGERALD, D J,
1966

"Surface Effects on p-n Junctions:
Characteristics of Surface Space-Charge
Regions under Non-Equilibrium Conditions",
Sol. St. Electron., 9, pp 783-806.

GUPTA, D C,
1968

"Diode Voltage-Capacitance Method
for Measuring Resistivity and Impurity
Profile in a Silicon Epitaxial Layer",
Sol. St. Technology, 11, 2, pp 31-34.

HAGER, C K,
1959

"Network Design of Microcircuits",
Electronics, 32, 36; pp 44-49.

HART, P B, ETTER, P J,
JERVIS, B W and
FLANDERS, J M,
1967

"Electrical Properties of Epitaxial
Silicon Films on α -Alumina", Brit.
J. Appl. Phys., 18, pp 1389-1398.

HASEGAWA, S, KAMINAKA, N,
NAKAMURA, T and
ITOH, T,
1969

"Analysis of Carrier Transport in Vacuum-
Evaporated Epitaxial Films of Silicon on
Spinel", J. Appl. Phys., 40, pp 4620-4623.

HEIMAN, F P,
1966

"Thin-Film Silicon-on-Sapphire Deep
Depletion MOS Transistors", IEEE Trans.
Electron Devices, ED-13, pp 855-862.

HEIMAN, F P,
1967a

"On the Determination of Minority Carrier
Lifetime from the Transient Response of
an MOS Capacitor", IEEE Trans. Electron
Devices, ED-14, pp 781-784.

HEIMAN, F P,
1967b

"Donor Surface States and Bulk Acceptor
Traps in Silicon-on-Sapphire Films",
Appl. Phys. Lett., 11, pp 132-134.

HEIMAN, F P and
WARFIELD, G,
1965

"The Effect of Oxide Traps on the MOS
Capacitance", IEEE Trans. Electron Devices,
ED-12, pp 167-178.

HILIBRAND, J and
GOLD, R D,
1960

"Determination of the Impurity Distribution
in Junction Diodes from Capacitance-
Voltage Measurements", RCA Rev., 21,
pp 245-252.

IPRI, A C,
1972

"Electrical Properties of Silicon Films on
Sapphire using the MOS Hall Technique",
J. Appl. Phys., 43, pp 2770-2775.

IPRI, A C and
ZEMEL, J N,
1973

"Impurity Centres in Silicon Films on
Sapphire", J. Appl. Phys., 44, pp 744-751.

IRVIN, J C,
1972

"Resistivity of Bulk Silicon and of
Diffused Layers in Silicon", Bell Syst.
Tech. Jnl., 41, pp 387-410.

ISENBERG, I, RUSSELL, B R
and GREENE, R F,
1948

"Improved Method for Measuring Hall
Coefficients", Rev. Scient. Instrum.,
19, pp 685-688.

ITOH, T, HASEGAWA, S
and WATENABE, H,
1968a

"Epitaxial Films of Silicon on Sapphire
formed by Vacuum Evaporation and their
Electrical Properties", J. Appl.Phys.,
39, pp 2969-2971.

ITOH, T, HASEGAWA, S and KAMINAKA, N, 1968b — "Electrical Properties of n-type Epitaxial Films of Silicon on Sapphire formed by Vacuum Evaporation", J. Appl. Phys., 39, pp 5310-5313.

ITOH, T, HASEGAWA, S and KAMINAKA, N, 1969 — "Epitaxial Films of Silicon on Spinel by Vacuum Evaporation", J. Appl. Phys., 40, pp 2597-2600.

JEFKINS, D M, 1970 — "Stress in Silicon Films Deposited Heteroepitaxially on Insulating Substrates with Particular Reference to Corundum", J. Phys. D: Appl. Phys., 3, pp 770-777.

JOYCE, B A, BENNETT, R J, BICKNELL, R W and ETTER, P J, 1965 — "The Epitaxial Deposition of Silicon on Quartz and Alumina", Trans. Met. Soc. AIME, 233, pp 556-561.

KENNEDY, D P and O'BRIEN, R R, 1969 — "On the Measurement of Impurity Atom Distributions by the Differential Capacitance Technique", IBM J. Res. Develop, 13, pp 212-214.

KENNEDY, D P, MURLEY, P C and KLEINFELDER, W, 1968 — "On the Measurement of Impurity Atom Distributions in Silicon by the Differential Capacitance Technique", IBM J. Res. Develop, 12, pp 399-409.

KUHN, M, 1970 — "A Quasi-Static Technique for MOS C-V and Surface State Measurements", Sol. St. Electron., 13, pp 875-885.

LAWSON, R W and JEFKINS, D M 1970 — "The Optimum Conditions for the Vacuum Deposition of Silicon on Sapphire", J. Phys. D: Appl. Phys., 3, pp 1627-1640.

LEDERHANDLER, S R and GIACOLETTO, L J, 1955 — "Measurement of Minority Carrier Lifetime and Surface Effects in Junction Devices", Proc. IRE, 43, pp 477-483.

LEHOVEC, K, 1968 — "Rapid Evaluation of C-V Plots for MOS Structures", Sol. St. Electron, 11, pp 135-137.

LEISTIKO, O and GROVE, A S, 1966 — "Breakdown Voltage of Planar Silicon Junctions", Sol. St. Electron., 9, pp 847-852.

LININGTON, P F, 1970 — "The Structure of Epitaxial Silicon Films on Sapphire", Proc. 7th International Congress of Electron Microscopy, Grenoble, pp 447-448.

MANASEVIT, H M and FORBES, D H, 1966 — "Single Crystal Silicon on Spinel", J. Appl. Phys., 37, pp 734-739.

MANASEVIT, H M,
MILLER, A, MORRITZ, F L
and NOLDER, R L,
1965

"Heteroepitaxial Silicon-Aluminium Oxide Interface. Part 1 - Experimental Evidence for Epitaxial Relationships of Single-Crystal Silicon on Sapphire; An Overview of the Growth Mechanism", Trans. Met. Soc. AIME, 233, pp 540-549.

MANASEVIT, H M and
MORRITZ, F L,
1967

"Gas Phase Etching of Sapphire with Sulfur Fluorides", J. Electrochem. Soc., 114, pp 204-207.

MANY, A, GOLDSTEIN, Y
and GROVER, N B,
1965

"Semiconductor Surfaces", published by North-Holland, Amsterdam, p 312.

MATARÉ, H F,
1969

"Heteroepitaxy", Scientia Electica, 15, pp 95-109.

MERCIER, J,
1970

"Physical and Electrical Investigations on Silicon Epitaxial Layers on Sapphire Substrates", J. Electrochem. Soc., 117, pp 666-671.

MERCIER, J,
1971

"Properties of Low Temperature Heteroepitaxial Thin Silicon Films Grown in Helium Atmosphere", J. Electrochem. Soc., 118, pp 962-966.

MEYER, N I and
GULPBRANDSEN, T,
1963

"Method for Measuring Impurity Distributions in Semiconductor Crystals", Proc. IEEE, 51, pp 1631-1637.

MORIN, F J and
MAITA, J P,
1954

"Electrical Properties of Silicon Containing Arsenic and Boron", Phys. Rev., 96, pp 28-35.

MUELLER, C W and
ROBINSON, P H,
1964

"Grown-Film Silicon Transistors on Sapphire", Proc. IEEE, 52, pp 1487-1490.

MURPHY, N St J, BERZ, F
and FLINN, I,
1969

"Carrier Mobility in Silicon MOST's", Sol. St. Electron., 12, pp 775-786.

NABER, C T and
O'NEAL, J E,
1968

"The Deposition of Silicon on Sapphire in Ultrahigh Vacuum", Trans. Met. Soc., AIME, 242, pp 470-479.

NAN, S and YI-HUAN, L,
1964

"X-Ray Measurement of the Thermal Expansion Coefficient of Germanium, Silicon, Indium Antimonide and Gallium Arsenide", Acta. Phys. Sinica, 20, pp 699-704.

MOLDER, R and
CADOFF, I,
1965

"Heteroepitaxial Silicon-Aluminium Oxide Interface. Part II - Orientation Relations of Single Crystal Silicon on Alpha-Aluminium Oxide", Trans. Met. Soc. AIME, 233, pp 549-556.

MOLDER, R L, KLEIN, D J
and FORBES, D H,
1965

"Twinning in Silicon Epitaxially Deposited on Sapphire", J. Appl. Phys., 36, pp 3444-3450.

PALS, J A,
1970

"Method of Measuring Mobility of Majority Carriers in Silicon Devices", Electronic Letters, 6, pp 503-504.

PASHLEY, D W,
1965

"The Nucleation, Growth, Structure and Epitaxy of Thin Surface Films", Advances in Physics, 14, pp 327-416.

van der PAUW, L J,
1958

"A Method of Measuring Specific Resistivity and Hall Effect of Discs of Arbitrary Shape", Philips Rev., 13, pp 1-9.

PEARSON, G L, READ, W T
and FELDMAN, W L,
1957

"Deformation and Fracture of Small Silicon Crystals", Acta. Met., 5, pp 181-191.

PETRITZ, R L,
1958

"Theory of an Experiment for Measuring the Mobility and Density of Carriers in the Space-Charge Region of a Semiconductor Surface", Phys. Rev., 110, pp 1254-1262.

REDDI, V G P,
1968

"Majority Carrier Surface Mobilities in Thermally Oxidised Silicon", IEEE Trans. Electron Devices, ED-15, pp 151-160.

REYNOLDS, F H and
ELLIOT, A B M,
1966

"Etching Corundum with Silicon", Phil. Mag., 13, pp 1073-1074.

REYNOLDS, F H and
ELLIOT, A B M,
1967

"Vacuum Deposition of Silicon on Corundum", Sol. St. Electron., 10, pp 1093-1102.

ROBINSON, P H and
DUMIN, D J,
1968

"The Deposition of Silicon on Single-Crystal Spinel Substrates", J. Electrochem. Soc., 115, pp 75-78.

ROBINSON, P H and
HEIMAN, F P,
1971

"Use of HCl Gettering in Silicon Device Processing", J. Electrochem. Soc., 118, pp 141-143.

ROBINSON, P H and
MUELLER, C W,
1966

"The Deposition of Silicon upon Sapphire Substrates", Trans. Met. Soc. AIME, 236, pp 268-274.

ROSS, E C and
WARFIELD, G,
1969

"Effects of Oxidation on Electrical Characteristics of Silicon-on-Sapphire Films", J. Appl. Phys., 40, pp 2339-2344.

RUNYAN, W R,
1965

"Silicon Semiconductor Technology",
published by McGraw Hill Book Co.

SAH, C T, NOYCE, R H,
and SHOCKLEY, W,
1957

"Carrier Generation and Recombination in
p-n Junctions and p-n Junction
Characteristics", Proc. IRE, 45, pp 1228-1243.

SAH, C T and
REDDI, V G K,
1964

"Frequency Dependence of the Reverse-
Biased Capacitance of Gold-Doped Silicon
p^+-n Step Junctions", IEEE Trans. Electron
Devices, ED-11, pp 345-349.

SCHEUPLIN, R and
GIBBS, P,
1966

"Surface Structure in Corundum: I,
Etching of Dislocations", J. Am. Ceram, Soc.,
43, pp 458-472.

SCHIBLI, E and
MILNES, A G,
1967

"Deep Impurities in Silicon",
Materials Sci. Engng., 2, pp 173-180.

SCHLOTTERER, H,
1967

Paper presented at the IEEE Semiconductor
Device Research Conference, Bad Nauheim,
April 1967.

SCHLOTTERER, H,
1968

"Mechanical and Electrical Properties of
Epitaxial Silicon Films on Spinel", Sol.
St. Electron., 11, pp 947-956.

SCHLOTTERER, H and
ZAMINER, Ch.,
1966

"Crystal Imperfections in the Epitaxial
Growth of Silicon on Magnesium-Aluminium
Spinel", Phys. Status Solidi., 15,
pp 399-411.

SEITER, H and
ZAMINER, Ch.,
1965

"Epitaxial Silicon Films on Mg-Al Spinel",
Z. Angew, Phys., 20, pp 158-161.

SMITH, R A,
1961

"Wave Mechanics of Crystalline Solids",
published by Chapman and Hall, London.

STEPHENS, D L and
ALFORD, W J,
1964

"Dislocation Structures in Single Crystal
$Al_2 O_3$", J. Am. Ceram. Soc., 47, pp 81-86.

STEVENSON, D T and
KEYES, R J,
1955

"Measurement of Minority Carrier Lifetime
in Germanium and Silicon", J. Appl. Phys.,
26, 190-195.

SZE, S M,
1969

"Physics of Semiconductor Devices",
published by John Wiley and Sons, New York.

TERMAN, L H,
1962

"An Investigation of Surface States at a Silicon/Silicon Oxide Interface employing Metal-Oxide-Silicon Diodes", Sol. St. Electron., 5, pp 285-299.

THOMAS, C O, KAHNG, D and MANZ, R C,
1962

"Impurity Distributions in Epitaxial Silicon Films", J. Electrochem. Soc., 109, pp 1055-1061.

TSUNODA, Y,
1966

"Vapor Etching of Sapphire by H_2-$SiHCl_3$ Gas Mixtures", J. Phys. Soc. Japan, 21, pp 1220-1221.

von GRUBE, G SCHNEIDER, A, ESCH, U and FLAD, M,
1949

Z. Anorg. Chem., 260, p. 120

WAXMAN, A, HENRICH, V E, SHALLCROSS, F V, BORKAN, H and WEIMER, P K,
1965

"Electron Mobility Studies in Surface Space-Charge Layers in Vapor-Deposited CdS Films", J. Appl. Phys., 36, pp 168-175.

WEISBERG, L R,
1962

"Anomalous Mobility Effects in Some Semiconductors and Insulators", J. Appl. Phys., 33, pp 1817-21.

WEISBERG, L R and MILLER, E A,
1968

"Vacuum Deposition of Single-Crystalline Silicon-on-Sapphire", Trans. Met. Soc. AIME, 242, pp 479-484.

WHELAN, M V,
1965

"Influence of Charge Interactions on Capacitance versus Voltage Curves in MOS Structures", Philips Res. Repts., 20, pp 562-577.

WILCOCK, J D,
1971

"MOST Digital Logic Circuits in Silicon Films Deposited on Sapphire Substrates", Sol. St. Electron., 14, pp 315-325.

WRIGLEY, C Y and KROKO, L J,
1969

"Properties of the Silicon-Sapphire Interface in Heteroepitaxy", Semiconductor Silicon, pp 329-338, published by Electrochem. Soc., New York.

YASUDA, Y and OHMURA, Y,
1969

"Epitaxial Growth of Silicon Films Evaporated on Sapphire", Jap. J. Appl. Phys., 9, pp 1098-1106.

ZAININGER, K H and WARFIELD, G,
1965

"Limitations of the MOS Capacitance Method for the Determination of Semiconductor Surface Properties", IEEE Trans. Electron Devices, ED-12, pp 179-193.

ZAMINER, Ch,
1968

"Formation of Crystal Structure Defects in Silicon Layers by Epitaxial Precipitation on Mg-Al Spinel", Z. Angew Phys., 24, pp 223-227.

ZEMEL, J N,
1958

"Surface Transport Theory", Phys. Rev., 112, pp 762-765.

APPENDIX 1

FABRICATION PROCESS FOR HALL EFFECT DEVICES

1   Clean slice in hot nitric acid and wash in double-distilled
de-ionised water.

2   Remove residual surface oxide in 5% solution of hydrofluoric acid
and wash thoroughly in double distilled de-ionised water.

3   Grow approximately 4000 $\overset{o}{A}$ of silicon dioxide by oxidising the slice
for 90 minutes at 950°C in steam.

4   Cut the windows for the source, drain and contact diffusions using
KPR1 resist.

5   Remove the resist by immersing the slice in a mixture of 90% sulphuric
acid and 10% hydrogen peroxide for 2 minutes and then wash in double-distilled
de-ionised water.

6   Remove the residual oxide in the diffusion windows by immersing the
slice for 30 seconds in a 5% solution of hydrofluoric acid and then wash
in double-distilled de-ionised water.

7   Diffuse phosphorus from a phosphorus glass source in a closed bottle
for 15 minutes at 1000°C to give a sheet resistivity of approximately
12Ω per square.

8   Oxidise for 120 minutes at 950°C in steam to 'drive-in' the phosphorus
diffusion.

9   Remove all the oxide and wash in double-distilled de-ionised water.

10   Boil for 15 minutes in nitric acid to grow a thin layer of oxide which
will prevent phosphorus transfer from the diffused regions during the
next high temperature process.

11 Grow approximately 3000 $\overset{o}{A}$ of oxide by oxidising the slice for 60 minutes at 950°C in steam.

12 Cut the pattern for the silicon islands using KPR1 resist.

13 Remove resist and residual oxide as in steps 5 and 6.

14 Cut silicon islands using a 33% solution of sodium hydroxide at 65°C which gives an approximate etching rate of 3000 $\overset{o}{A}$/min.

15 Remove masking oxide and wash in double-distilled de-ionised water.

16 Boil for 15 minutes in nitric acid to prevent phosphorus transfer.

17 Transfer the slice to a water-cooled quartz reactor and grow 400 $\overset{o}{A}$ of oxide by oxidising the slice for 60 minutes at 950°C in dry oxygen.

18 Deposit 800 $\overset{o}{A}$ of silicon nitride in the same reactor by the nitridation of silane by ammonia at 850°C.

19 Deposit 2000 $\overset{o}{A}$ of silicon dioxide, still in the same reactor, by the oxidation of silane by carbon dioxide at 850°C.

20 Cut contact windows in the top oxide using KPR2 resist, which is more viscous than KPR1 and adheres better to the silicon island edges.

21 Remove resist as described in step 5.

22 Remove the silicon nitride in the contact windows by boiling under reflux in orthophosphoric acid and then remove the bottom oxide and wash in double-distilled de-ionised water.

23 Evaporate 1μm of aluminium from a tungstem filament while rocking the slice backwards and forwards in order to ensure good coverage on the edges of silicon islands.

24   Cut the aluminium using AZ 1350H resist.

25   Remove the resist as described in step 5.

26   Bake the slice for 15 minutes at $460^{o}C$ in air in order to improve

the contact resistance.

Unless otherwise stated the etch used for oxide removal is a buffered solution

of hydrofluoric acid containing.

      32g   Ammonia fluoride

      58ml   Distilled water and

      10ml   Hydrofluoric acid

and the etch for aluminium removal contains

      228ml   Phosphoric acid

        9ml   Nitric acid

      45ml   Acetic acid and

      15ml   Distilled water.

APPENDIX 2

DEPLETION APPROXIMATION

An analysis of the charge distribution in an MIS structure with an n-type substrate (Das 1969), shows that the charge in the silicon is given by.

$$Q_s = Q_o \, (2f \, (u_s))^{\frac{1}{2}} \qquad \cdots \cdots \cdots \text{(A1)}$$

where $Q_o = \left(\dfrac{\epsilon_{si}}{L_{D_i}}\right) \left(\dfrac{kT}{q}\right) = 1.23 \times 10^{-11}$ coulomb cm^{-2},

L_{D_i} is the intrinsic Debye length and

$$f(u_s) = \Big[\cosh (u_s - u_F) - \cosh u_F + u_s \sinh u_F \Big] \qquad \cdots \cdots \cdots \text{(A2)}$$

If $u_F < -4$, (ie the doping level is greater than 8×10^{11} cm^{-3}) then (A2) reduces to

$$2f(u_s) = \left\{ \exp(-u_F) \underbrace{[-u_s - 1}_{A} \underbrace{+ \exp(u_s)}_{B} \underbrace{+ \exp - (u_s - 2u_F)]}_{C} \right\} \qquad \cdots \cdots \text{(A3)}$$

The above expression may be separated into three parts as shown. Part B increases rapidly when the surface is accumulated, part C inreases rapidly when the surface is inverted and part A is predominant in the range.

$$-3 > u_s > 3 + 2 u_F$$

This range is the range of validity of the 'depletion approximation' and within the range virtually all the charge in the semiconductor is in the fixed positive charge of the ionised donor atoms in the depletion region.

$$Q_s = Q_o \left\{ [\exp (-u_F)] \, [-u_s - 1] \right\}^{\frac{1}{2}} \qquad \cdots \cdots \cdots \text{(A4)}$$

Now if $u_F \leqslant -4$ the first integration of Poissons equation gives

$$\frac{du}{dx} = \frac{\sqrt{2}}{L_D} \, [-u - 1 + \exp (u) + \exp - (u - 2u_F)]^{\frac{1}{2}} \qquad \cdots \cdots \cdots \text{(A5)}$$

(see equation (7) in Das 1969)

181

The depletion depth, x_d, is defined as the postion at which the Fermi level attains its bulk value, ie $u = 0$. Therefore integration of (A5) between the limits $u = u_s$, $x = 0$ and $u = 0$, $x = x_d$ gives

$$x_d = L_D \left\{ 2 \ (- u_s - 1) \right\}^{\frac{1}{2}} \qquad \dots\dots\dots\dots (A6)$$

From (A6) and (A4) it follows that

$$Q_s = q \ (N_D) \ x_d \qquad \dots\dots\dots\dots (A7)$$

ie the total charge in the semiconductor is due to the net ionised donor charge in the depletion region.

The general expression for the semiconductor capacitance is

$$C_s = - \frac{q}{kT} \ \frac{\partial Q_s}{\partial u_s}$$

$$= \frac{\epsilon si}{L_D} \ \frac{\left\{ 1 - \exp(u_s) + \exp - (u_s - 2u_F) \right\}}{\left\{ 2 \ [-u_s - 1 + \exp(u_s) + \exp - (u_s - 2u_F)] \right\}^{\frac{1}{2}}} \qquad \dots\dots\dots (A8)$$

which in the range of the 'depletion approximation' reduces to

$$C_s \approx \frac{\epsilon si}{L_D} \ \cdot \ \frac{1}{\left\{ 2 \ (-u_s - 1) \right\}^{\frac{1}{2}}} = \frac{\epsilon si}{x_d} \qquad \dots\dots\dots\dots (A9)$$

Therefore within the range of the 'depletion approximation' we have

$$\frac{1}{C} = \frac{1}{C_{ox}} + \frac{x_d}{\epsilon si} \qquad \dots\dots\dots\dots (A10)$$

provided that the capacitance due to fast states is small. In the case of a high frequency C-V plot the charge in the inversion layer cannot follow the ac signal. Therefore the inversion term can be deleted from (A8) and hence the equation (A10) is valid beyond the limit $u_s = 3 + 2u_F$. The other limit $u_s < -3$ still holds.

APPENDIX 3

FABRICATION PROCESS FOR N-CHANNEL JUNCTION FIELD-EFFECT DEVICES

1   Clean slice in hot nitric acid and wash in double-distilled
deionised water.

2   Remove residual surface oxide in 5% solution of hydrofluoric acid and
wash thoroughly in double-distilled deionised water.

3   Grow 5000 Å of silicon dioxide by oxidising the slice for 2 hours at
950°C in steam.

4   Cut the windows for the source and drain diffusion using KPRl resist.

5   Remove the resist by immersing the slice in a mixture of 90% sulphuric
acid and 10% hydrogen peroxide for 2 minutes and then wash in double-
distilled deionised water.

6   Remove the residual oxide in the diffusion windows by immersing for
30 seconds in a 5% solution of hydrofluoric acid and then wash in double-
distilled deionised water.

7   Diffuse phosphorus from a phosphorus glass source in a closed bottle for
10 minutes at 1000°C to give a sheet resistivity of approximately $20\Omega/\square$.

8   Oxidise the slice for 90 minutes at 950°C to drive in the phosphorus
diffusion.

9   Cut the pattern in the oxide for defining the gate region using KPRl resist.

10   Remove the resist as described in step 5.

11   Immerse the slice for 30 seconds in a 5% solution of hydrofluoric acid
to remove any residual oxide from the diffusion windows.

12 Diffuse boron from a diborane source for 30 minutes at $900^{\circ}C$ to give a sheet resistivity of approximately 50 $\Omega/\square$.

13 In order to remove the thin layer of 'boron rich' glass from the diffusion windows boil the slice for 10 minutes in nitric acid, wash in double-distilled deionised water and then immerse for 10 seconds in oxide etch before a further wash in double-distilled deionised water.

14 Oxidise the slice for 60 minutes at $900^{\circ}C$ in steam to drive in the boron diffusion.

15 Deposit 2000 $\overset{o}{A}$ of silicon dioxide by the oxidation of silane with carbon dioxide at $850^{\circ}C$.

16 Cut the pattern for the silicon islands using KPR1 resist.

17 Remove the resist as in step 5.

18 Immerse the slice for 30 seconds in a 5% solution of hydrofluoric acid to remove any residual oxide from the silicon areas which are to be etched.

19 Cut the silicon islands using a 33% solution of sodium hydroxide at $65^{\circ}C$ which gave an approximate etching rate of 3000 $\overset{o}{A}$/minute.

20 Remove the masking oxide and wash in double-distilled deionised water.

21 Boil for 10 minutes in concentrated nitric acid and then wash in double-distilled deionised water to grow a thin layer of oxide which will prevent phosphorus transfer from the phosphorus diffused regions during the next high temperature process.

22   Oxidise the slice for 45 minutes in steam at 900°C.

23   Deposit 1000 Å of silicon nitride by the nitridation of silane with ammonia at 850°C.

24   Deposit 2000 Å of silicon dioxide by the oxidation of silane with carbon dioxide at 850°C.

25   Cut the pattern in the top oxide for defining the contact windows. KP2 resist is used for this stage because it is more viscous and has superior adhesion to the edges of the silicon islands.

26   Remove resist as in step 5.

27   Continue the cutting of the contact windows through the silicon nitride using orthophosphoric acid boiling under reflux at 180°C.

28   Complete the cutting of the contact windows by cutting the bottom oxide and then wash in double-distilled deionised water.

29   Deposit 1 μm of aluminium from a tungsten filament while rocking the slice backwards and forwards in order to ensure good coverage on the edges of the silicon islands.

30   Cut the aluminium to define the contact pattern using AZ1350H resist.

31   Remove resist as in step 5.

32   Heat the slice for 15 minutes at 460°C in air in order to improve the contact resistance,

The oxide and aluminium etches used in the above processing sequence are the same as those used in Appendix 1.

A bulk silicon test-piece is included with the silicon-on-sapphire
slice at the phosphorus diffusion stage and the subsequent oxidation stage.
At the completion of the processing this test-piece is angle-lapped and
stained to enable a measurement of the junction depth to be made. The
junction depth was always in the range 0.45μm to 0.6μm.

APPENDIX 4

LIST OF SYMBOLS

A_g area of metal electrode on dielectric (cm^2)

A_J area of p-n junction (cm^2)

B strength of applied magnetic field (gauss)

C small signal series capacitance of MIS structure per unit area (F cm^{-2})

C_d depletion layer capacitance per unit area (F cm^{-2})

C_i equivalent capacitance of dielectric per unit area (F cm^{-2})

C_J capacitance of p-n junction (F)

C_S silicon space charge capacitance per unit area (F cm^{-2})

d depth of conducting channel (cm)

D_n diffusion constant for electrons $(\text{cm}^2\text{s}^{-1})$

D_p diffusion constant for holes $(\text{cm}^2\text{s}^{-1})$

E_{FN} quasi Fermi level for electrons (eV)

E_{FP} quasi Fermi level for holes (eV)

E_i intrinsic Fermi level (eV)

E_t energy level of traps (eV)

f occupancy factor of recombination centres

g_{do} field-effect transistor drain conductance limit as $V_D \rightarrow 0$ (mhos)

g_{max} field-effect transistor drain conductance limit as $V_D \rightarrow 0$ and $V_G \rightarrow 0$ (mhos)

G_s sheet conductivity of device channel (mhos/square)

I_C channel current (A)

I_D drain current of field-effect transistor (A)

k Boltzmann's constant $(\text{eV}/^\circ\text{K})$

L channel length (cm)

L_D $L_{Di}/(0.5 \exp u_F)^{0.5}$, extrinsic Debye length (cm)

L_{Di} $(\epsilon_{Si} k T/2q^2 n_i)^{0.5}$, intrinsic Debye length (cm)

l_k directional cosine

L_n $(D_n \tau_n)^{0.5}$, diffusion length for electrons (cm)

L_p $(D_p \tau_p)^{0.5}$, diffusion length for holes (cm)

m^* electron effective mass (kg)

m_o electron rest mass (kg)

n free electron concentration (cm^{-3})

N_D donor concentration (cm^{-3})

N_{Di} dislocation density (cm^{-2})

n_i intrinsic carrier concentration (cm^{-3})

n_o steady state electron concentration (cm^{-3})

n_p minority carrier electron concentration is p-type material (cm^{-3})

N_{ss} surface state density per eV (cm^{-2} eV^{-1})

N_{st} surface state density per unit area (cm^{-2})

N_T concentration of deep traps (cm^{-3})

$n(z)$ carrier concentration at position z (cm^{-3})

$\langle n_\square \rangle$ "effective number of carriers" per unit area in the channel (cm^{-2})

p probability of specular scattering

p_n minority carrier hole density in n-type material (cm^{-3})

p_o steady state hole concentration (cm^{-3})

Q charge per unit area on an electrode (C cm^{-2})

Q_i charge per unit area in the dielectric (C cm^{-2})

Q_M charge per unit area on the metal electrode (C cm^{-2})

Q_S charge per unit area in the silicon space-charge region (C cm^{-2})

Q_{SS} charge per unit area in surface states (C cm^{-2})

r ratio of Hall mobility to the conductivity mobility

R Hall constant (cm^3 C^{-1})

r_a	electron capture rate (s^{-1})
r_b	electron emission rate (s^{-1})
r_c	hole capture rate (s^{-1})
r_d	hole emission rate (s^{-1})
r_m	ratio of bulk mean free path to the extrinsic Debye length
R_t	terminal series resistance (ohms)
S	stress ($N\ m^{-2}$)
S_{ij}	compliance coefficient ($m^2\ N^{-1}$)
s_o	surface recombination velocity ($cm\ s^{-1}$)
t	film thickness (cm)
T	absolute temperature ($^{\circ}K$)
T_o	minority carrier lifetime (s)
T_t	MOS transient time constant (s)
U	net recombination rate (s^{-1})
u_F	reduced Fermi potential ($= q\phi_F/kT$)
u_s	reduced surface potential ($= q\phi_s/kT$)
Vb	built-in potential of a p-n junction (V)
V_D	drain voltage (V)
V_G	potential on metal electrode in any MIS structure (V)
V_H	Hall voltage (V)
V_J	potential across p-n junction (V)
V_K	the voltage at which the knee occurs in the drain characteristic of an MOS transistor
V_P	pinch-off voltage of junction field-effect transistor (V)
v_{th}	thermal velocity of carriers ($cm\ s^{-1}$)
w	channel width (cm)
x_d	depth of depletion region (cm)
x_{dmax}	maximum value of x_d (cm)
x_j	depth of p-n junction (cm)

x_t	distance between p-n junction and intersection of quasi Fermi level and trap level (cm)
Y	Youngs modulus (N m^{-2})
z	distance from the sapphire interface (cm)
α	coefficient of linear expansion ($^{\circ}$K^{-1})
α_m	magnetoconductive correction factor
λ	bulk mean free path (cm)
	$= Y_1 (1-v_2)/Y_2 (1-v_1)$
ϵ_{si}	dielectric constant of silicon $= 1.06 \times 10^{-12}$ F cm^{-1}
μ_B	bulk mobility (cm^2 v^{-1} s^{-1})
μ_D	dislocation mobility (cm^2 v^{-1} s^{-1})
$<\mu_H>$	'mean mobility' in the channel (cm^2 v^{-1} s^{-1})
μ_n	conductivity mobility for electrons (cm^2 v^{-1} s^{-1})
μ_s	surface mobility (cm^2 v^{-1} s^{-1})
μ_{sc}	space-charge mobility (cm^2 v^{-1} s^{-1})
$\mu(z)$	mobility at position z (cm^2 v^{-1} s^{-1})
v	Poissons ratio
σ	capture cross section (cm^2)
σ_n	capture cross section for electrons (cm^2)
σ_p	capture cross section for holes (cm^2)
σ_{ss}	capture cross section of surface states (cm^2)
τ	time constant for trap level (s)
τ_n	minority carrier lifetime for electrons (s)
τ_p	minority carrier lifetime for holes (s)
ϕ	electrostatic potential (eV)
ϕ_F	Fermi potential with respect to the bulk intrinsic value (eV)
ϕ_i	intrinsic Fermi potential (eV)
ϕ_{ms}	contact potential between the metal and the semiconductor (eV)

ϕ_n quasi Fermi potential for electrons (eV)

ϕ_p quasi Fermi potential for holes (eV)

ϕ_s surface potential with respect to the bulk (eV)

APPENDIX 5

WORK PUBLISHED IN CONJUNCTION WITH THIS THESIS

AN INVESTIGATION OF CARRIER
TRANSPORT IN THIN SILICON-ON-
SAPPHIRE FILMS USING MIS
DEEP DEPLETION HALL EFFECT
STRUCTURES

A B M ELLIOT AND
J C ANDERSON, SOLID-
STATE ELECTRONICS,
Vol 15, pp 531-545
1972.

AN INVESTIGATION OF CARRIER TRANSPORT IN THIN SILICON-ON-SAPPHIRE FILMS USING MIS DEEP DEPLETION HALL EFFECT STRUCTURES

A. B. M. ELLIOT

Post Office Research Dept, Dollis Hill, London, England

and

J. C. ANDERSON

Department of Electrical Engineering, Imperial College of Science and Technology, London, England

Abstract — A deep depletion MIS structure has been used to obtain measurements of the carrier concentration and mobility variations through thin epitaxial films, (0·6–2 μm), of *n*-type silicon on sapphire substrates. Both the mobility and the net donor concentrations decrease in the direction of the sapphire interface. The room temperature mobility in the material close to the sapphire interface is shown to be limited by scattering at charged defects but in accumulation layers adjacent to the oxide interface a phonon scattering limitation is observed and a large proportion of the oxide interface scattering is shown to be specular. Under certain conditions shallow donor states can be generated at the sapphire interface in densities of the order of 10^{12} cm^{-2}. A new technique for measuring the energy distribution of fast interface states at the oxide interface in a film with nonuniform doping is described. The energy distributions obtained on the silicon-on-sapphire films investigated in the present work are very similar to those obtained in bulk silicon by earlier workers.

Résumé — Une structure MIS de couche diélectrique basse a été utilisée pour mesurer les variations de densité et de mobilité des porteurs dans des couches épitaxiales minces, (0·6–2 μm), de silicium de type *n* sur saphir. La mobilité aussi bien que les concentrations nettes des donneurs diminuent dans la direction de l'interface du saphir. On montre que la mobilité à température ambiante dans le matériel près de l'interface du saphir est limitée par dispersion à des imperfections chargées, mais dans la région d'accumulation, contiguë à l'interface à oxyde on observe une limitation de la mobilité à cause d'une dispersion des phonons, et on montre qu'une grande proportion de la dispersion de l'interface à oxyde est spéculaire. On peut sous certaines conditions générer des états peu profonds des donneurs à l'interface du saphir aux densités par ordre de 10^{12} cm^{-3}. Une nouvelle méthode a été élaborée pour déterminer la distribution d'énergie des états interfaciales rapides à l'interface à oxyde dans une couche avec des distributions non-uniformes de dopant. Les distributions d'énergie obtenues sur les couches en silicium-sur-saphir que nous avons étudiées dans ces travaux, sont très semblables à celles trouvées pendant les recherches antérieures opérées dans le silicium de masse.

Zusammenfassung — Eine stark verarmte MIS Struktur wurde zur Gewinnung von Messungen der Schwankungen der Trägerkonzentration und Beweglichkeit über dünne epitaktische Schichten (0·6–2 μm) von *n*-Typ-Silizium auf Saphir benutzt. Die Beweglichkeit sowie die Netto donatorenkonzentrationen nehmen in der Richtung der Grenzfläche des Saphirs ab. Es wird gezeigt, dass die Raumtemperaturbeweglichkeit in der Grenzfläche des Saphirs enganliegenden Material durch Streuung bei geladenen Störstellen begrenzt wird, aber im Akkumulationsgebiet neben der Oxidgrenzfläche wird eine Begrenzung der Beweglichkeit durch Phononenstreuung beobachtet, und man zeigt, dass ein grosser Anteil, der Streuung der Oxidgrenzfläche spiegelnd ist. Unter gewissen Umständen können flache Donatorenzustände auf der Saphirgrenzfläche in einer Dichtenordnung von 10^{12} cm^{-3} erzeugen werden. Eine neue Methode zur Messung der Energieverteilung von schnellen Grenzflächenzustanden auf der Oxidgrenzfläche einer Schicht mit ungleichmässiger Verteilung der Dotierungsatome wird beschrieben. Die in dieser vorliegenden Arbeit beobachteten Energieverteilungen in Siliziumschichten auf Saphir stimmen im allgemeinen mit jenen von früheren Untersuchungen in Massiv-Silizium überein.

NOTATION

B	strength of magnetic field
C	small signal series capacitance per unit area of MIS structure
C_i	equivalent capacitance of insulator per unit area
C_s	silicon space charge capacitance per unit area
d	thickness of conducting channel
F_s	function of u_s and N_D, (see Fig. 4.7 of Ref. [16])
G_s	sheet conductivity of device channel
I_c	channel current
k	Boltzmann's constant
l	channel length
L_{Di}	$\left(\dfrac{\epsilon_{si}kT}{2q^2n_i}\right)^{1/2}$, intrinsic Debye Length
L_D	$\dfrac{L_{Di}}{(\frac{1}{2}\exp u_F)^{1/2}}$, extrinsic Debye length
m_o	electron rest mass
m^*	electron effective mass
n	electron concentration
n_i	intrinsic carrier concentration
$\langle n_\square \rangle$	'effective number of carriers' per unit area in the channel
$N(z)$	carrier concentration at position z
N_D	net donor concentration
N_{ss}	surface state density per eV
q	charge on the electron
Q_M	charge on the metal top electrode per unit area
Q_i	charge in the insulator per unit area
Q_s	charge in the silicon space-charge region per unit area
Q_{ss}	charge in surface states per unit area
r	ratio of the Hall mobility to the conductivity mobility
r_m	ratio of bulk mean free path to the extrinsic Debye length
R	Hall constant
T	absolute temperature
u_F	reduced Fermi potential ($= q\phi_F/kT$)
u_s	reduced surface potential ($= q\phi_s/kT$)
V_G	potential on metal electrode in any MIS structure
V_H	Hall voltage
x_d	depth of depletion region
x_{dmax}	maximum value of x_d
α	magnetoconductive correction factor
λ	bulk mean free path
ϵ_{si}	dielectric constant of silicon $= 1 \cdot 06 \times 10^{-12}$ F/cm
$\mu(z)$	mobility at position z
μ_s	surface mobility
μ_B	bulk mobility
$\langle \mu_H \rangle$	'mean mobility' in the channel
ϕ_F	Fermi potential with respect to bulk intrinsic value
ϕ_{MS}	contact potential between metal and semiconductor
ϕ_S	surface potential with respect to bulk
ω	probability of specular scattering

1. INTRODUCTION

THIN epitaxial films of silicon on sapphire substrates are used for the fabrication of high speed MOS integrated circuits [1–3]. Significant improvements over bulk silicon circuits can be achieved by using films only 1–2 μm thick and by diffusing all junctions through to the sapphire interface. The improvements result both directly from the reduction of junction capacitances and also from the ability to use gated silicon resistors [2] as load components in place of the conventional MOST load. Complementary circuits [3] are also simpler to fabricate with this technology. The characteristics of both the deep depletion type of MOS transistor [4] used in some complementary circuits and the gated resistor depend on the variation of carrier transport properties through the film. Therefore a detailed analysis of the variations is necessary to predict the performance of circuits using these components.

Two investigations of the variations of carrier density and mobility as a function of distance from the sapphire interface have been reported by Dumin and Robinson. In the first investigation [5] the films studied were autodoped p-type films from 14 to 28 μm thick and the measurements were made with 6 terminal Hall samples as the films were progressively thinned by polishing. The second investigation [6], which also used 6 terminal Hall samples, reported measurements on 5–7 μm thick n and p type films which were thinned to 1·5 μm by anodic oxidation. The results of the second investigation showed that for both types of film the net majority carrier concentrations and mobilities decreased as the sapphire interface was approached. The net majority carrier density decrease was believed to result from compensation by an increasing density of deep trapping levels associated with crystal defects near the sapphire interface.

The thinnest layers used in Dumin and Robinson's investigations were 1·5 μm thick. Since a layer of this thickness with a donor concentration of 5×10^{15} cm^{-3} has only $7·5 \times 10^{11}$ donors/cm^2, the trapping of carriers at both the silicon interfaces is likely to significantly influence the transport properties. Since the sapphire interface states will remain invariant during a set of measurements they will not influence the calculation of the properties of a removed layer unless the film is sufficiently thin for the space charge layer to penetrate the entire film. But the new air interface which is exposed at each removal step is unlikely to exhibit a constant surface state density and carrier density errors in excess of 10 per cent

at a film thickness of 1·5 μm are probable for fairly moderate surface state variations of 7·5 × 10¹⁰ cm⁻². Therefore, even if accurate thinning techniques could be developed for 1·5 μm thick films, the errors involved in using the conventional 6 terminal Hall sample technique could become very large as the sapphire interface is approached.

The use of a deep depletion MIS transistor structure enables the thickness of the conducting layer to be varied by applying the appropriate gate potential to drive a depletion region in from the oxide interface, thereby producing an upper boundary to the conducting layer which is in the bulk of the silicon and therefore free of interface states. The MIS Hall and conductivity measurements as a function of gate voltage are combined with $C-V$ measurements to provide both the carrier concentration and mobility profiles.

A new method is described of measuring the energy distribution of the fast interface states at the oxide interface on layers with non-uniform doping and the results obtained on typical silicon-on-sapphire films are presented. In the concluding section the influence of the bulk and interface properties measured in this investigation on the performance of silicon on sapphire MOST devices and integrated circuits is discussed.

2. EXPERIMENTAL TECHNIQUES

The structures of the deep depletion MIS Hall effect transistor and the MIS capacitor used in this investigation are shown in Fig. 1, [the two additional deep depletion transistors on the left of Fig. 1(a) were not used]. Both the channel length and width of the Hall effect transistor are 300 μm, as drawn on the photolithographic mask. Sideways diffusion and silicon undercutting will tend to reduce the length and width respectively, but the errors in both will be less than 1·5 per cent. The MIS dielectric is an oxide–nitride double layer with approximately 400 Å of oxide and 800 Å of nitride. Spreads of up to ±10 per cent in the capacitance/unit area of the dielectric are observed but this does not affect the accuracy of the calculations because the measured value of the dielectric capacitance in heavy accumulation is substituted in the analysis. The contacting fingers, numbered 1, 2 and 3 in Fig. 1(a) are diffused into the edge of the channel under the gate electrode to enable four terminal resistivity

and Hall measurements to be made. The source, drain and contact regions are all formed by a heavy phosphorus diffusion which is driven right through to the sapphire interface. The metallisation is aluminium.

In a deep depletion MIS device [4], (in contrast to the more usual surface inversion layer device), the source and drain are in ohmic contact with the bulk of the film and the conduction channel exists in the bulk of the film. This conduction channel is restricted by applying a potential to the gate of the device such as to drive a depletion region into the film, (e.g. a negative potential on an n-channel device). If the maximum depletion depth at heavy inversion for the particular film carrier concentration is greater than the film thickness then the device can be completely turned off.

The capacitor has the interdigitated structure shown in Fig. 1(c) in order to minimise the potential drop in the silicon electrode. When the depletion region is driven back into the silicon electrode of a conventional dot capacitor structure the spreading resistance of this electrode is usually sufficiently large to introduce significant errors in the capacitance measured at 1 MHz. The effect of the distributed resistance in the interdigitated capacitor in Fig. 1(c) was calculated by treating the structure as a simple $R-C$ transmission line. The analysis showed that with the dielectric thickness employed in this investigation the capacitance error is less than 2·5 per cent provided that the sheet resistance of the conducting layer below the depletion region is greater than 10⁵ Ω/square, which corresponds to a residual layer thickness of 0·1 μm at a film resistivity of 1 Ωcm. The area of the interdigitated capacitor, as defined on the photolithographic mask, is 4·912 × 10⁻¹ cm² but since the areas of processed capacitors were up to 7·5 per cent less than this value the dimensions of all the completed specimens used in this study were measured to an accuracy of better than 2 per cent. The silicon films were (100) orientation grown on (1̄102) sapphire substrates by thermal decomposition of silane*. All the films were phosphorus doped with net donor concentrations in the range 3 × 10¹⁵ cm⁻³ to 3 × 10¹⁶ cm⁻³. The film thicknesses after processing were in the range

---

*The films were supplied by the Allen Clark Research Centre of the Plessey Company.

0·6–1·7 μm, as measured on a Talysurf to an accuracy of $\pm 0·05 \mu$m.

The carrier transport properties of the films were derived from the following measurements made over a range of gate voltages from heavy inversion to heavy accumulation. Four-terminal resistivity measurements were made by using the source and drain for current contacts and points 1 and 2 in Fig. 1(a) for voltage contacts. The potential difference between points 1 and 2 was always within 5 per cent of the value expected from the outer terminal voltage, assuming zero contact resistance at the source and drain. Evidence that the small differences that did exist were probably due to small non-uniformities in film resistivity rather than contact resistance was obtained from the fact that the difference always decreased when the device was driven into heavy accumulation where the contact resistance error would be expected to be at its greatest. The Hall voltages were obtained by using a digital voltmeter with an accuracy of $\pm 10 \mu$V to measure the change in voltage between the source and contact 3 when the magnetic field was reversed. Typical Hall voltages were in the range 0·25–3·0 mV. The linearity of the Hall voltage with magnetic field was checked up to 6000 G and all subsequent measurements were made at 4460 G.

Referring to Fig. 1(a), the sheet conductivity of the device channel is given by

$$G_s = \frac{s}{l} \cdot \frac{I_c}{V_{12}} \, \Omega^{-1} \tag{1}$$

where I_c is the channel current and V_{12} is the potential difference between the voltage contacts. The Hall coefficient is given by

$$R = 10^8 V_{II} \frac{d}{I_c B} \cdot \frac{1}{\alpha} \, \text{cm}^3 \text{C}^{-1} \tag{2}$$

where d = thickness of conducting channel, V_{II} is the Hall voltage and α is a magnetoconductive correction factor for the shorting of the Hall field by the source and drain electrodes[7]. For an infinitely long specimen $\alpha = 1$ and for the square geometry used in this work $\alpha = 0·68$. Equations (1) and (2) can be combined by using the relationships

$$R = \frac{1}{nq} \tag{3}$$

and

$$\sigma = nq \langle \mu_{II} \rangle = \frac{G_s}{d} \tag{4}$$

to give

$$\langle \mu_{II} \rangle = \frac{s}{l} \cdot \frac{V_{II}}{V_{12}} \cdot \frac{10^8}{B\alpha} \, \text{cm}^2 \text{V}^{-1} \, \text{sec}^{-1} \tag{5}$$

and

$$\langle n_{\square} \rangle = nd = \frac{\alpha I_c B}{q V_{II}} \times 10^{-8} \, \text{cm}^{-2} \tag{6}$$

where $\langle \mu_{II} \rangle$ = 'mean mobility' in the channel and $\langle n_{\square} \rangle$ = 'effective number of carriers per cm²' in the channel. Note that in equations (3) and (4) it has been assumed that the conductivity mobility equals the Hall mobility.

In order to estimate the uncertainties in $\langle \mu_{II} \rangle$ and $\langle n_{\square} \rangle$ consider first the uncertainties in the individual parameters which appear in equations (5) and (6). Maximum errors in I_c and V_{II} are less than 2 and 4 per cent respectively and the uncertainty in the s/l ratio is less than 1 per cent. The maximum error in α due to the 1·5 per cent uncertainty in the square geometry is less than 1 per cent. The uncertainties in V_{12} and B are insignificant, ($< 0·3$ per cent). Using the above uncertainties, examination of equations (5) and (6) shows that the maximum errors in $\langle \mu_{II} \rangle$ and $\langle n_{\square} \rangle$ are both less than 7·5 per cent and that the major part of this is due to the uncertainty in V_{II}. The relatively large uncertainties of 4 per cent in V_{II} only exist when the depletion region is driven deep into the film. For all measurements where $\langle \mu_{II} \rangle$ is greater than 100 cm² V^{-1}sec^{-1} the uncertainty in V_{II} is less than 1 per cent and the maximum possible errors in $\langle \mu_{II} \rangle$ and $\langle n_{\square} \rangle$ less than 4·5 per cent. Non-uniformities in sheet resistivity could give rise to further errors in $\langle \mu_{II} \rangle$ and $\langle n_{\square} \rangle$ but the sample used in this investigation is more than an order of magnitude smaller than typical 6 terminal Hall samples and therefore long-range non uniformities are likely to be less significant. The use of four-terminal resistivity measurements to verify the absence of large localised inhomogeneities by checking the constancy of the potential gradient along the sample has already been referred to.

In Section 3 a method of calculating the true average mobility and total carrier density per cm²
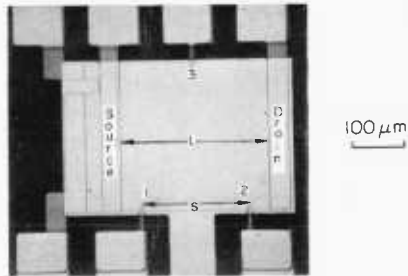
(a) Plan view of Hall effect device

Sapphire
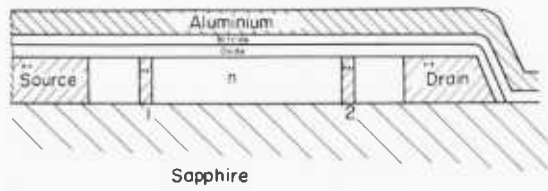
(b) Cross section of Hall effect device

(c) Plan view of capacitor

Fig. 1. Structures of the deep depletion MIS Hall effect transistor and the MIS capacitor. (a) Plan view of Hall effect transistor. (b) Cross-section of Hall effect transistor. (c) Plan view of capacitor.

from the 'mean mobility' and 'effective carrier density' is described.

Hall measurements alone simply enable the mobility and carrier density to be calculated as a function of the gate voltage. C-V measurements are made at a frequency of 1 MHz on the adjacent interdigitated capacitor to determine the relationship between the gate voltage, V_G, and the depletion depth, x_d. The C-V measurements on a typical device were checked at a number of points over the full range of bias voltages on a 100 kHz bridge and all the results were within 3 per cent of the 1 MHz values, confirming that the interdigitated capacitor is not significantly affected at 1 MHz by the distributed resistance in the silicon electrode. In the range of the 'depletion approximation' [8], where the reduced surface potential $u_s < -3$ for an n-type specimen, the value of x_d can be obtained from a high-frequency C-V plot by using the equation

$$\frac{1}{C} = \frac{1}{C_i} + \frac{x_d}{\epsilon_{si}} \qquad (7)$$

where C and C_i are the capacitance per cm² of the MIS capacitor and dielectric respectively and $\epsilon_{si} = 1.06 \times 10^{-12}$ F cm⁻¹ is the dielectric constant for silicon. The uncertainty in x_d is determined mainly by the uncertainty in the area of the capacitor and is therefore of the order of 3 per cent.

The major limitation of this method is that under equilibrium conditions there is a maximum depth for the depletion region, which is given by the equation

$$x_{d\max} = 2\left(\frac{\epsilon_{si}}{qN_D}\right)^{1/2} \left\{\frac{kT}{q} \ln\left(\frac{N_D}{n_i}\right)\right\}^{1/2}$$

$$= 2L_D\left\{\ln\left(\frac{N_D}{n_i}\right)\right\}^{1/2} \qquad (8)$$

where L_D is the extrinsic Debye length. Typical values for $x_{d\max}$ are approximately 1 μm at $N_D = 10^{15}$ cm⁻³ and 0.1 μm at $N_D = 10^{17}$ cm⁻³.

All the measurements were made in a conventional long-tailed cryostat with automatic temperature control accurate to ±1°K. The cryostat was used to enable the temperature dependence of mobility to be determined and to facilitate the measurement of 80°K C-V plots for the capacitors.

3. MOBILITY AND CARRIER CONCENTRATION

The majority of the specimens examined in this investigation exhibited room temperature variations of $\langle \mu_H \rangle$ and $\langle n_\square \rangle$ of the form shown in Fig. 2. (exceptions to this behaviour will be discussed in Section 5). The maximum mobility over a range of wafers varied from 275 to 420 cm² V⁻¹ sec⁻¹ with spreads of up to ±7 per cent occurring across a single wafer. The value of $\langle n_\square \rangle$ at flatband varied

Specimen 505/1/9
Film thickness = 1·1 μm
T = 298°K
Mean net donor density ≈ 3×10¹⁵ cm⁻³
—✕— Experimental results
- - - - Theoretical relationship for ω=0
△ Theoretical results for ω=0·895

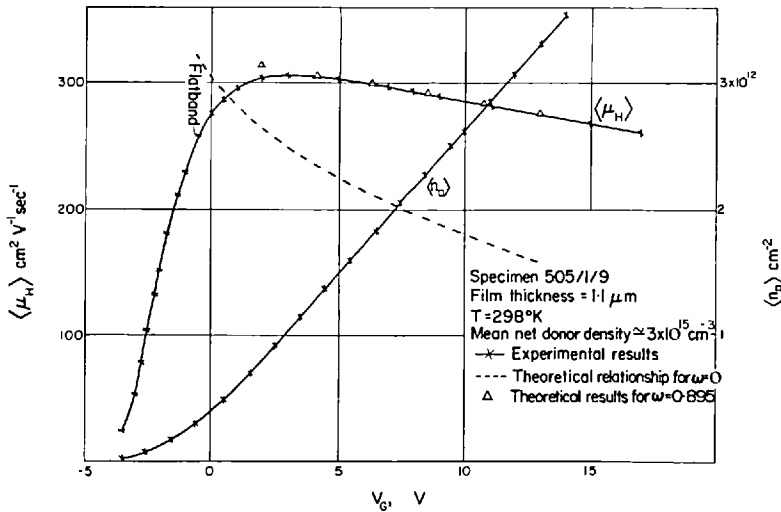Fig. 2. 'Mean mobility' and 'effective number of carriers per cm² of channel' as a function of gate voltage for a typical device.

by up to ± 12 per cent across a wafer. In deciding on the magnitude of the source to drain voltage to be used for these measurements a compromise must be reached between the conflicting requirements of having a large Hall voltage to minimise measurement errors and of simultaneously ensuring a relatively small change in surface potential along the device. A value of 300 mV was chosen which gives a Hall voltage measurement accuracy of better than 4 per cent under worst conditions. The effect of the surface potential variation along the channel is most significant when the position of the depletion layer edge is changing rapidly with surface potential. The above drain voltage of 300 mV leads to a maximum variation in depletion depth along the channel of 500 Å, near the intrinsic surface condition. At the heavy accumulation and heavy inversion limits the errors arising from surface potential variations are small.

If the carrier concentrations and mobility vary in the direction perpendicular to the surface of the film, then $\langle \mu_{H} \rangle$ and $\langle n_{\square} \rangle$ are not the true average mobility and total number of carriers per cm² respectively. In this case the values of $\langle \mu_{H} \rangle$ and $\langle n_{\square} \rangle$ are given by[9]

$$\langle n_{\square} \rangle = \frac{d}{Rq} = \frac{\left(\int_{0}^{d} N(z) \mu(z) \, dz \right)^{2}}{\int_{0}^{d} N(z) \mu^{2}(z) \, dz} \qquad (9)$$

and

$$\langle \mu_{H} \rangle = \frac{\int_{0}^{d} N(z) \mu^{2}(z) \, dz}{\int_{0}^{d} N(z) \mu(z) \, dz} \qquad (10)$$

where $N(z)$ is the carrier concentration per cm³ at a distance z from the sapphire interface.

When the mobility is independent of position the above equations reduce to

$$\langle \mu_{H} \rangle = \mu \qquad (11)$$

and

$$\langle n_{\square} \rangle = \int_{0}^{d} N(z) \, dz. \qquad (12)$$

Information on the carrier transport properties of the silicon films is obtained by using equations (9) to (12) to analyse the accumulation and depletion region results in Fig. 2. In the heavy accumula-

tion region the majority of the carrier transport is in a thin region near the oxide interface where it would be reasonable to assume that the mobility is constant with respect to position and therefore equations (11) and (12) can be used. Since the gate capacitance is equal to the dielectric capacitance in this region, the total carrier density will vary linearly with gate voltage and its expected rate of change can be calculated from a direct capacitance measurement in heavy accumulation after removal from the cryostat, (the stray parallel capacitance in the three terminal direct measurement is less than 0·1 pF, i.e. less than 1 per cent of the dielectric capacitance). A comparison between this calculated rate of change of total charge density with gate voltage and the rate of change of free charge density with gate voltage as obtained from Hall measurements shows that they agree to ±5 per cent. A similar comparison of inversion layer charge densities in bulk silicon by Fowler, Fang and Hochberg[10] resulted in an equally close agreement. Fang and Fowler[11] concluded from this agreement that there was little trapping of electrons in the corresponding range of surface potentials and that r, the ratio of the Hall mobility to the conductivity mobility is constant, but it would also appear to signify that r is approximately equal to unity. Since heavily accumulated surfaces are degenerate the relaxation time is expected to be constant[12] and Zemel[13] has shown that for a constant relaxation time the theoretical value of r is 0·87 for diffuse surface scattering and unity for specular surface scattering. Because the maximum errors in both charge densities are less than 4·5 per cent and the experimental values of r are within 5 per cent of unity the difference between this and 0·87 is believed to be significant and probably indicates a large proportion of specular scattering at the oxide interface. An estimate of the probability of specular scattering is obtained from the following analysis of the mobility variation with gate voltage in the accumulation region.

The mobility in the accumulation region slowly decreases as the gate voltage increases as shown in Fig. 2. Similar type of behaviour has been reported for accumulation layers in bulk silicon by Reddi[14] and also for inversion layers by Murphy, Berz and Flinn[15] and by Fang and Fowler[11]. The decrease is probably due to enhanced oxide-interface scattering since at higher gate voltages

there will be more interface collisions both because of the large electric field perpendicular to the surface and because a larger fraction of the carriers are within a mean free path of the oxide interface.

The results of Reddi, which were shown to be basically in agreement with a diffuse scattering mechanism at the oxide interface, exhibit a more rapid decrease of mobility with increasing surface field than the present results. In order to test whether the differences could be attributed to a greater probability of specular scattering the present results were analysed on the basis of the approximate equation for partially specular scattering derived by Many, Goldstein and Grover[16].

$$\frac{\mu_s}{\mu_B} = \frac{1}{1 + (1 - \omega)\left(\frac{r_m F_s}{u_s}\right)(1 + u_s)^{1/2}} \quad (13)$$

where μ_s and μ_B are surface and bulk mobility respectively, u_s is the surface potential in units of kT, ω is the probability that an electron reaching the surface will be specularly reflected, F_s is a function of surface potential and the net donor density (see Fig. 4.7 in Ref. [16]) and r_m is the ratio of the bulk mean free path to the extrinsic Debye length, given by the equation

$$r_m = \frac{\lambda}{L_{Di}} = \mu_B \left(\frac{m^* N_D}{2\pi \epsilon_{si}}\right)^{1/2} \quad (14)$$

where N_D is the net donor density and m^* is the electron effective mass, $(= 0.26 \, m_0)$. Equation (13) has been shown to be a good approximation to a more rigorous numerical analysis both for entirely diffuse scattering[17] and also for partially specular scattering at flatband[16]. In order to evaluate equation (13) values must be determined for μ_B and r_m. Although μ_B is not constant with respect to depth in silicon-on-sapphire films an approximate value appropriate to the material near the oxide interface can be obtained from Fig. 2 in the following manner. At flatband equation (13) reduces to

$$\frac{\mu_s}{\mu_B} = 1 - (1 - \omega)r_m. \quad (15)$$

Since it shall be shown that $(1 - \omega)r_m < 5 \times 10^{-3}$ the theoretical value of μ_s at flatband is within 0.5 per cent of μ_B. Unfortunately at flatband it is

no longer true that the majority of the carrier transport is in a layer close to the oxide interface and in fact it is seen from Fig. 2 that the value of $\langle \mu_H \rangle$ has started to decrease due to significant contributions from the lower mobility regions of the silicon film closer to the sapphire interface. However an approximate estimate for μ_s at flatband can be obtained by extrapolating the value of $\langle \mu_H \rangle$ in heavy accumulation back to the flatband voltage. The value obtained in this way from Fig. 2 was 326 cm² V⁻¹ sec⁻¹ which is only 25 per cent of the good single-crystal bulk silicon value at the same impurity concentration. Although the net donor concentration also varies as a function of depth in the film, a mean value of 3×10^{15} cm⁻² obtained from Fig. 3 can be used to calculate a value of 3.4×10^{-2} for r_m from equation (14). The value of u_s at any given gate voltage is obtained from the corresponding $\langle n_\square \rangle$ value in Fig. 2 by using the graphical relationship between surface charge and surface potential derived by Whelan[18]. Substitution of $\omega = 0$ and the above values of μ_B and r_m in equation (13) generates the theoretical mobility-gate voltage relationship for entirely diffuse surface scattering shown by the dashed line in Fig. 2, which lies well below the experimental points and changes much more rapidly with gate voltage. Whereas, if ω is taken as 0.885 then the excellent agreement between the experimental results and the theoretical prediction shown in Fig. 2 is obtained. This confirms the high probability of specular scattering suspected from the r value of unity.

Near the flatband condition not only are N and μ functions of z but also the function $N(z)$ varies with the gate voltage, hence the results are very difficult to analyse in this region. As the gate voltage is reduced and the interface conditions move from accumulation into depletion the slope of the $\langle n_\square \rangle$ vs. V_G plot decreases. A qualitative explanation for this is that for a given change in gate voltage the mean position of the modulated charge moves back from the oxide interface (this also exhibits itself as a decrease in capacitance.) and furthermore part of the change in surface charge may be a change in trapped charge which does not contribute to $\langle n_\square \rangle$.

Since inversion layer charge does not contribute to the channel current, in the range $u_s < -3$ the charge density in the channel is modulated solely by the movement of the edge of the depletion
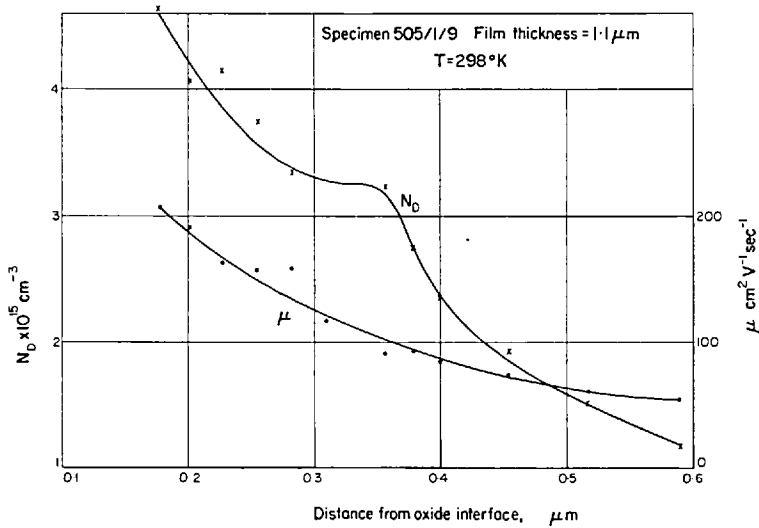
Fig. 3. Mobility and carrier concentration as a function of distance from the oxide interface.

region, [see for example equation (10) of Das[19]). Hence the functions $N(z)$ and $\mu(z)$ in equations (9) and (10) are independent of gate voltage and only the limits of the integrations are changed.

Let $\langle n_{01} \rangle$ signify that the integrations are between the limits 0 and d_1, $\langle n_{02} \rangle$ between the limits 0 and d_2 etc.

From equations (9) and (10)

$$\langle n_{01} \rangle \langle \mu_{01} \rangle = \int_0^{d_1} N(z)\mu(z) \, \mathrm{d}z$$

and

$$\langle n_{01} \rangle \langle \mu_{01} \rangle^2 = \int_0^{d_1} N(z)\mu^2(z) \, \mathrm{d}z.$$

Therefore

$$\int_{d_2}^{d_1} N(z)\mu(z) \, \mathrm{d}z = \langle n_{01} \rangle \langle \mu_{01} \rangle - \langle n_{02} \rangle \langle \mu_{02} \rangle \qquad (16)$$

and

$$\int_{d_2}^{d_1} N(z)\mu^2(z) \, \mathrm{d}z = \langle n_{01} \rangle \langle \mu_{01} \rangle^2 - \langle n_{02} \rangle \langle \mu_{02} \rangle^2. \qquad (17)$$

In the limit where $d_1 - d_2$ is small equation (10) gives

$$\mu(\overline{d_{12}}) = \frac{\langle n_{01} \rangle \langle \mu_{01} \rangle^2 - \langle n_{02} \rangle \langle \mu_{02} \rangle^2}{\langle n_{01} \rangle \langle \mu_{01} \rangle - \langle n_{02} \rangle \langle \mu_{02} \rangle} \qquad (18)$$

where $\overline{d_{12}}$ signifies the value mid-way between d_1 and d_2.

From equation (9) it also follows that

$$N(\overline{d_{12}}) = \frac{((n_{01}) \langle \mu_{01} \rangle - \langle n_{02} \rangle (\mu_{02}))^2}{\langle n_{01} \rangle \langle \mu_{01} \rangle^2 - \langle n_{02} \rangle \langle \mu_{02} \rangle^2} \cdot \frac{1}{\Delta x_d} \qquad (19)$$

where Δx_d = increase in the depletion depth.

Typical carrier concentration and mobility profiles calculated from the results shown in Fig. 2 using equations (7), (18) and (19) are presented in Fig. 3. Since the difference terms in equations (18) and (19) are similar in magnitude to the maximum possible errors in the individual parameters these calculations would be virtually meaningless if the errors were random. But because the errors are largely systematic an _a priori_ estimate of the inaccuracies in $N(\overline{d_{12}})$ and $\mu(\overline{d_{12}})$ is not possible and the validity of the analysis must be judged by comparison with other independent measurements and on the repeatability of the results. Figure 3 shows a tendency for the net donor concentration to decrease towards the sapphire interface, which is probably due to compensation by deep acceptor levels of the type reported by Heiman[20] and Dumin[21]. All the wafers used in this investigation had received post-growth heat treatments of 15·5 hr at 1100–1200°C in a gas stream of 25 per cent oxygen in nitrogen in order to remove aluminium by redistribution into the growing oxide film. According to the analysis of Ross and Warfield[22] this should reduce the

aluminium level below 10^{14} cm^{-3}. The mean doping level for the maximum depletion depth calculated from the C_{max}/C_{min} ratio of the interdigitated capacitor[23] was 3.5×10^{15} cm^{-3} which is in good agreement with the results in Fig. 3.

The mobility decreases steadily in the direction of the sapphire interface towards a limiting value of approximately 50 cm^2 V^{-1} sec^{-1}. The long high temperature post-growth heat treatment would be expected to generate aluminium–oxygen complexes[22] which may be the scattering centres responsible for this low interface mobility. Extrapolation of the mobility results in Fig. 3 to the oxide interface would give a value in good agreement with the estimate of 326 cm^2 V^{-1} sec^{-1} used for μ_B in the analysis of the accumulation region.

4. TEMPERATURE DEPENDENCE OF MOBILITY

The temperature dependence of the Hall mobility in the heavy accumulation region was investigated at a gate voltage of $+10$ V and a drain voltage of 300 mV. The 'mean mobility' was measured as the specimens were cooled down from room temperature to below 100°K and as they were allowed to warm back up to room temperature, with the temperature controlled to ± 1°K during each measurement. The results for

a single specimen all fitted on the single curve labelled (A), shown in Fig. 4. In the vicinity of 300°K the slope of the curve approaches conformity with a $T^{-1.5}$ law, as has also been reported for electron inversion layers in bulk silicon[11], which probably indicates a predominance of phonon scattering. This phonon scattering must be a bulk mechanism since Greene[24] has shown that surface electron–phonon interactions produce no diffusivity when the electron gas is degenerate, as is the case in heavy accumulation. This conclusion is also consistent with the high specular reflection coefficient of 0·885 measured in the previous section. As the temperature is lowered the mobility increases less rapidly than for pure phonon scattering and below 130°K it begins to decrease with decreasing temperature, suggesting a predominance of charged impurity scattering. This type of behaviour has been observed in bulk silicon[25] and in thin silicon-on-sapphire films[26] and may arise simply from the different temperature dependence of the two scattering mechanisms. The transition from phonon scattering to impurity scattering occurs at a low mobility and relatively high temperature for a carrier concentration of 3×10^{15} cm^{-3} which probably indicates the presence of a much higher density of charged scattering centres due to compensation

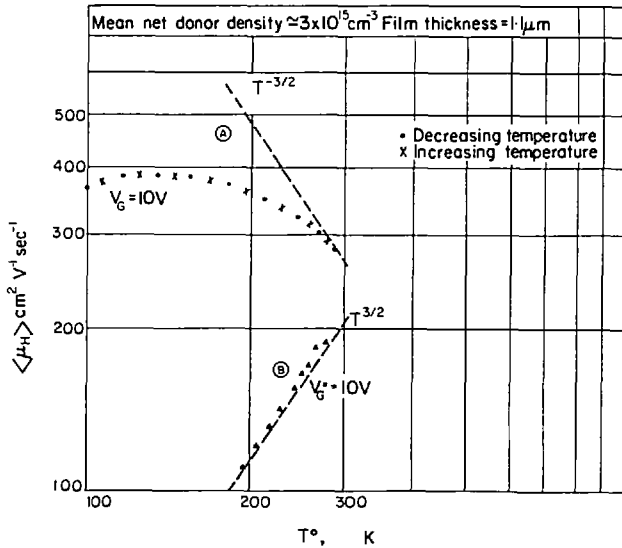Fig. 4. Temperature dependence of mobility. Curve A: $V_G = +10$ V, the surface is heavily accumulated. Curve B: $V_G = -10$ V, the surface is inverted and $x_d = x_{d\,max}$.

by deep acceptors[26]. Another factor in the case of an accumulated surface is the temperature dependence of the penetration depth of the accumulation layer. At a given surface potential the penetration depth is proportional to the extrinsic Debye length, which is in turn proportional to $T^{1/2}$ and inversely proportional to the carrier density in the bulk to the half power. Below 130°K the carrier density decreases exponentially with temperature, outweighing the $T^{1/2}$ term and giving rise to a rapid increase in the penetration depth of the accumulation layer, thereby increasing the ratio of charged impurity scattering events to phonon scattering events. The $T^{1.5}$ dependence of the $V_G = -10$ V curve (B) in Fig. 4 is evidence of the predominance of charged impurity scattering as the sapphire interface is approached. At $V_G = -10$ V the surface is inverted and conduction between the source and drain can only occur through the 0·5 μm region of the film below the depletion layer. This measurement is complicated by the long time constants required to reach equilibrium at room temperature[27] but the problem can be overcome by heavily inverting the surface at room temperature and taking all the measurements as the temperature is lowered. Then the attainment of an equilibrium state is not limited by the speed of the generation process. The measurements were restricted to the temperature range 200–300°K

because below 200°K there is appreciable deionisation of the donors and the depletion region would be driven back towards the sapphire interface. Hence the $\langle \mu_H \rangle$ value would be an 'effective mean' over a different depth of channel. The $T^{1.5}$ temperature dependence for carrier transport near the sapphire interface indicates that although space charge scattering from neutral aluminium–oxygen complexes[22] probably exists it is not the major scattering mechanism in this region because a $T^{-0.5}$ dependence is to be expected for space charge scattering[28] in a temperature range where the carrier concentration is not changing significantly.

5. ELECTRONIC INTERFACE LAYERS

Deviations from the general behaviour described in the previous section were observed on a number of specimens. At room temperature they exhibited $\langle \mu_H \rangle$ plots which changed more rapidly with V_G and $\langle n_\square \rangle$ plots which tended to level off below the flatband value, some specimens even showed a gradual increase in $\langle n_\square \rangle$ as the gate voltage was made more negative. A typical set of results is shown in Fig. 5. At lower temperatures the minimum in the $\langle n_\square \rangle$ plot is even more marked.

As the negative voltage on the gate is increased, the total negative charge in the channel must decrease. Therefore the minimum in the $\langle n_\square \rangle$ plot
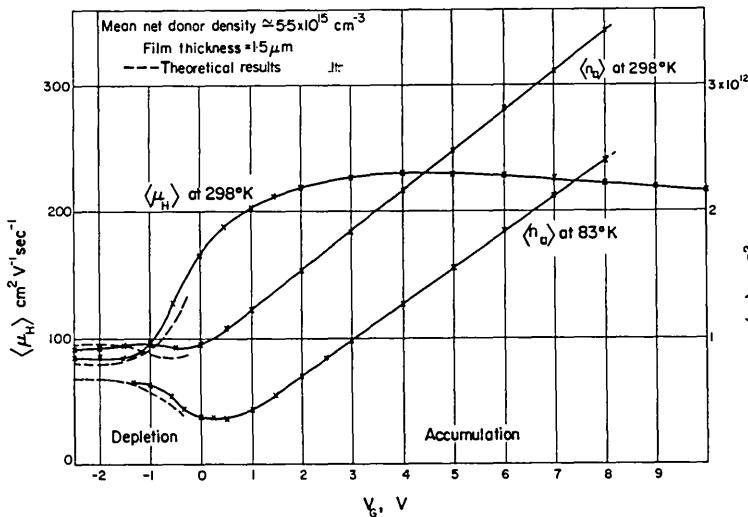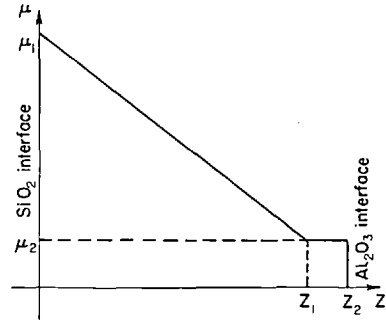

Fig. 5. Experimental and theoretical 'mean mobility' and 'effective number of carriers per cm²' as a function of gate voltage for a device with a high density of donor interface states.
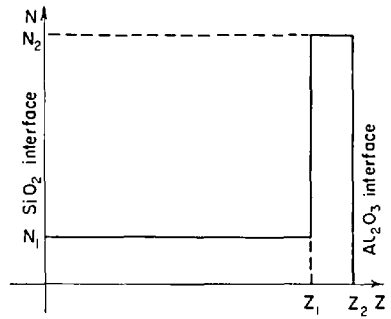
must be an anomaly of the averaging effect. If a high density of low mobility electrons is present at the sapphire interface the denominator in equation (9), which contains a μ^2 term, will decrease rapidly as the depletion region is driven back through the higher mobility regions near the oxide interface. It may decrease more rapidly than the numerator causing the $\langle n_\square \rangle$ value to increase with decreasing V_G. All the wafers had received a 30 min heat treatment at 460°C in a moist atmosphere and it is possible that shallow donor states generated at the sapphire interface were responsible for the high density of low mobility electrons since Heiman[20] has reported that 15 min heating at 500°C in a moist atmosphere is sufficient to generate interface donor states on a 'contaminated' wafer, although the contamination responsible has not yet been identified. Wrigley and Kroko[29] have pointed out that the experimental evidence could equally well be explained by the generation of positive interface charge.

Simple mobility and carrier concentration profiles of the type shown in Fig. 6 were assumed for the specimen in an attempt to explain the results in Fig. 5. Values of the parameters in Fig. 6 must be determined before an analysis can be undertaken. The value of N_1 was taken as the mean donor concentration in the depletion range as calculated from the C_{max}/C_{min} ratio of the capacitor C-V plot. The thickness of the film was measured at $1\cdot5$ μm using a Talysurf. The value of the mobility at the oxide interface was taken as the extrapolation of the 'mean mobility' in the accumulation region to flatband, which is a valid approximation since $r_{in} \ll 1$. Since the interface region between $Z = Z_1$ and $Z = Z_2$ is likely to be small compared with the film thickness, its exact dimensions are not important but the total charge in it and the mobility of the carriers are. These are two unknown quantities, although it is clear that the mobility in this region will be less than the minimum value of the 'mean mobility'.

A range of values was tried for the two unknowns to see if a reasonable fit with the experi-

(a) Mobility profile

(b) Carrier concentration profile

Fig. 6. Model for donor interface states calculation. (a) Mobility profile. (b) Carrier concentration profile.

mental results could be obtained by analysing the model in Fig. 6 using equations (9) and (10).

The parameters in Table 1 give theoretical results, (shown by the dashed lines) which are in good agreement with the experimental results, (shown by solid lines) in Fig. 5 even to the extent of duplicating the shallow dip in the $\langle n_\square \rangle$ plot.

The $\langle n_\square \rangle$ points agree to better than 8 per cent and the $\langle \mu_H \rangle$ points to better than 11 per cent. The above results give an interface donor state density of 9×10^{11} cm^{-2} and the value was always within 7×10^{11} cm^{-2} to $2\cdot5 \times 10^{12}$ cm^{-2} which lies within the 10^{11} to 10^{13} range reported by Heiman [20]. The sapphire interface mobility of 30 cm^2 V^{-1}

Table 1

Mobility (cm^2 V^{-1} sec^{-1})		Distance (cm)		Carrier density (cm^{-3})	
μ_1	μ_2	Z_1	Z_2	N_1	N_2
230	30	$1\cdot49 \times 10^{-4}$	$1\cdot50 \times 10^{-4}$	$5\cdot5 \times 10^{15}$	$9\cdot0 \times 10^{17}$

sec^{-1} is less than the 100 cm^2 V^{-1} sec^{-1} reported by Wrigley and Kroko[29], who measured electron field-effect mobilities in inversion layers adjacent to the sapphire interface by using 75 μm thick sapphire substrates as the dielectric in MIS transistors. Since the films used by Wrigley and Kroko had an average electron mobility of 400 cm^2 V^{-1} sec^{-1} they may have had a lower overall density of crystal defects than those used in this investigation. The difference between the 30 cm^2 V^{-1} sec^{-1} interface mobility for films with electronic layers at the sapphire interface and the 50 cm^2 V^{-1} sec^{-1} value for films without interface inversion layers is not really significant because of the approximate nature of the above analysis, but a reduction is to be expected due to enhanced scattering at ionised donor centres on the Heiman model or due to enhanced interface scattering on the Wrigley and Kroko model if the sapphire interface is not entirely specular.

As the temperature is reduced the carrier density in the bulk of the film reduces but not at the sapphire interface because the donor density in the Heiman model is approaching degeneracy and the electron density in the Wrigley and Kroko model is determined by the positive interface charge. This increase in the ratio of interface carrier density to the bulk carrier density explains the larger dip in the $\langle n_\square \rangle$ plot at lower temperatures and when the appropriate 83°K values are substituted in the profile model the results obtained are again in good agreement with the experimental observations, as shown in Fig. 5.

Even with this simple model it has been possible to explain all the features of the unusual behaviour of this type of specimen. Although an accurate measurement of the density of interface donor states has not been obtained, an estimate of the approximate magnitude has been possible.

6. FAST STATE DENSITY AT THE OXIDE INTERFACE

The conventional high frequency[30, 31], low frequency[32] and quasi-static techniques[33] for determining the surface state density as a function of position in the band-gap all require the comparison of an experimental curve with an ideal curve. The ideal curves can readily be computed if the substrate doping is known and uniform but these techniques cannot be applied to substrates with non-uniform and unknown doping

profiles such as exist with silicon-on-sapphire films. Hall measurements can be used in the following manner to provide the additional information necessary to analyse the high-frequency C–V curves for the case of non-uniform substrates.

The charge neutrality condition for an MIS capacitor gives the relationship

$$Q_M + Q_{SS} + Q_S + Q_i = 0 \tag{20}$$

where Q_M = charge in the gate metal, Q_{SS} = charge in surface states, Q_S = charge in the semiconductor and Q_i = charge in the insulator.

When the gate voltage is changed the first three terms in equation (20) may change but Q_i is independent of V_G. The value of ΔQ_M for an incremental change ΔV_G in the gate voltage can be determined from the MIS capacitance since

$$C = \frac{dQ_M}{dV_G} \text{ therefore } \Delta Q_M = C\,\Delta V_G.$$

Hence if ΔQ_S can be evaluated, the value of ΔQ_{SS} follows simply from

$$\Delta Q_{SS} = -(\Delta Q_M + \Delta Q_S). \tag{21}$$

Since ΔQ_S for a non-uniform substrate cannot be evaluated theoretically it must be measured. In the range of the 'depletion approximation', where $-3 > u_s > 2u_F + 3$, the ΔQ_S is entirely due to ionised donors in the depletion region and it follows from equation (19) that

$$\Delta Q_S = q \frac{(\langle n_{01} \rangle \langle \mu_{01} \rangle - \langle n_{02} \rangle \langle \mu_{02} \rangle)^2}{\langle n_{01} \rangle \langle \mu_{01} \rangle^2 - \langle n_{02} \rangle \langle \mu_{02} \rangle^2}. \tag{22}$$

Hence in the region of the 'depletion approximation' by substitution of ΔQ_S and ΔQ_M in equation (21) the value of ΔQ_{SS} for any ΔV_G can be calculated. The relationship between V_G and ϕ_S is

$$V_G - \phi_{MS} - \phi_S = -\frac{(Q_{SS} + Q_i)}{C_i} \tag{23}$$

and from equation (21)

$$\Delta \phi_S = \Delta V_G + \frac{\Delta Q_M}{C_i}. \tag{24}$$

A plot of $N_{SS} = \Delta Q_{SS}/q\Delta \phi_S$ as a function of ϕ_S,

with an arbitrary zero for ϕ_S, can now be determined. The approximate position of the true ϕ_S zero, (i.e. flatband), is readily determined from the C-V plot by the method of Lehovec[34], although strictly speaking that analysis also assumes a uniform doping level in the substrate.

A typical set of N_{SS} results is shown in Fig. 7. These results are very similar in form to those obtained by Kuhn[33] on bulk silicon and the values are approximately 50 per cent lower.

By using the Brown and Gray[35] technique, C-V measurements on the capacitor at 300°K and 80°K enable a value for the total number of interface states between the 300°K Fermi level and the 80°K Fermi level to be calculated. These Fermi levels for the specimen in Fig. 7, ($N_D = 1.5 \times 10^{16}$ cm^{-3}) are at 0·21 eV and 0·05 eV respectively from the conduction band edge. Hence the $\Delta Q_{SS}/q$ value of 2.5×10^{11} cm^{-2}, measured by the Brown and Gray experiment corresponds to a mean N_{SS} value of 1.55×10^{12} cm^{-2} eV^{-1} which is well above the mid-gap values obtained by the above technique. This rapid increase of the interface state density near the band edge is a feature of all the reported work on the Si–SiO$_2$ interface. Clearly neither the high compressive stress nor the high density of crystal faults in silicon-on-sapphire films substantially effects the density or energy distribution of silicon–oxide interface states.

7. DISCUSSION AND CONCLUSIONS

The carrier transport properties of thin silicon films on sapphire vary significantly through the thickness of the film with the majority of the films having the following characteristics. Firstly the following rather anomalous transport behaviour at the oxide interface must be explained. The electron mobility maxima in the light accumulation region are in the range 275 cm^2 V^{-1} sec^{-1} to 420 cm^2 V^{-1} sec^{-1} and since the r_m values for all the films are. below 5×10^{-2} these mobility values which are only 22–34 per cent of good single crystal bulk mobility, should be a fairly good indication of film quality near the oxide interface. But in contrast to this result the channel mobilities in conventional inversion-layer p-channel MOST's fabricated in similar films are close to the channel mobilities obtained in bulk silicon[2]. Boleky[3], who obtained channel mobilities in p-channel MOST's which were even slightly greater than corresponding bulk silicon MOST's concluded that either the quality of the top surface of the film is very comparable to that of good single-crystal bulk silicon or that the inversion layer mobility is relatively independent of crystalline quality. Since the present results indicate that the former explanation is not the reason the second explanation is examined on the basis of Grover, Goldstein and Many's[36] development of the
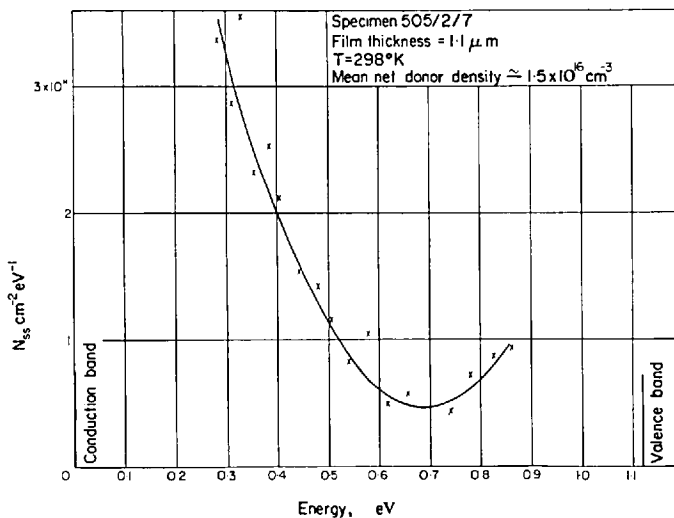

Fig. 7. Energy distribution of fast states at the oxide interface.

Schrieffer surface transport theory. Although this analysis assumes diffuse surface scattering it is considered because a complete analysis of minority carrier transport is inversion layers for partially specular scattering has not yet been developed. The effect of partially specular scattering is discussed qualitatively later.

At a typical carrier concentration of 5×10^{15} cm^{-3} the bulk hole mobility for good quality single crystals is 470 cm^2 V^{-1} sec^{-1} which from equation (14) gives a value of 0·063 for r_m. Substituting these values into the analysis of Grover Goldstein and Many a value of 0·45 is obtained for the ratio of surface mobility to bulk mobility, (μ_s/μ_B), at a surface potential $u_s = -2u_F - 6$, which corresponds to a typical 'on-state' for an MOST device. This gives a surface mobility of 211·5 cm^2 V^{-1} sec^{-1}. If the bulk hole mobility is reduced to 30 per cent of the good single-crystal value, (which is the measured reduction for the bulk electron mobility near the oxide interface), then the value for r_m is reduced to 0·019 and consequently the μ_s/μ_B ratio is increased to 0·72. This corresponds to a surface mobility of 101·5 cm^2 V^{-1} sec^{-1}, which is 48 per cent of the good single-crystal value. Although the increase in the μ_s/μ_B ratio partially compensates for the μ_B reduction the theoretically predicted differences in surface mobility between silicon-on-sapphire and good single-crystal devices are still significantly greater than the experimentally measured differences. Furthermore, if partially specular scattering occurs at the oxide interface in the inversion layer, the μ_s/μ_B ratios will be larger and the predicted μ_s differences between bulk and silicon-on-sapphire films will be even greater. Therefore some other explanation for the similarity in the channel mobilities must be found.

Schlotterer[37] has shown that the compressive stress in silicon-on-sapphire films, which is of the order of 8×10^9 dynes/cm^2 and is due to the thermal expansion mismatch between the silicon and sapphire can give rise to electron mobility reductions of approximately 50 per cent in (100) films on (1$\bar{1}$02) sapphire at a carrier concentration of 5×10^{15} cm^{-2} whereas the same stress level has little effect on hole mobilities for this orientation of film. Therefore it is probable that the bulk hole mobility for silicon-on-sapphire material adjacent to the oxide interface is not reduced by nearly as much as the bulk electron mobility and that this is

at least part of the reason for the relatively high channel mobilities in silicon-on-sapphire p-channel MOST. This hypothesis is supported by the fact that Boleky[3] did observe a reduction in the channel mobility of n-channel silicon-on-sapphire MOST's to 70 per cent of the value obtained in good bulk silicon devices. The piezoresistive reduction of electron mobility also helps to explain why the accumulation layer mobility exhibits a $T^{-1·5}$ temperature dependence, which is typical of phonon scattering, although the mobility values are much less than good single-crystal silicon. A reduction of mobility of this magnitude due to charged or neutral crystal defects would not be compatible with a $T^{-1·5}$ temperature dependence.

Silicon-on-sapphire devices have the same form of energy distribution of interface states in the band gap as bulk silicon devices[33] and the density of states for similar processing procedures is of the same order of magnitude.

Moving away from the oxide interface towards the sapphire interface the defect properties of the silicon-on-sapphire material become more important. The net donor concentration decreases as the sapphire interface is approaahced. This is probably due to compensation by an increasing concentration of deep acceptor levels since the aluminium has been gettered by long thermal oxidation processes. The steady decrease of mobility and its $T^{1·5}$ temperature dependence near the sapphire interface indicates that in this region charged defect scattering is probably predominant. The defects may be either impurities or appropriate crystal structure defects which decrease in density as the film growth proceeds[38]. Any device which operates by carrier transport near the sapphire interface will have its characteristics limited by the defects in the silicon-on-sapphire film. Therefore the deep depletion MOS transistors of the type reported by Heiman[4] and Boleky[3] are likely to have low channel mobilities near the threshold voltage. Boleky[3] has measured a channel mobility of 140 cm^2 V^{-1} sec^{-1} for a deep depletion p-channel MOST in light accumulation, 2 V above threshold, but no values closer to the threshold have been reported.

The long term stability of silicon-on-sapphire devices is suspect. If interface electronic layers an be generated in a few minutes at 460°C they may well be generated at significant rates at lower temperatures. This will cause large leakage

currents to develop in conventional n-channel devices and the breakdown voltage of the drain diode to decrease in p-channel devices. Further investigation is required both to determine the activation energy for this process and to define precisely the 'clean' conditions necessary to prevent the electronic interface layer generation.

REFERENCES

1. E. C. Ross and C. W. Mueller, *IEEE Trans. Electron Devices*, 13, 379 (1966).
2. J. D. Wilcock, *Solid-St. Electron.* 14, 315 (1971).
3. E. J. Boleky, *RCA Rev.* 31, 372 (1970).
4. F. P. Heiman, *IEEE Trans. Electron Devices*, 13, 855 (1966).
5. D. J. Dumin and P. H. Robinson, *J. electrochem Soc.* 113, 469 (1966).
6. D. J. Dumin and P. H. Robinson, *J. Crystal Growth*, 3, 214 (1968).
7. I. Isenberg, B. R. Russell and R. F. Greene, *Rev. Scient. Instrum.* 19, 685 (1948).
8. A. S. Grove, *Physics and Technology of Semiconductor Devices*, p. 267, John Wiley, New York (1967).
9. R. L. Petritz, *Phys. Rev.* 110, 1254 (1958).
10. A. B. Fowler, F. Fang and F. Hochberg, *IBM Jl. Res. Dev.* 8, 427 (1964).
11. F. Fang and A. B. Fowler, *Phys. Rev.* 169, 619 (1968).
12. R. A. Smith, *Wave Mechanics of Crystalline Solids*, p. 333, Chapman and Hall, London (1961).
13. J. N. Zemel, *Phys. Rev.* 112, 762 (1958).
14. V. G. P. Reddi, *IEEE Trans. Electron Devices*, 15, 151 (1968).
15. N. St. J. Murphy, F. Berz and I. Flinn, *Solid-St. Electron.* 12, 775 (1969).
16. A. Many, Y. Goldstein and N. B. Grover, *Semiconductor Surfaces* p. 312, North-Holland, Amsterdam (1965).
17. Y. Goldstein, N. B. Grover, A. Many and R. F. Greene, *J. appl. Phys.* 32, 2540 (1961).
18. M. V. Whelan, *Philips Res. Rep.* 20, 620 (1965).
19. M. B. Das, *Solid-St. Electron.* 12, 305 (1969).
20. F. P. Heiman, *Appl. Phys. Lett.* 11, 132 (1967).
21. D. J. Dumin, *Solid-St. Electron.* 13, 415 (1970).
22. E. C. Ross and G. Warfield, *J. appl. Phys.* 40, 2339 (1969).
23. B. E. Deal, A. S. Grove, E. H. Snow and C. T. Sah, *J. electrochem. Soc.* 112, 308 (1965).
24. R. F. Greene, *Surface Sci.* 2, 101, (1964).
25. F. J. Morin and J. P. Maita, *Phys. Rev.* 96, 28 (1954).
26. D. J. Dumin and E. C. Ross, *J. appl. Phys.* 41, 3139 (1970).
27. A. Goetzberger and E. H. Nicollain, *Bell Syst. tech. J.* 46, 513 (1967).
28. L. R. Weisberg, *J. appl. Phys.* 33, 1817 (1962).
29. C. Y. Wrigley and L. J. Kroko, *Semiconductor Silicon*, p. 329, Electrochem. Soc., New York (1969).
30. L. M. Terman, *Solid-St. Electron.* 5, 285 (1962).
31. K. H. Zainiger and G. Warfield, *IEEE Trans. Electron Devices* 12, 179 (1965).
32. C. N. Berglund, *IEEE Trans. Electron Devices*, 13, 107 (1966).
33. M. Kuhn, *Solid-St. Electron.* 13, 873 (1970).
34. K. Lehovec, *Solid-St. Electron.* 11, 135 (1968).
35. D. M. Brown and P. V. Gray, *J. electrochem. Soc.* 115, 760 (1968).
36. N. R. Grover, Y. Goldstein and A. Many, *J. appl. Phys.* 32, 2538 (1961).
37. H. Schlotterer, *Solid-St. Electron.* 11, 947 (1968).
38. D. W. Pashley, *Advn. Phys.* 14, 327 (1965).