



Point defects and diffusion in thin films of GaAs

R.M. Cohen

Material Science and Engineering Department, University of Utah, 160 S Central Campus DR Rm 304, Salt Lake City, UT 84112-0560, USA

Received 6 September 1996; accepted 12 March 1997

Abstract

Impurity diffusivity in thin films of GaAs is affected by native defect concentrations which were grown into the film, and which enter the film from both the substrate and the external ambient. Variations in measured diffusivity are related to native defects whose concentrations gradually relax to equilibrium values via different kinetic pathways. The time required for native defects to equilibrate depends upon the experimental design which in turn determines which region of the phase diagram that an annealing experiment begins and ends in. The Gibbs phase rule is applied to several commonly used experimental designs to show why the equilibrium native defect concentrations are generally not defined solely by temperature. The necessary conditions which define an equilibrium state, and the sufficient conditions which determine whether that equilibrium state can be approximated within a short time, are explicitly discussed. Experiments begun far from equilibrium are often associated with unusually high or low time-dependent diffusivities, but the results are often irreproducible as the native defect concentrations drift for extended periods of time. Experiments begun close to equilibrium are generally associated with reproducible diffusivities because the native defect concentrations reach constant values in a time which is short compared with the anneal period. Some diffusion results can be adequately described by equilibrium models, and the so-called 'Fermi-level effect' model is shown to apply as a special case of solid-vapor equilibrium. Near-equilibrium results from several groups provide strong evidence that a Ga vacancy, with a charge of -1, controls group III element marker diffusion in n-type, intrinsic, and p-type GaAs at $T > 800^\circ\text{C}$. By relating an anneal ambient to a region of the ternary phase diagram, it becomes clear why the group II-related diffusion and other variables are extremely sensitive to the experimental design. Pinning of the Fermi energy at the surface during vapor phase epitaxy appears to explain why nonequilibrium concentrations of dopants and charged point defects can be grown into GaAs and affect diffusion in post-growth anneals. The relatively large random measurement noise often accompanying interdiffusion appears to be largely associated with the presence of Al in GaAs. We conclude that this behavior is associated with a residual contaminant, most likely oxygen. It also appears likely that small amounts of an oxidant have been present in many closed ampoule anneals and affected the reported interdiffusion results. We conclude that poorly understood metallurgical reactions are probably responsible for the enormous range of interdiffusion observed in device structures annealed under SiO_2 and Si_3N_4 glass caps. © 1997 Elsevier Science S.A.

Keywords: Native point defect; Diffusion mechanisms; Interdiffusion; Fermi energy pinning; Dopant solubility; Point defect solubility; Vacancy; Interstitial; Dielectric cap; Experimental design

1. Introduction

Control over atomic diffusion in thin films is important for the processing of electronic and electro-optic devices. This review shall discuss several key materials issues which are related to diffusion in thin films of GaAs which have been grown by modern epitaxial methods. We shall show that diffusion in thin films and bulk GaAs differs because (1) both the vapor and the substrate may act as a source or sink of native defects, i.e. vacancies and interstitials, for an extended period of time, (2) Fermi level pinning during epitaxial growth causes the concentration of all charged species, i.e. impurities and native point defects, to deviate significantly from their equilibrium values during and after growth, and (3) the dopant solubility limit may easily be exceeded in a film during vapor phase epitaxy. In order to keep our discussion relatively simple, we shall avoid discussing the diffusion measured after ion implantation. An excellent review by Deppe and Holonyak [1] published a decade

ago began to address some of the important issues associated with diffusion in GaAs thin films. In addition, some of the fundamental issues associated with diffusion and native defects in bulk GaAs have been nicely reviewed by Casey and Pearson [2], Kendall [3], and Shaw [4]. Tan and Gosele [5] and Cohen [6] have interpreted several reported diffusion results in terms of a 'Fermi-level effect' model. The applicability of these models will be clarified in this review.

We shall address the issue of why a unique equilibrium state needs to be clearly defined, and how an equilibrium model can be applied to diffusional mass transport, i.e. to an inherently nonequilibrium process. It will be shown that impurity transport can be correctly described by an equilibrium model when the concentration of appropriate native defects are close to their equilibrium value. We shall discuss the importance of a properly designed experiment, i.e. one which meets both the *necessary* and the *sufficient* conditions for the native defect concentrations to approach their equilibrium values, in order to use an equilibrium interpretation of the data. It shall also become clear that the so-called 'Fermi-level effect' model is applicable only as a special case when the native defect concentrations have equilibrated with an external phase, and when the experimental design has allowed this equilibration to occur at a sufficiently rapid rate.

It will be repeatedly emphasized that thermodynamics only relates the equilibrium concentration of atoms in different phases, and that thermodynamics can never be used to deduce a particular diffusion mechanism or reaction pathway. Indeed, the equilibrium concentration of any species is independent of whether the solid-vapor equilibrium is reached via vacancies, interstitials, or a cosmic wormhole. To reach firm conclusions about the kinetics of diffusion, i.e. to determine which native defect mediates the diffusion process, requires that controlled nonequilibrium defect concentrations be used. In order to interpret the results from such experiments as a perturbation upon an equilibrium state, one must begin with a very clear understanding of which equilibrium state (out of several possible) correspond to the chosen experimental design.

Thus, this review will begin by developing several important equilibrium relationships between solid and vapor. Section 2 will develop equations applicable to the native defects in intrinsic (undoped) GaAs, and relate those to properly designed experiments. Section 3 will develop equations applicable to the native defects in extrinsic (doped) GaAs, and application made to appropriately designed experiments. Some of those relationships may be familiar, but some are shown here for the first time. Section 4 will discuss evidence for Fermi level pinning at the growing GaAs surface and several practical results of that pinning. Section 5 will discuss the results obtained from several well-defined annealing experiments and we shall interpret several features about the point defects involved in diffusion. Section 6 will summarize the current picture of diffusion and the implications for controlling diffusion in modern epitaxial structures.

2. Equilibrium in intrinsic GaAs

Kroger [7] summarized several useful relationships between carrier concentration and point defect concentrations in bulk GaAs more than 20 years ago. These basic concepts provide the foundation for charged point defect models [6,8] of diffusion. However, it is clear that such models do not explain many observed results, and that there is significant confusion over the basic description of diffusion mechanisms. Poor experimental technique may account for some of the variation in reported diffusivities, but the results summarized in Section 5 strongly suggest that other sources of error have crept into the measurements. These include (1) application of an equilibrium model to highly non-equilibrium experiments, (2) errors in the application of an equilibrium model to near-equilibrium

experiments, and (3) the appearance of unanticipated variables which affect the results because of the particular experimental design.

An element of time implicitly enters any consideration of equilibrium. For example, a pn junction reaches true equilibrium only when donor and acceptor atoms have uniformly redistributed themselves throughout the entire volume of the semiconductor, i.e. when there is no longer a pn junction. However, when describing device operation, we make a convenient equilibrium approximation which is remarkably useful: we assume that the n- and p-type regions of the semiconductor are two distinct material phases with the ions frozen into each phase and insoluble in the other, and that the fast moving species (electrons and holes) equilibrate quickly across the space charge region. Similarly, we shall consider several experiments here in which the native defects move much more rapidly than the ions which set the Fermi energy. In such cases, it may be appropriate to model diffusion assuming that the native defect concentrations have approached their equilibrium values in a time which is short compared with the time of the experiment. Of course, experiments are required to determine whether or not native defect equilibrium actually can be approximated. We shall see that the results of several experiments indicate that this assumption is sometimes acceptable in uniformly doped layers which do not contain group II impurities, and for experimental designs which do not interfere with the process of point defect equilibration. When native defect isoconcentration can be approximated, then interdiffusion measurements may provide useful fundamental information. It will also be shown that when group II impurities are present, any interdiffusion is almost always associated with strong nonequilibrium conditions and an equilibrium model cannot be used to interpret the results.

This review will discuss evidence which shows that a few experiments do exist in which a native defect equilibrium approximation appears justifiable. While useful information can be obtained from these experiments, it will become clear that such experiments are fundamentally limited, i.e. they can *never* distinguish between a vacancy and interstitial mechanism without additional information. To fully understand the nature of the point defects controlling diffusion, it is necessary to perform controlled nonequilibrium experiments. Since nature will always bring a nonequilibrium experiment back to one equilibrium state (out of several possible), it is important to have a clear understanding of the equilibrium state which will result from the experimental conditions chosen. We emphasize that the commonly used words ‘in equilibrium’ are a shortened form of the phrase ‘in equilibrium with’. It is *fundamentally incorrect* to discuss GaAs as simply in equilibrium with itself, and it is *always* necessary to consider its equilibrium state as defined with respect to at least one other distinct material phase, i.e. vapor, liquid, and/or solid. To appreciate the reasons for the above statements, and their implications for a well designed experiment, it is necessary to discuss the Gibbs phase rule.

2.1. The Gibbs phase rule applied to annealing experiments

Any discussion of equilibrium in a semiconductor isolated from the rest of the world is of little value for both practical and theoretical reasons. We shall show this below by applying the Gibbs phase rule to some simple material systems. The phase rule is a calculation of the number of degrees of freedom, F , i.e. the number of intrinsic thermodynamic variables, which are available to be set (and which must be set) if one is to control the equilibrium state of a system.

The phase rule comes about by considering a system consisting of a total number, C , of elements, and a total number, P , of distinct material phases (vapor, liquid, and solid). Within any given phase in this system, one might adjust the composition for each of the C elements in that phase. Since one may do this in each of the P phases, the total number of variables is the product, $CP + 2$ (temperature and total pressure give the 2). However, the number of independent variables will be less than $(CP + 2)$ because several equations of constraint exist for two reasons: (1) In each phase, the total

composition must add up to 100%. Thus, $(CP + 2)$ must be reduced by 1 for each of the P phases. (2) The composition across phases are related because the chemical potential of each element is equal in each phase at equilibrium. Since there will be $(P - 1)$ such equations for each element present, the number of independent variables will be reduced by $C(P - 1)$. Thus, the number of available degrees of freedom are $F = (CP + 2) - P - C(P - 1) = C - P + 2$.

Control of exactly F independent thermodynamic variables is a *necessary* condition if one intends to define a unique equilibrium state for the system which can be approached quickly and reproducibly. For example, a system consisting of a solid phase in contact with its own vapor phase gives $F = C$. In this system, $F = 1$ for an elemental semiconductor. Setting the temperature defines the equilibrium state of all the remaining variables in the entire system, i.e. vapor pressure, vacancy concentration, etc. Setting any fewer variables obviously leaves the system undefined. Of course, it is physically possible to control more than $F = 1$ variables in an elemental solid + vapor system, say, by setting the temperature and pumping on the vapor to keep the total pressure below the vapor pressure of the solid. However, if the total pressure remains below the vapor pressure, then the solid will lose mass by sublimation until it disappears. Indeed, as seen in explicit examples below for GaAs, controlling either too few or too many variables generally causes not only visible changes in the solid, but also large uninterpretable changes in the measured diffusivity.

However, a choice of F variables is not a *sufficient* condition to assure that the system actually approaches a desired equilibrium state in a useful time. One must also design an experiment to make it possible for the dependent thermodynamic variables to approach their equilibrium values in a reasonably short time. For example, an elemental semiconductor cannot equilibrate when left exposed to flowing gas at high temperatures in an open tube system because the elemental partial pressure above the solid could never reach its equilibrium vapor pressure. However, an elemental semiconductor wafer may be placed face-to-face with an identical wafer (often called a proximity cap) and the partial pressure will closely approach the vapor pressure. The proximity cap does not eliminate sublimation, but rather it slows sublimation by limiting atomic diffusion through the small cross-sectional area between the two wafers out to the nearby open tube. We shall encounter evidence below that rapid sublimation can have a large effect on diffusion, and that a proximity cap is a very useful tool. Indeed, for all semiconductors, it is important that both the *necessary* and *sufficient* conditions for equilibrium be met if one intends to control the equilibrium concentrations of the native defects and to approach those concentrations in the shortest possible time. The alternative will invariably lead to results which are irreproducible and uninterpretable within a simple model.

In the remainder of Section 2, we shall compare the classic closed ampoule with the open tube systems commonly used for annealing today, and discuss how each may meet the sufficient condition for approximating a well-defined equilibrium state in GaAs. We shall develop the first expressions for the equilibrium concentration of impurities and native defects which depend directly upon the *degrees of freedom which are actually controlled*, discuss deviations from equilibrium in capped semiconductors, and describe some simple models of diffusion. Except where otherwise stated, simplifying assumptions used throughout this review include the ideal gas law and unity activity coefficients.

2.1.1. Some relevant examples

The phase rule is intimately related to the fact that the chemical potential of each species in each phase must be constant for a system to be at equilibrium. The chemical potential of a species, μ , relates the change in Gibbs free energy of the system, G , to a change in the number of molecules of that species, n , i.e. $\mu = (\partial G / \partial n)_{T,P}$, at constant temperature and pressure. Under a total pressure of 1 atm, a pure substance in a condensed phase is generally considered to be in its standard state and its chemical potential is designated $\mu^{\text{Std}}(T)$.

If equilibrium exists between an element contained in a solid phase, s, and the same element in the vapor, v, then one may describe that equilibrium by

$$\mu_s^{\text{Std}}(T) + kT \ln a_s = \mu_v^{\text{Std}}(T) + kT \ln a_v \quad (1)$$

where a_s and a_v represent the activity of the element in the solid and vapor phases, respectively. Since we are assuming that the ideal gas law and unity activity coefficients apply, then the concentration, N , of the element in an ideal solid solution can be directly related to its activity in the vapor, $a_v = P/P^{\text{Std}}$ via,

$$\mu_s^{\text{Std}}(T) + kT \ln N = \mu_v^{\text{Std}}(T) + kT \ln(P/P^{\text{Std}}) \quad (2)$$

where $P^{\text{Std}} = P^\circ(T)$, the vapor pressure of the element, if $P^\circ(T) < 1 \text{ atm}$. If $P^\circ(T) > 1 \text{ atm}$ at the chosen temperature, then the condensed phase of the pure substance does not exist at 1 atm, and the standard state is customarily chosen to be $P^{\text{Std}} = 1 \text{ atm}$ for the pure gas.

Standard conditions, i.e. a total pressure of 1 atm, will not be attained if the vapor pressure of each component is low. However, one can always add an inert gas to bring the total pressure up to $P_{\text{tot}} = 1 \text{ atm}$. The additional degree of freedom arising from the addition of this new component (the inert gas) is immediately satisfied because the partial pressure of that gas is directly controlled and the remaining degrees of freedom are unaffected by any chemical reaction with the inert gas. The $P dV$ work done on the solid in raising the total pressure in the system from P° to 1 atm can be shown to be negligibly small because of the low compressibility of condensed phases. Since an inert gas has no significant effect on a system, the addition of an inert gas to reach 'standard' conditions is often ignored. We too shall ignore the inert gas which is usually present in open tube systems except when it is required to make a practical point.

2.1.1.1. Elemental As

Elemental As is worth considering because it provides a simple example for the application of the phase rule and it is an integral component of GaAs. We shall begin by considering a system consisting only of As in the solid and vapor phases, i.e. $F = 1$. Setting one variable is *necessary* to determine the equilibrium values of all other thermodynamic variables. For example, if we choose $T = 900 \text{ K}$, then the necessary condition to reach equilibrium is satisfied. However, setting $T = 900 \text{ K}$ is not a *sufficient* condition to bring the solid to equilibrium. For example, someone must have designed the experiment to make it *possible* for the As partial pressure to reach its equilibrium vapor pressure, which is $P_{\text{As}}^\circ \approx 1.7 \text{ atm}$.

If a small lump of As had been placed inside of a very large closed ampoule, then the solid would sublime before P_{As}° was attained. In this case, we would end up with only a vapor phase, and thus $P = 1$ and $F = 2$. $F = 2$ demands that 2 degrees of freedom be controlled to define the state of the system. In this case, we inadvertently did that by selecting both the temperature and the number of moles of arsenic per unit volume in the ampoule (i.e. the final As partial pressure). This illustrates a pair of general observations related to a closed system: (1) the number of condensed phases usually changes when the wrong number of degrees of freedom are controlled, and (2) after a sufficiently long time, nature does bring a system to equilibrium, although it may be one that is unexpected.

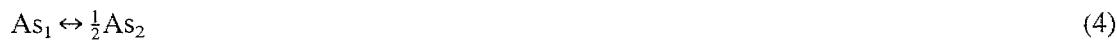
In an open system, if we set T and attempt to keep $P_{\text{As}} < P_{\text{As}}^\circ$, then the solid also sublimes away. If we attempt to set $P_{\text{As}} > P_{\text{As}}^\circ$, perhaps by adding As to the vapor from a high temperature Knudsen cell, then condensation of As would occur as nature acted to reduce the As partial pressure to P_{As}° . In an open system, it is not possible to precisely set $P_{\text{As}} = P_{\text{As}}^\circ$, and thus at least a small amount of sublimation, growth, or new phase formation (possible when more components are present) must

occur in practice. The extent to which this occurs depends upon experimental design and the potential effect on experimental results needs to be carefully considered.

The case of As is complicated by the existence of molecular species, i.e. monomers, dimers, and tetramers in the vapor. If these are the only species present, then the total As partial pressure, P_{As} , is a sum of these partial pressures,

$$P_{\text{As}} = P_{\text{As}_1} + P_{\text{As}_2} + P_{\text{As}_4} \quad (3)$$

At an arbitrary value of P_{As} , the equilibrium concentration of the individual species are related through reactions such as



or



Eq. (4) is described by

$$\mu_{\text{As}_1}^{\text{Std}} + kT \ln (a_{\text{As}_1}) = \frac{1}{2}\mu_{\text{As}_2}^{\text{Std}} + kT \ln (a_{\text{As}_2})^{1/2} \quad (6)$$

and Eq. (6) allows the activities of the monomer and dimer to be related to a free energy of formation of the dimer from the monomer, ΔG_f° , via

$$\frac{a_{\text{As}_2}^{1/2}}{a_{\text{As}_1}} = \exp\left(\frac{\mu_{\text{As}_1}^{\text{Std}} - \frac{1}{2}\mu_{\text{As}_2}^{\text{Std}}}{kT}\right) = \exp\left(-\frac{\Delta G_f^\circ}{kT}\right) = K(T) \quad (7)$$

A similar expression can be written relating the monomer and tetramer activities. However, for most practical temperatures, the customarily chosen standard state for the tetramer is 1 atm because of its high vapor pressure. Since this standard state is different than that for the dimer and monomer (which is less than 1 atm), we can avoid the awkwardness of using different standard states by burying the standard state pressure for each species in the equilibrium constant of Eq. (7) and similar expressions. Thus, we directly relate the actual partial pressures to each other,

$$K_8(T) = \frac{P_{\text{As}_2}^{1/2}}{P_{\text{As}_1}} \quad (8)$$

and

$$K_9(T) = \frac{P_{\text{As}_4}^{1/4}}{P_{\text{As}_1}} \quad (9)$$

Combining Eqs. (8) and (9) relates the dimer and tetramer,

$$K_{10}(T) = \frac{P_{\text{As}_4}^{1/4}}{P_{\text{As}_2}^{1/2}} \quad (10)$$

From the three independent equations, Eqs. (3), (8) and (9), the four equilibrium pressures have uniquely determined values if any one of them is known at the chosen temperature. Therefore, an equilibrium model can always be expressed in terms of $P_{\text{total}} = P_{\text{As}}$ or the partial pressure of any single As species. Actual values for the K s above lead to a prediction that $P_{\text{As}} \approx P_{\text{As}_4}$ at high P_{As} and $P_{\text{As}} \approx P_{\text{As}_2}$ at low P_{As} . It is often useful to consider this when interpreting an experiment where P_{As} is directly controlled through the addition of an arsenic-bearing chemical source.

2.1.1.2. GaAs

When applying the phase rule to the GaAs solid-vapor system, one obtains $F=2$, i.e. two thermodynamic state variables must be controlled for the system to reach a well defined and reproducible equilibrium state of ones choosing. Since the standard state is usually defined at a total pressure of 1 atm, and since many annealing systems operate at 1 atm total pressure, we will assume that an inert gas has been added to the system to keep the total pressure at 1 atm. For this solid-vapor system, $C=3$ (Ga, As, inert gas), $P=2$, and thus $F=3$. Since the composition of solid GaAs is not something that we directly control, the choices for control are T , P_{As} , P_{inert} and P_{Ga} . Of course, P_{Total} is determined by $P_{\text{Total}} = P_{\text{As}} + P_{\text{Ga}} + P_{\text{inert}}$. Since P_{inert} cannot be a result of any chemical reaction, it must be set by us and thus only two degrees of freedom remain available. The practical choices for these two degrees of freedom are T and P_{As} . The choice of these variables satisfies the necessary, but not the sufficient conditions to reach equilibrium. To be sufficient, the experimental design must make it possible for P_{Ga} to actually reach its equilibrium value, and this topic is discussed further in Section 2.1.2. Whether or not the equilibrium P_{Ga} can be reached in practice, it is important to appreciate that its value is well-defined once T and P_{As} have been selected. This is easily seen by considering the reaction between the solid, s, and vapor, v,



which is described at equilibrium by

$$\mu_{\text{GaAs},s}^{\text{Std}} = \mu_{\text{Ga},v} + \frac{1}{4}\mu_{\text{As}_{4,v}} = \mu_{\text{Ga}}^{\text{Std}} + kT \ln a_{\text{Ga}} + \frac{1}{4}\mu_{\text{As}_4}^{\text{Std}} + kT \ln a_{\text{As}_4}^{1/4} \quad (12)$$

This gives the corresponding mass action expression

$$K_{13}(T) = P_{\text{Ga}} P_{\text{As}_4}^{1/4} \quad (13)$$

or, if desired, with the aid of Eq. (10),

$$K_{14}(T) = P_{\text{Ga}} P_{\text{As}_2}^{1/2} \quad (14)$$

These show that the equilibrium value of P_{Ga} is uniquely defined once the two degrees of freedom T and either P_{As_4} or P_{As_2} are chosen. Since P_{As_4} is related to the partial pressure of the As monomers through Eq. (9), it is clear that P_{Ga} can be related to any of the individual As species. Indeed, a convenient plot of the maximum and minimum equilibrium partial pressure of the individual species is shown in Fig. 1. [9]. The lower (upper) limits to the different arsenic partial pressures are set when a third phase of liquid Ga, saturated with As (condensed As, saturated with Ga) coexists with the GaAs. Of course, the appearance of an additional condensed phase reduces the number of degrees of freedom from two to one, and thus these partial pressures and liquid compositions are all determined solely by the choice of temperature.

Additional appreciation of the necessary and sufficient conditions associated with the phase rule may be gained by considering a different choice for the 2 degrees of freedom for the GaAs solid-vapor system. The necessary condition, a choice of two independent thermodynamic variables, could be satisfied by arbitrarily choosing P_{As_4} and P_{Ga} (within the limits summarized by Fig. 1). However, the sufficient condition to bring the system to equilibrium is satisfied only when one makes it possible for all other dependent variables to attain their equilibrium value, i.e. bringing the ambient temperature to an appropriate value of T (via Eq. (13)). Of course, the above experiment is not a practical one. In practice the simplest variables to control are selected, and these are the temperature and the partial pressure of the element with a conveniently high partial pressure, i.e. T and P_{As} .

We have focused above upon solid-vapor interactions because most of the discussion to follow will make use of this viewpoint. However, it also is useful to consider the solid-liquid-vapor equilib-

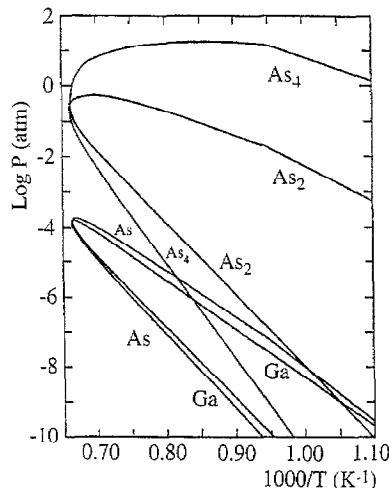


Fig. 1. Vapor pressure curves for Ga, and for the monomer, dimer, and tetramer of As, vs. $1/T$, as adapted from Ref. [9], with kind permission from Elsevier Science Ltd, The Boulevard, Langford Lane, Kidlington OX5 1GB, UK. The phase rule makes it clear that once a single partial pressure has been chosen at a selected temperature, then the values of all other equilibrium vapor pressures are constrained by the equations discussed in Section 2.1.1.

rium for GaAs. For this three phase system, $F = 1$, and the equilibrium activity of each variable is defined once the temperature is chosen. Fig. 2(a) shows a T - X slice of the phase diagram for the Ga + As system. On this scale, GaAs appears to be a 'line compound', i.e. the solidus line, S, at $x = 0.5$ is of negligible thickness and suggests that there is no variation in the stoichiometry of the solid. This diagram also shows the liquidus line labeled 'L', i.e. the line which defines all of the possible liquid compositions which may coexist at equilibrium with the solid and vapor. At equilibrium, no liquid or solid may exist with a composition between the solidus and liquidus lines. A dashed line is shown connecting the solid to the Ga-rich liquidus at a temperature, T_1 . This line illustrates that the solid may equilibrate with a Ga-rich liquid, saturated with As at a mole fraction, x_1 . It should also be clear that it is possible for the solidus to equilibrate instead with an As-rich condensed phase at temperature, T_1 .

Of course, GaAs will contain vacancies on the Ga and As sublattice, interstitials of both kinds of atoms, and antisites (atoms on the wrong sublattice), and thus the solid composition may vary slightly about $x = 0.5$. Calculations have been made to estimate the width of the solid composition region, but the accuracy of such calculations is unknown and they will not be discussed here. A schematic illustration of the anticipated range of variation in solid composition is shown in Fig. 2(b). For three

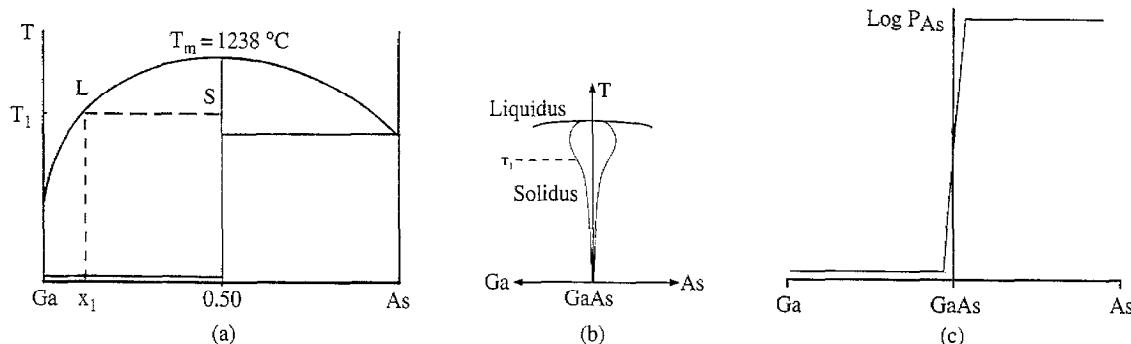


Fig. 2. (a) A T - X slice of the Ga-As phase diagram at constant total pressure. (b) An expansion of the solidus line illustrates anticipated deviations in the crystal stoichiometry between Ga-rich and As-rich ambients. (c) A P - X slice of the phase diagram at a constant temperature. All of the variation in equilibrium values of P_{As} occurs between the edges of the solidus.

phase equilibrium, $F = 1$ and a choice of temperature determines the two points on the solidus line, S, which define the solid composition in equilibrium with either a Ga-rich or As-rich condensed phase. To the extent that it is accurate, the dotted line shown at T_1 indicates that a Ga-rich liquid will equilibrate with a Ga-rich solid. Since there are the same number of lattice sites on the Ga- and As-sublattices, any variation in the crystal stoichiometry is accommodated by a mixture of vacancy, interstitial, and antisite defects. Since atomic diffusion depends upon the concentration of these native point defects, this figure makes it clear why diffusivities may vary by orders of magnitude with very minor changes in host crystal composition. In addition, as discussed in Section 3.1.1, the width of the solidus line is expected to widen in a highly doped crystal.

The maximum and minimum values of P_{As_4} are set by the three phase equilibrium, and have values which can be determined in Fig. 1. If the solid GaAs equilibrates only with the vapor, i.e. no liquid exists, then the average composition lies between the solidus lines. Without liquid present, $F = 2$ and the variables controlled in practice are T and P_{As} . In the composition region between the edges of the solidus, the partial pressures can vary substantially as the solid composition varies from Ga-rich to As-rich. This behavior is illustrated in Fig. 2(c), which shows the P -X slice of the GaAs phase diagram at a temperature, T_1 .

It is also worth considering what will happen to the solid-vapor system in a closed ampoule if we control only T , and do not select the second degree of freedom. For temperatures below approximately 640 °C, Fig. 1 shows that congruent sublimation is possible from GaAs, i.e. the same number of Ga and As atoms may leave the solid and move to the vapor. Since the dimer form of As predominates at low pressures, one expects the congruent sublimation to give an equation of constraint, $P_{\text{Ga}} = 2P_{\text{As}_2}$. Thus, ideally, the total $P = P_{\text{Ga}} + P_{\text{As}_2}$ will be determined by Fig. 1 at a chosen temperature. Of course, even the slightest chemical reaction with the system walls or with a background contaminant would significantly change these partial pressures. As discussed below, such reactions probably occur in all practical situations. In addition, $T < 640$ °C is a relatively low temperature for the study of diffusion in GaAs, and thus we shall not discuss this approach further.

At $T > 640$ °C, congruent sublimation does not occur, and the choice not to control a second degree of freedom is often an unfortunate one. At these temperatures, $P_{\text{As}_2} \gg P_{\text{Ga}}$ when equilibrium is eventually reached. Thus, for the solid to equilibrate with the vapor, much more As than Ga must sublime. The result can only be three phase equilibrium between solid, vapor, and at least a small amount of Ga-rich liquid (with a composition defined by the liquidus line). We note here that during the past decade, several groups have missed this point entirely, and we will elaborate in Section 5 on the implications of this for their reported results.

In summary, control of only one degree of freedom, T , when two are called for, does not prevent a system from reaching an equilibrium state. However, because of the very different vapor pressures of As and Ga, As will readily leave the solid at typical annealing temperatures and liquid Ga will remain on the surface. The solid must lose material for this to occur, and the resulting small droplets of Ga on the solid surface may cause roughening of both the surface and the diffusion fronts below the surface. In addition, the rate at which a second condensed phase is formed (i.e. the rate at which equilibrium is approached) depends strongly on small differences in the experimental design. This leads to a nonreproducible variation in the time required for an experiment to reach equilibrium, and is one cause of time-dependent variability sometimes reported for the measured diffusivity.

2.1.2. Open versus closed systems

As discussed above, it is necessary for one to control the correct number of *independent* thermodynamic variables to *define the equilibrium state*. The defined equilibrium state *will be attained* when the *dependent* thermodynamic variables actually reach their equilibrium values. The time

required for this to occur, or whether it even can occur, depends upon the experimental design. This section compares two commonly used experimental designs to study diffusion: closed ampoules and open tubes with flowing gases. We shall show why an open system closely approximates a closed system when a proximity cap is used. However, the ultimate closed system, a semiconductor covered with a sealed cap, is a special case of a system without a vapor. When this one degree of freedom is removed, there are such profound implications for the native defect concentrations that we have chosen to discuss it separately in Section 2.4.

2.1.2.1. Closed ampoule

We shall initially consider the annealing of pure GaAs in an ideal closed ampoule (one which does not chemically interact with Ga or As), and thus it is necessary to control 2 thermodynamic variables. The temperature will be taken as one of the controlled variables and we shall assume that it is above the congruent sublimation temperature of approximately 640 °C. We shall begin with a discussion of the effect of annealing a wafer with three commonly used ambients, i.e. in the presence of added As, added Ga, or nothing at all added to the ampoule. We shall also discuss conditions which receive rather little attention in the literature, but which appear to occasionally play a role in practice, i.e. ambients which are associated with the decomposition of the GaAs, chemical reactions with quartz, and reactions with residual oxidants.

In practice, the simplest choice for a well-controlled experiment is to add a piece of As to the ampoule prior to sealing. Because the vapor pressure of As is several atmospheres at the temperatures we are considering, it is easy to physically handle a piece of As which will sublime completely at typical annealing temperatures. A choice of T and P_{As} satisfies the necessary condition to define the equilibrium state. Solid-vapor equilibrium will be attained when P_{Ga} reaches its equilibrium value through the sublimation of GaAs, and it is simple to calculate how much sublimation is required. P_{Ga} will always be less than the vapor pressure of pure Ga at all temperatures. For example, at say $T = 900$ °C, $P_{\text{Ga}} < P_{\text{Ga}}^{\circ} \approx 4 \times 10^{-7}$ atm. Thus, if a sample with 1 cm² surface area were placed in an ampoule with a volume of 10 cm³, sublimation will total less than 5% of one monolayer.

In contrast, GaAs sitting in or near pure liquid Ga will decompose and supply As to the Ga-rich liquid until the liquid composition reaches its equilibrium value, as defined by the liquidus line in Fig. 2(a). A direct observation of this decomposition will be presented later in this section. Of course, the crystal eventually reaches an equilibrium state for three-phase equilibrium (solid–liquid–vapor). Section 5 will show that the diffusion process is strongly affected either by the sublimation process itself, or by the time dependent ambient pressures above the GaAs, and the experimental evidence will make it clear that equilibrium assumptions do not apply well to this experimental design.

If nothing but GaAs has been placed in an ampoule, then at say, $T = 900$ °C, the GaAs begins to decompose to provide both Ga and As to the vapor. The GaAs decomposition will stop when Eq. (14) holds true, i.e. $K_{14}(T) = P_{\text{Ga}}P_{\text{As}_2}^{1/2}$. Fig. 1 shows that the maximum P_{Ga} is approximately 5% of the minimum P_{As_2} required for the solid to equilibrate with the vapor. This means that the large majority of Ga atoms produced by the decomposing GaAs must form a liquid phase. The amount of liquid which appears will depend primarily upon temperature and ampoule volume. For many of the closed ampoule experiments whose results have been published, it will be shown below that the quantity of Ga expected to appear is much less than 1 μmol. Thus, samples annealed under these conditions are expected to reach three-phase equilibrium much more quickly than samples annealed with intentionally added Ga metal. This means that diffusivities measured in these experiments are much more likely to be representative of native defect equilibrium associated with Ga-rich GaAs than diffusion measured in the presence of intentionally added Ga. Experimental evidence will be shown in Section 5 which indicates that several groups have seriously misinterpreted Ga-rich ampoule data.

2.1.2.1.1. Decomposition of GaAs in liquid Ga

Olmsted et al. [10] added approximately 3 mg of Ga to several ampoules and annealed the liquid plus solid GaAs at $T = 855^\circ\text{C}$. At this temperature, it has been determined [11] that the composition of the Ga-rich liquid has an As mole fraction, $x_{\text{As}} \approx 0.04$, on the liquidus line. Thus, their Ga-rich liquid should have contained approximately 0.12 mg As when it equilibrated with the GaAs. Since Olmsted et al. reported that their ampoules had a nominal volume of 1 cm^3 , it is reasonable to assume that the surface area of their samples did not exceed 1 cm^2 . If we assume that the total surface area of a given sample was exactly 1 cm^2 , then simple calculations show that approximately $0.6\text{ }\mu\text{m}$ of GaAs should have disappeared from all surfaces, assuming that the rate of mass transport was the same from all surfaces. Fig. 3 shows the Al marker concentration profiles taken by secondary ion mass spectroscopy (SIMS) reported by Olmsted et al. [10] for (a) an as-grown superlattice (SL), and pieces annealed either (b) under As overpressure, or (c) in the presence of liquid Ga. It is clear that approximately 500 \AA of the sample has disappeared after annealing in the presence of Ga. However, this loss of GaAs from the epitaxial surface is roughly an order of magnitude lower than the calculated estimate.

The unexpectedly small amount of decomposition discussed above might be related to the placement of the sample, i.e. perhaps most of the decomposition took place on the back surface because the face was protected against the quartz. However, the results of Hsieh et al. [12] suggest that something else may be happening to protect the surface. Hsieh et al. also studied interdiffusion in GaAs in ampoules containing either (a) 20 mg As, (b) nothing else, or (c) 20 mg Ga. These studies were conducted at temperatures of up to $T = 925^\circ\text{C}$, where one expects a mole fraction of As, $x_{\text{As}} \approx 0.07$, for a Ga-rich liquid at equilibrium with GaAs [11]. A mirror-like surface was reported for cases (a)

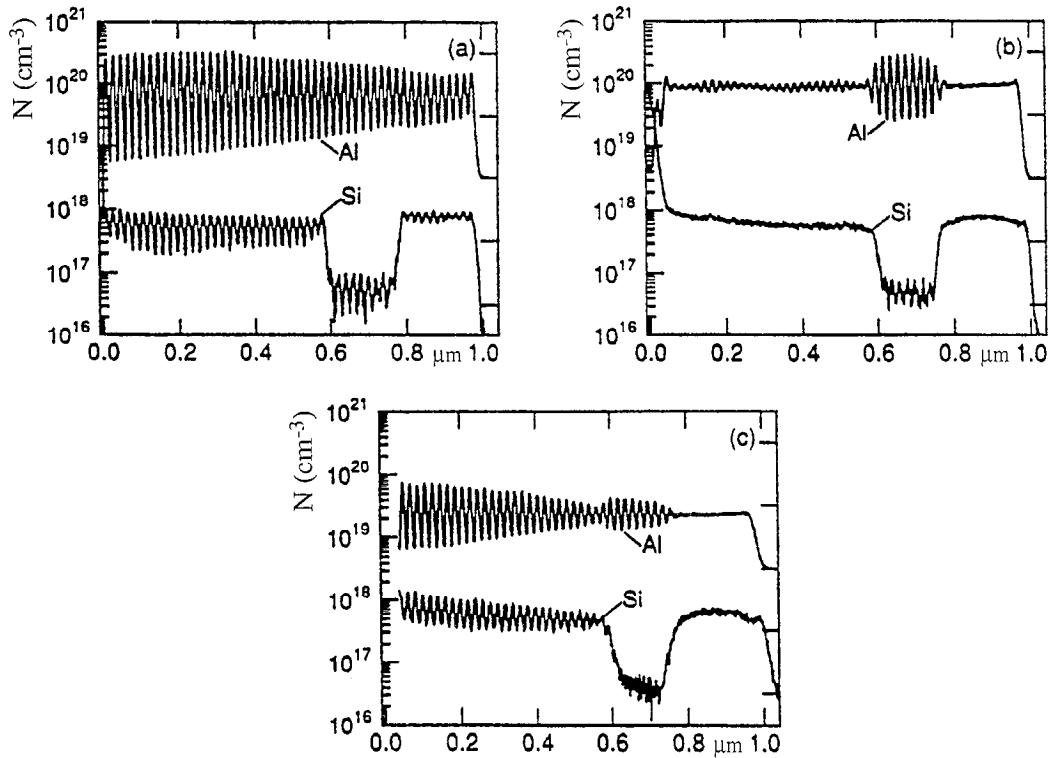


Fig. 3. SIMS concentration profiles of Al and Si reported by Olmsted et al. for (a) the as-grown SL, (b) after annealing the SL under $P_{\text{As}} \approx 1\text{ atm}$ for $t = 0.5\text{ h}$ at $T = 855^\circ\text{C}$, and (c) after annealing the SL in the presence of Ga liquid at $T = 855^\circ\text{C}$ for $t = 24\text{ h}$. Approximately 500 \AA of the sample in (c) appears to have decomposed. Reprinted with permission, from Ref. [10], copyright 1993, American Institute of Physics.

and (c) at all temperatures, but significant damage to the surface was reported for case (b) whenever $T \geq 700$ °C. Given their reported ampoule volume of ≤ 4 cm³, it is unlikely that their sample areas were much in excess of 1 cm². Assuming that their sample surface area was 1 cm², then approximately 6 μm should have been removed from the GaAs surface through the decomposition process to bring the original 20 mg of Ga metal to its equilibrium composition of 7% As. However, the quantum wells used to study the diffusion process were grown only 0.15 μm below the original surface, and yet remained intact for their photoluminescence (PL) studies. No particular effort to protect the surface from sublimation was mentioned, and the lack of etching of group (c) is almost certainly a result of a protective coating which formed on the GaAs. It is worth considering the possible contaminants present in a quartz closed ampoule which might lead to the formation of a coating.

2.1.2.1.2. Chemical interactions with quartz

In practice, chemically reactive elements are present to some degree during annealing, but one usually makes the optimistic assumption that they do not affect the diffusion results. At a minimum, Si and O from the quartz must be considered, assuming that spectroscopic quality quartz (which contains low concentrations of water and certain metals) was used as the starting material, and efforts were made to avoid hydrating the glass with the H₂O produced by the flame of a torch. In recent years, these points have been seldom mentioned in reports of diffusion in GaAs, and we consider it likely that H₂O (at a minimum) has been present in the ampoule. However, we shall initially discuss the most optimistic case, i.e. a dry quartz ampoule.

Without H₂O present, Ga is known [13] to reduce quartz to produce Si and the volatile gallium suboxide, Ga₂O, via



and also the less volatile SiO, via



Indeed, these reactions were studied in the early days of the technology of liquid phase epitaxy (LPE) when quartz boats and containers were used at temperatures close to the melting temperature of GaAs. Cochran et al. [13] showed that these reactions could not be neglected even at $P_{\text{As}} = 1$ atm. In addition, they demonstrated that rapid etching of quartz and gradual growth of Si crystals on the etched walls occurred, particularly in the presence of liquid Ga. They also noted a relation between the etching and the observed Si contamination in the GaAs. In diffusion ampoules, we expect that the suboxides will reach equilibrium with the other species in the ampoule. However, apart from any unintended doping with Si and perhaps a modest increase in the GaAs sublimation (to provide the extra Ga for the suboxides), there is probably little net effect on the equilibrium state of the crystal because the suboxides are volatile and they are believed not to condense on the GaAs surface.

2.1.2.1.3. Chemical reactions with residual oxidants

Thurmond et al. [14] have determined the Ga–As–O ternary phase diagram. They reported that at $T > 600$ °C, a hard crystalline phase of gallia with negligible vapor pressure forms on the GaAs surface, and this phase is known as $\beta\text{-Ga}_2\text{O}_3$. Monteiro et al. [15] determined that at similar temperatures, most of the As leaves the oxide, although precipitates of both As and its oxides may remain in the gallia for extended times. Hughes and Ludeke [16] concluded that As₂O₃ reacted with GaAs to form gallia and free As. Based upon published free energies of formation, we have confirmed that these results are consistent with the known thermochemistry for these phases. The most important point to note is that exposure of GaAs to an oxidant at elevated temperatures will cause the growth of

an oxide which is dense, hard, and chemically unreactive, and which is largely composed of Ga_2O_3 . Such an oxide would effectively halt the semiconductor decomposition. It would presumably cut off the exchange of Ga and As atoms between the solid and vapor, which implies that the point defects would be unable to reach their equilibrium concentrations defined by the ambient, and diffusion results would be difficult to reproduce.

Potential sources which may supply oxidants to the ambient vapor inside of a quartz ampoule include (1) H_2O left in the quartz after either manufacture or exposure to an oxyhydrogen (or similar) flame used to seal the ampoule, (2) O_2 or H_2O adsorbed on, and absorbed in, any solid As source added to the ampoule, (3) oxides of Ga formed by reaction with O_2 or H_2O in the air and held as a scum in any intentionally added Ga source, and (4) the native oxides of Ga and As which appear on the GaAs surface after exposure to air and moisture. A few reports in recent years have addressed item (4), i.e. the removal of native oxide from their samples prior to annealing. However, none of them have addressed the first three.

At typical anneal temperatures, the solubility limit in SiO_2 for O_2 is $\sim 10^{16} \text{ cm}^{-3}$ and for H_2O is $\sim 10^{19} \text{ cm}^{-3}$, and thus H_2O appears to be of more serious concern. To obtain reproducible results, Palfrey et al. [17] considered it essential to remove residual water by baking the quartz tubing for several hours in flowing N_2 at the diffusion temperature prior to annealing. It is worth considering a simple example. For a hypothetical spherical ampoule with inside radius of 1 cm, this ampoule would have an internal volume of $\approx 4 \text{ cm}^3$. If the wall were 3 mm thick, then the volume of SiO_2 would be $\approx 3.5 \text{ cm}^3$. Of course, a real ampoule will not be spherical, and thus the volume of glass would be larger than this for the same inner volume. If all of the quartz contained H_2O at its solubility limit, and if only 1% of that H_2O entered the ampoule, then the steam partial pressure inside the ampoule would be $\sim 10^{-2} \text{ atm}$, and a hard coating of Ga_2O_3 , presumably mixed with As_2O_3 would be expected to form on the GaAs surface.

Cochran et al. observed that oxidants entering an LPE system with 'high purity' Ga metal required significant effort to be removed. Pumping on the Ga, or flowing a dry inert gas through it, at $T > 600^\circ\text{C}$ appeared to be necessary. To date, we have not found any reports of diffusion which discuss cleansing of the Ga source prior to adding it to an ampoule. Presumably O_2 and H_2O may enter a system as a contaminant in metallic As as well. We are unaware of any reports which have measured the level of these impurities after exposure to air, or have considered any associated effect on diffusion.

If the native defect concentrations cannot be defined with respect to an external phase, then diffusion cannot be related to an equilibrium model. Indeed, it will be shown in Section 2.4 that the native defect concentrations will depend upon the history of the substrate and epitaxial layer, and will generally drift with time. Presumably, this problem becomes even more severe when Al is used as a diffusion marker because of the very strong bond between Al and O. For example, moisture in a room temperature environment causes [18] AlAs layers to become hydrolyzed and largely oxidized over a time scale of days. Measurements [19,20], made at temperatures of $400\text{--}500^\circ\text{C}$ have shown that H_2O can diffuse tens of microns laterally through submicron thick AlAs layers as the layers convert to an oxide of Al in a few minutes. In these anneals, the oxidation of the adjacent GaAs is much slower, and often assumed not to occur. The extent to which trace quantities of oxygen affect the interdiffusion of Al in GaAs remains to be determined, but some of the experimental noise in the interdiffusion coefficients summarized in Section 5 is probably related to the combined presence of oxygen and aluminum in GaAs.

2.1.2.2. Open system

An open tube system used for annealing often consists of a quartz tube through which gas can flow over and around a sample and its proximity cap contained inside of a small oven made of graphite (or other nonreactive material), such as that shown in Fig. 4(a) [21]. Such ovens are often heated by

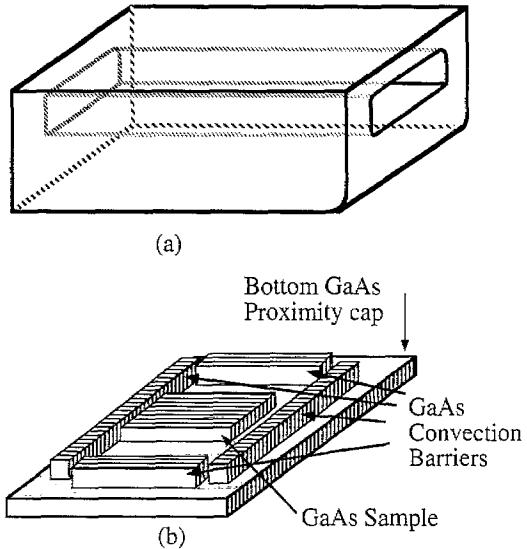


Fig. 4. (a) Sketch of a rectangular graphite oven used for open tube annealing. This oven was used in place of the growth susceptor in a growth reactor, and its snug fit inside of the rectangular quartz allowed relatively low gas flows to be used. A thermocouple placed in a hole (not shown) monitors the graphite temperature, and controls the power to an IR lamp below the oven. At higher anneal temperatures, a gold coated mirror was placed above the oven to reduce the radiation losses. (b) At higher anneal temperatures, GaAs samples are often placed between two GaAs 'caps' and surrounded by loose fitting GaAs bars acting as convection barriers. The convection barriers stop the cold flowing gas from cooling the sample, and reduce Ga sublimation losses.

the absorption of IR energy, and their temperature is electronically controlled. Since diffusion studies are almost always performed at temperatures where energy transfer occurs largely by radiant energy, it is important that a sample be exposed to radiant energy from as wide a solid angle as possible, i.e. not simply placed on top of a growth susceptor. When the samples and their proximity caps are placed in the slot shown in the oven of Fig. 4(a), they are bathed in a reasonably uniform thermal ambient. Reichert et al. [21] routinely placed their GaAs samples face down on a polished sapphire substrate or on a GaAs cap and found that no damage to the GaAs surface occurred at $T \leq 700$ °C for anneal times of 1–2 h. At $T = 800$ °C, Chen [22] placed both the GaAs sample and proximity cap on a sapphire substrate in the oven. After annealing, they found that mass transport occurred from the bottom piece of GaAs to the top piece. Since this implied that a thermal gradient existed in the oven, GaAs convection barriers were added to reduce the flow of cold gas over the sample(s), as shown in Fig. 4(b). With a second proximity cap (not shown in Fig. 4(b)) placed over the samples and convection barriers, all measurable mass transport between the sample and its proximity cap was eliminated, and pristine surfaces were obtained at all temperatures studied, including $T = 900$ °C.

An open system is appealing because of its inherent simplicity and cleanliness. At first glance, performing an anneal for minutes or hours under flowing gas appears to be risky because the GaAs might simply collect the impurities entering with the carrier gas and As source. However, in modern chemical sources, the concentrations of metallic contaminants in the source gases are so low that they can be neglected because they do not change the carrier concentration. Of course, in any system, there are invariably unwanted low concentrations of O₂ and H₂O present. If GaAs were to be oxidized by the uncontrolled microscopic quantities of oxidant present in a flowing carrier gas, then Ga₂O₃ would cap the surface and either slow or stop solid–vapor equilibration. However, these oxides are thermodynamically unstable in the presence of H₂, which is the carrier gas used in most open systems. We note that in the early days of the Si technology, solid Ga₂O₃ was reduced by H₂ to provide an in situ source, Ga₂O, which was transported through the vapor and decomposed over hot Si to provide Ga for

use as a p-type dopant [23]. More recently, Kisker et al. [24] have shown that the native oxide remains intact on the GaAs surface at $T > 500^\circ\text{C}$ when only N_2 is present, but that the oxide will disappear at the same temperature if H_2 is present. Thus, we conclude that when H_2 is the carrier gas in a reasonably clean system, the kinetics of oxide decomposition is sufficiently rapid that it is impossible for a layer of oxide to build up on a GaAs surface. We emphasize that this is quite different than the situation inside of a quartz ampoule where the accidental presence of any residual oxidant is likely to cause the formation of Ga_2O_3 on a GaAs surface.

As discussed above for the GaAs solid-vapor system, the phase rule implies that a choice of T and P_{As} satisfies the necessary conditions for equilibrium to be defined. These two variables determine a unique equilibrium value for P_{Ga} via Eq. (13), and one must make it possible for the actual P_{Ga} above the sample surface to attain this equilibrium value. The only practical approach to accomplish this is to allow the natural sublimation of the GaAs wafer to supply the Ga needed to reach equilibrium. This is usually accomplished by using a GaAs proximity cap, i.e. by placing the sample face-to-face with another piece of GaAs. The primary purpose of the proximity cap is to slow down the rate of sublimation of the sample by reducing the loss of Ga to the flowing gas ambient in the open tube (which contains no added Ga). The net result is beneficial in two ways: the solid and the vapor next to the solid quickly approach a predictable equilibrium, and the low sublimation rate causes the original good morphology of the sample surface to be preserved after long anneal times.

It is worth making an order of magnitude estimate of the sublimation rate to appreciate how effective a proximity cap actually is. To keep the calculations simple, we shall consider a pair of round 1 cm radius GaAs wafers placed in the intimate proximity of each other, and assume that they are separated by a distance of 1 μm , as indicated by the cross-sectional cut shown through the wafers sketched in Fig. 5(a). The small diameter chosen will lead to an overestimate of the sublimation loss, but the spacing between real wafers is more difficult to estimate since it depends upon imperfections such as warpage, and we simply take a uniform value of 1 μm to be a reasonable average estimate.

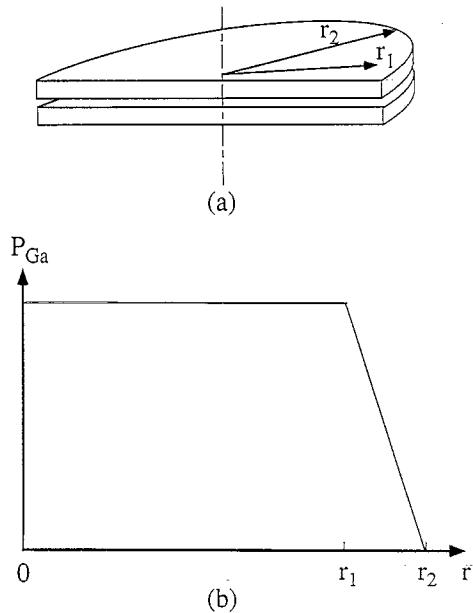


Fig. 5. (a) For the calculation of sublimation losses, a GaAs sample wafer is assumed to be in the intimate proximity of another wafer acting as its 'proximity cap', as illustrated in this schematic cross section. (b) The assumed profile of P_{Ga} vs. radial position, in the space between the wafer and its proximity cap, used to make an order of magnitude estimate of the diffusion limited loss of Ga to the flowing gas ambient of the reactor. This assumption has the effect of underestimating sublimation near the edge of the wafer, but overestimating sublimation near the center.

This spacing between wafers is roughly $100 \times$ the mean intermolecular collision distance at atmospheric pressure, and thus it is reasonable to use a gas phase diffusivity of [25] $D_{\text{As}} \approx D_{\text{Ga}} \approx 10 \text{ cm}^2 \text{ s}^{-1}$ near $T = 900 \text{ }^\circ\text{C}$. If $P_{\text{As,ext}}$ is set in the open tube ambient, then the characteristic diffusion length, $L = (Dt)^{1/2}$, indicates that the As will diffuse into the space between the wafers and bring P_{As} over the sample close to $P_{\text{As,ext}}$ in a time which is roughly on the order of 1 s. A comparable amount of time will be required for the equilibrium P_{Ga} profile to be established.

We note that the equations describing the steady state Ga sublimation rate and gas phase diffusional flux are coupled and position dependent, and that they do not result in linear equations. However, we will linearize the problem in order to make an order of magnitude calculation of the sublimation rate in the central portion of the wafer, i.e. in the region where any electronic devices would be. We will assume that the P_{Ga} profile is as shown in Fig. 5(b), which sets P_{Ga} at its equilibrium value from the center of the wafer out to a radius, r_1 . Between radii of r_1 and r_2 , it will be assumed that no sublimation occurs. This is equivalent to assuming a constant Ga flux, J , flows through the gas in the annulus, $J = D(\partial C/\partial r)$, and that P_{Ga} falls linearly to zero (its value in the open tube ambient). Using this equation, we shall make an effort to *overestimate* the loss of Ga from the central region of the wafers by taking $\delta r = r_2 - r_1 = 10^{-3} \text{ cm}$. We shall assume that $T = 900 \text{ }^\circ\text{C}$ and $P_{\text{As}} = 1 \times 10^{-3} \text{ atm}$ are the two degrees of freedom chosen, and that an inert carrier gas is present simply to keep the total pressure at 1 atm. From Fig. 1, the equilibrium $P_{\text{Ga}} \approx 4 \times 10^{-8} \text{ atm}$. The diffusion limited flux, J (atoms s^{-1}), of Ga into the open tube can then be estimated as

$$J = \frac{D_{\text{Ga}}}{RT} \frac{P_{\text{Ga}}}{\delta r} N_A A \approx 1.5 \times 10^{12} \text{ Ga atoms s}^{-1} \quad (17)$$

where $R = 82 \text{ atm cm}^3 \text{ K}^{-1} \text{ mol}^{-1}$ is the gas constant, N_A is Avogadro's number, and A is the cross-sectional area across which Ga transports, i.e. the wafer perimeter times the space between the wafers. A wafer of the assumed size has approximately 10^{15} Ga atoms in the top monolayer (ML). Thus, for the assumptions given, one expects that 1 ML will sublime from each face-to-face wafer in about 30 min. The calculation is seriously in error near the edge of a wafer, but experience has shown that this order of magnitude estimate is valid in the central region of samples of modest size.

2.2. Equilibrium concentration of native point defects

2.2.1. Entropy and the existence of native point defects

Kroger [26] has given an elegant, yet simple, argument for the existence of a well defined equilibrium concentration of vacancies in an elemental crystal. Considering an elemental crystal whose substitutional sites contain either atoms of type 'a' or vacancies 'v', then the free energy, G , of the crystal depends upon (1) the total number of atoms, N_a , and the total number of vacancies, N_v , (2) the chemical potential of each in their standard states, μ_a^{Std} and μ_v^{Std} , respectively, (3) the temperature, and (4) the configurational entropy (entropy of mixing), i.e.

$$G = N_a \mu_a^{\text{Std}} + N_v \mu_v^{\text{Std}} - TS_{\text{Config}} \quad (18)$$

Since the entropy of mixing for an ideal solution is related to the activities of the atoms and vacancies via

$$S_{\text{Config}} \approx k(N_a \ln[a_a] + N_v \ln[V]) \quad (19)$$

one may describe the chemical potentials,

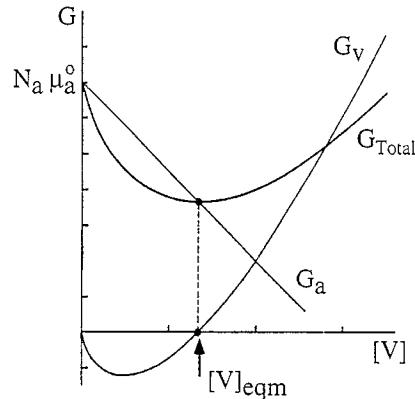


Fig. 6. Schematic illustration of free energy, G , for the atoms, vacancies, and crystal vs. the vacancy concentration. Because of entropy, the lowest free energy of the system occurs for a nonzero vacancy concentration at temperature, $T > 0$. Adapted from Ref. [26].

$$\mu_v \equiv \left. \frac{\partial G}{\partial N_v} \right|_{T, P, N_a} = \mu_v^{\text{Std}} + kT \ln [V] \quad (20)$$

$$\mu_a \equiv \left. \frac{\partial G}{\partial N_a} \right|_{T, P, N_v} = \mu_a^{\text{Std}} + kT \ln [a_a] \quad (21)$$

For low vacancy concentrations, $[a_a] \approx 1 - [V]$, and the free energy of the crystal may be identified with the separate contributions,

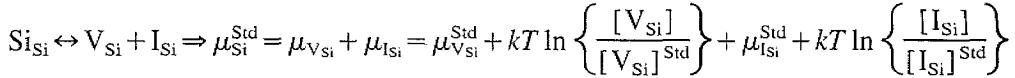
$$G = G_a + G_v = N_a \mu_a + N_v \mu_v \quad (22)$$

Thus, as a small concentration of vacancies increases, G_a will decrease slowly, as shown in Fig. 6. However, $G_v = 0$ at $N_v = 0$, and the entropic contribution causes G_v to drop rapidly with the addition of a small vacancy concentration. This, of course, is exactly the same behavior which describes an impurity in any solid and is the physical reason why the solubility of any impurity never attains a value of zero at $T > 0$. The minimum in the total free energy of the crystal, G , defines the equilibrium state of the crystal, and as can be seen in Fig. 6, this equilibrium state requires a nonzero concentration of vacancies. Using analogous arguments, one can show that other native point defects, such as interstitials, have nonzero equilibrium concentrations in elemental or compound semiconductors. Indeed, it is *only* as $T \rightarrow 0$ that the entropic contribution to the free energy disappears and that the equilibrium concentration of point defects $N \rightarrow 0$. In practice, of course, point defects are grown into a crystal at elevated temperatures, and are frozen into place (on human time scales) at low temperatures, and thus are present in all crystals at all applied temperatures.

2.2.2. Neutral native point defects in Si

The remainder of Section 2 and all of Section 3 will introduce several important concepts associated with point defects in semiconductors. To keep the discussion as simple as possible, we shall focus on neutral point defects in Section 2, and discuss charged point defects in Section 3.

For an elemental semiconductor such as Si, in the presence of only its own vapor, $F = 1$. Setting T is sufficient to determine one unique equilibrium value for each of the other thermodynamic quantities, e.g. the vapor pressure, P_{Si}^0 , the concentration of Si vacancies, and the concentration of Si interstitials. Considering the so-called Frenkel reaction which relates vacancy-interstitial (V-I) generation or annihilation, to substitutional Si, Si_{Si} ,

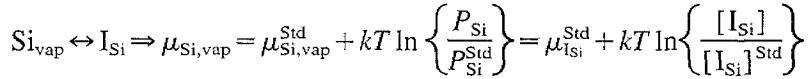


provides the useful mass action expression,

$$K_{23}(T) = [\text{V}_{\text{Si}}][\text{I}_{\text{Si}}] \quad (23)$$

which relates the product of the equilibrium vacancy and interstitial concentrations to an equilibrium constant, K_{23} , which is determined only by the chosen temperature. For an elemental semiconductor, it is common to choose the equilibrium state to be the standard state, and thus if solid–vapor equilibrium exists, then the concentrations $[\text{I}_{\text{Si}}]$ and $[\text{V}_{\text{Si}}]$ have the values $[\text{I}_{\text{Si}}]^{\text{Std}}$ and $[\text{V}_{\text{Si}}]^{\text{Std}}$, respectively. We note that Eq. (23) is sometimes misinterpreted as meaning that the crystal is in equilibrium whenever the product of $[\text{V}_{\text{Si}}]$ and $[\text{I}_{\text{Si}}]$ is equal to K_{23} . It is important to understand that each defect concentration has *one, and only one*, equilibrium value. Of course, there are times when it may be reasonable to use Eq. (23) to estimate one defect concentration in terms of the other. However, whether this is appropriate depends upon the atomistic model assumed, and details of several models are introduced in Section 2.5 and later sections.

Although K_{23} determines the product of the vacancy and interstitial concentrations, a single equation cannot define two separate concentrations. A second independent reaction is required. One possible reaction to consider is that between the vapor and interstitials,



Once the temperature has been chosen, solid–vapor equilibrium is defined, and the Si partial pressure is simply its vapor pressure, $P_{\text{Si}} = P_{\text{Si}}^{\circ}(T) \equiv P^{\text{Std}}$. This defines the standard state for the interstitials, $[\text{I}_{\text{Si}}] = [\text{I}_{\text{Si}}]^{\text{Std}}$, and

$$K_{24}(T) = [\text{I}_{\text{Si}}]^{\text{Std}} \quad (24)$$

uniquely defines $[\text{I}_{\text{Si}}]^{\text{Std}}$ once the temperature is chosen. Taking these two equations together, it is clear that the concentration of both interstitials and vacancies are uniquely defined by the temperature.

It is also worth reflecting on the implications of this logic because it can be tempting to believe that the equilibration process actually works as described above. However, one could describe a fundamentally different reaction,



which could have been taken together with Eq. (23) (or with Eq. (24)) and viewed as determining the concentration of the two defects. We also note that these latter two expressions are equivalent to the so-called Schottky reactions which relate the vacancy or interstitial concentrations to surface reactions. These results show both the beauty and the problem of thermodynamics: (1) the mass action expressions obtained from *any chemical reaction* relating a species in two phases will correctly describe the relative equilibrium concentrations of that species in the two phases, and (2) nothing at all can be learned about the kinetics, i.e. the reaction pathway *actually taken* by the point defects to approach equilibrium.

2.2.3. Neutral native point defects in GaAs

For the GaAs system composed of only solid and vapor phases, $F = 2$, and the equilibrium concentration of all point defects must be defined upon setting two degrees of freedom. In practice, one usually sets T and P_{As} . Because of the existence of two sublattices, the native defects are more

complex than those found in Si. There are six types of native point defects: vacancies on a Ga or As site, V_{Ga} or V_{As} , interstitial atoms of Ga or As, I_{Ga} or I_{As} , and atoms substituting on the wrong sublattice, Ga or As antisites, Ga_{As} or As_{Ga} , respectively. Assuming only uncharged point defects in this section, six reactions are necessary to define the equilibrium concentration of the six point defects with respect to the two degrees of freedom, T and P_{As} . We arbitrarily choose to use the following six equations to relate the equilibrium concentrations:

$$\frac{1}{4}As_{4,vap} \leftrightarrow I_{As} \Rightarrow K_{26}(T) = \frac{[I_{As}]}{P_{As_4}^{1/4}} \quad (26)$$

$$Ga_{vap} \leftrightarrow I_{Ga} \Rightarrow K_{27}(T) = \frac{[I_{Ga}]}{P_{Ga}} \quad (27)$$

$$As_{As} \leftrightarrow I_{As} + V_{As} \Rightarrow K_{28}(T) = [I_{As}][V_{As}] \quad (28)$$

$$Ga_{Ga} \leftrightarrow I_{Ga} + V_{Ga} \Rightarrow K_{29}(T) = [I_{Ga}][V_{Ga}] \quad (29)$$

$$Ga_{Ga} + I_{As} \leftrightarrow As_{Ga} + I_{Ga} \Rightarrow K_{30}(T) = \frac{[As_{Ga}][I_{Ga}]}{[I_{As}]} \quad (30)$$

and

$$As_{As} + I_{Ga} \leftrightarrow Ga_{As} + I_{As} \Rightarrow K_{31}(T) = \frac{[Ga_{As}][I_{As}]}{[I_{Ga}]} \quad (31)$$

We emphasize that for the chosen two degrees of freedom, T and P_{As} , there is one, and only one, equilibrium concentration for each defect. Using Eq. (13), Eqs. (26)–(31) may be rewritten to demonstrate explicitly that the concentration of each native defect is only a function of the chosen two degrees of freedom, and this is summarized in Table 1, which shows the dependence upon P_{As} at a given temperature. The temperature dependence of the native defect concentrations occurs through the equilibrium constants, and these are not shown in Table 1. The above relationships may also be useful in understanding how a perturbation in one point defect concentration can temporarily affect the concentration of other defects. Perturbations upon equilibrium states will be discussed in Sections 4 and 5. If desired, mass action expressions describing the concentration of defect complexes, i.e. divacancies, etc., may be built up from the six basic reactions above. However, there appears to be little evidence that such species are needed to explain diffusion data, and we shall not discuss them further.

Table 1
Predicted dependence of equilibrium native point defect concentrations upon P_{As_4} in intrinsic GaAs

| Equations | Prediction |
|--------------------------------|-------------------------------------|
| (26) \Rightarrow | $[I_{As}] \propto P_{As_4}^{1/4}$ |
| (13), Eq. (27) \Rightarrow | $[I_{Ga}] \propto P_{As_4}^{-1/4}$ |
| (26), (28) \Rightarrow | $[V_{As}] \propto P_{As_4}^{-1/4}$ |
| (27), (29) \Rightarrow | $[V_{Ga}] \propto P_{As_4}^{1/4}$ |
| (26), (27), (30) \Rightarrow | $[As_{Ga}] \propto P_{As_4}^{1/2}$ |
| (26), (27), (31) \Rightarrow | $[Ga_{As}] \propto P_{As_4}^{-1/2}$ |

2.2.4. Charged native point defects in intrinsic GaAs

Point defects are believed to exist in a mixture of charge states, even in an intrinsic semiconductor. When solid–vapor equilibrium is reached, the concentrations of these point defects are also determined by the choice of two degrees of freedom. The concentration of neutral defects and their charged analogues can be related via the electron concentration, n , via reactions such as,



where superscripts are used to indicate the charge state of the point defects. For the intrinsic crystal, $n = n_i(T)$, and thus two degrees of freedom also define these concentrations, i.e.

$$\text{Eq. (32)} \Rightarrow [V_{\text{Ga}}^{-1}] \propto [V_{\text{Ga}}^0] = f(T, P_{\text{As}}) \quad (33)$$

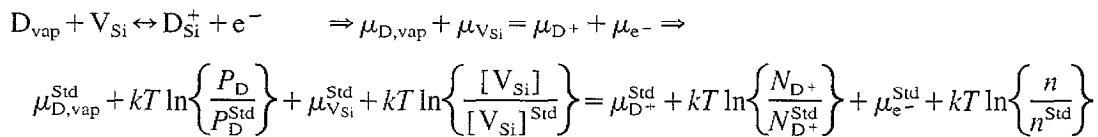
For doped crystals, Section 3.2.2 will discuss charged native defects in detail and show that their concentration depends upon three degrees of freedom, i.e. $f(T, P_{\text{As}}, P_{\text{dop}})$, or (sometimes) equivalently $f(T, P_{\text{As}}, n)$.

2.3. Equilibrium dopant solubility in an intrinsic crystal

Applying equilibrium expressions to native defects, and in turn to the diffusion of impurities, is likely to be fruitless unless one has a clear understanding of the extent to which the impurity atom concentration has equilibrated with the vapor. As used in this paper, the term ‘solubility’ refers to the dependence of the *equilibrium* impurity concentration upon the chosen degrees of freedom, i.e. including the ambient partial pressures. It is useful to distinguish the term ‘solubility’ from the term ‘solubility limit’, which refers to the maximum equilibrium impurity concentration obtained in the semiconductor. Since the crystal is usually extrinsic when the solubility limit is reached, we shall discuss the solubility limit in Section 3.2.3. In this section, we shall consider experiments performed at relatively high temperature and low dopant partial pressure, i.e. when $n = p = n_i$. As shown below, solubility can sometimes be described by a so-called distribution coefficient, but often it cannot be.

2.3.1. Substitutional impurity in Si

As discussed above, there is one degree of freedom for a Si wafer–vapor system. Setting the temperature thus determines the equilibrium values of P_{Si} and all charged and uncharged native defect (vacancy and Si interstitial) concentrations. However, an additional degree of freedom appears with the addition of an impurity to the system. Control of both T and the dopant partial pressure, P_{dop} , are sufficient to define the equilibrium concentration of the dopant in the solid. Since any reaction pathway may be used to relate the dopant partial pressure to the dopant concentration at equilibrium, we choose to use the neutral vacancy to relate a generic donor in the vapor, D_{vap} , to ionized donors in substitutional sites, D_{Si}^+ ,



which gives the mass action expression,

$$K_{34}(T) = \frac{N_{D^+}^{\text{Std}} n}{P_D [V_{\text{Si}}]} \quad (34)$$

Eq. (34) relates the partial pressure of a donor, P_D , directly to the ionized donor concentration in the wafer, N_D^+ . In the intrinsic crystal, $n = n^{std} = n_i$, and $N_D^+ + N_D$, and thus, $N_D \propto P_D$. For a system which uses a carrier gas to maintain a constant total pressure, the solubility can then be described by a single number, i.e. a distribution coefficient, k , which relates the mole fraction of dopant in the solid, X_s , to the mole fraction of dopant in the vapor, X_v ,

$$k = \frac{X_s}{X_v} = \frac{N_D/N_{Si}}{P_D/P_{Tot}} = \text{const.} \quad (35)$$

where N_{Si} = atomic density of Si, and P_{Tot} is the total pressure in the system.

2.3.2. Substitutional impurity in GaAs

There are three degrees of freedom for a system containing a GaAs wafer, vapor, and dopant. Thus, the solubility of the dopant must depend upon the three independent variables chosen.

Since the results of a thermodynamic prediction must be independent of the actual reaction pathway, we arbitrarily choose the neutral Ga vacancy to mediate between the impurity in the vapor and solid phases. For a generic donor atom, D, assumed to reside only on the Ga sublattice,



or

$$K_{36}(T) = \frac{N_D^+ n}{P_D [V_{Ga}]} \quad (36)$$

This result is superficially similar to the case in Si. However, the equilibrium vacancy concentration in Si is set by the temperature alone, and $N_D^+ = f(T, P_D)$. In GaAs, Table 1 shows that $[V_{Ga}] \propto P_{As_4}^{1/4}$, and thus Eq. (36) predicts that the solubility, $N_D^+ = f(T, P_{As}, P_D)$, is set by the three degrees of freedom chosen. Equations analogous to Eq. (36) may be written for acceptors or even neutral impurities residing on the group III sublattice. Since this section is concerned only with the intrinsic crystal, i.e. $n = p = n_i$, a pair of expressions can be shown to describe the solubility of any impurity in terms of T , P_{As} , and P_{dop} ,

$$N_{dop} \propto P_{dop} P_{As_4}^{1/4} \quad \text{for impurity on group III sublattice} \quad (37)$$

and

$$N_{dop} \propto P_{dop} P_{As_4}^{1/4} \quad \text{for impurity on group V sublattice} \quad (38)$$

Eqs. (37) and (38) show that a distribution coefficient is, strictly speaking, not applicable to GaAs even at the lowest impurity concentrations. However, this is often ignored because of the weak dependence upon P_{As} and because only one value, or a small range of values, of P_{As} are used in practice.

2.3.3. Interstitial impurity in GaAs

It should be noted that interstitial impurity atoms also have a well defined solubility if solid-vapor equilibrium applies. For a generic impurity atom, d, equilibrium between the neutral impurity interstitials, I_d , and atoms in the vapor, d_{vap} , can be described via



or

$$K_{39} = \frac{[I_d]}{P_d} \quad (39)$$

and thus for neutral interstitials,

$$[I_d] \propto P_d \quad (40)$$

The neutral interstitial impurity has no explicit dependence upon P_{As} . However, charged interstitial impurity concentrations will be affected by P_{As} , as shown in Section 3.2.1. Evidence that some important interstitial atoms are charged in GaAs will be presented in Sections 4 and 5.

2.4. Point defect concentrations under a sealed cap

It is often considered desirable to encapsulate a semiconductor to prevent the As from subliming during annealing and to prevent the formation of a Ga-rich roughened surface. We shall briefly consider the thermodynamic description of both a Si wafer and a GaAs wafer sealed by an ideal nonreactive capping material. It will be shown that without an external reservoir of atoms in contact with the solid, the native defect concentrations are significantly different than those predicted by the previous equilibrium expressions. The large differences in practice will be shown in Section 5.3 which summarizes experimental results for capped annealing.

The fundamental question underlying native defect equilibration, when annealing under an ideal sealed cap, can perhaps be best summarized as

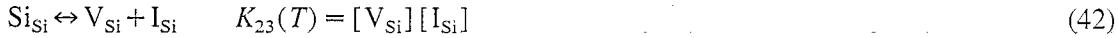
$$\mu_{\text{Defect}}(\text{Semiconductor}) = \mu_{\text{What?}}(\text{Where?}) \quad (41)$$

because there are no atoms in another phase acting as a reservoir to define the chemical potential of the defects. We shall define an ‘ideal’ capping material as one which has (1) no vapor pressure, (2) no solubility in the semiconductor, (3) no solubility for the semiconductor or its dopants, (4) no metallurgical reactions with the semiconductor, except that it forms (5) covalent bonds to all of the atoms at the semiconductor surface (thus no Schottky reactions occur at the surface and no trapped gas phase can form). As discussed in Section 5, real caps do not appear to approximate ideal caps for long anneal times. However, during the initial period of a typical anneal cycle, atomic exchange between the solid and vapor may be effectively cut off and it is worth considering the implications of this.

2.4.1. Ideal sealed cap covering Si

For a Si wafer which is isolated from the world by an ideal elemental capping material, we have $C = 3$ (Si, ideal cap material, and an inert gas used to set the total pressure), and $P = 3$ (wafer, cap, and gas). This gives $F = 2$, the same as for a Si wafer and vapor in the presence of an inert gas and no capping material. Thus, it is a formal *necessary condition* to choose two independent thermodynamic variables, i.e. the temperature and total pressure, to *define* the equilibrium values of the dependent thermodynamic variables.

However, defining the values of the dependent variables is not *sufficient* to actually bring the variables to those values. For example, it is clear that the vapor pressure of the Si is defined by the choice of T and P , but no atoms can sublime from the wafer if it is covered by an ideal cap and thus the vapor never equilibrates with the wafer. Similarly, the equilibrium native defect concentrations are defined, but there is no means for them to attain those values in an experiment. The interstitial and vacancy concentrations are physically defined by at least two of the following three reactions,



Only the first reaction is physically available when the wafer is encapsulated by an ideal cap. The latter two reactions cannot occur if the wafer cannot exchange atoms with the vapor. One may also consider alternative reactions to vapor-defect exchange, but these involve the surface and the surface is tied up by unbreakable covalent bonds (at least for the ideal cap). Indeed, for any experiment, it is physically impossible for the defect concentrations to reach their equilibrium values solely through a single *chemical pathway*, i.e. Eq. (42). However, there is an equation of constraint which describes the conservation of atoms in a unit volume, i.e.

$$N_{\text{Si}} = N_{\text{sites}} + [\text{I}_{\text{Si}}] - [\text{V}_{\text{Si}}] \quad (45)$$

When Eq. (45) is used with Eq. (42) the two equations define the two steady-state native defect concentrations uniquely. However, those concentrations are not the equilibrium concentrations and they depend upon the history of the crystal. This can be appreciated by considering a simple example.

At the beginning of an anneal, the native defect concentrations clearly depend upon the past history of the crystal. We shall assume that the native defects have reached their equilibrium concentrations, $[\text{V}]_1$ and $[\text{I}]_1$, during a prior anneal at temperature, T_1 , and thus $K_{23}(T_1) = [\text{V}]_1[\text{I}]_1$. If the encapsulated crystal is annealed at $T_2 > T_1$, then $K_{23}(T_2) > K_{23}(T_1)$ because of the increased disorder expected at higher temperatures. We may write

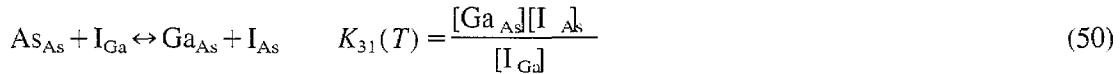
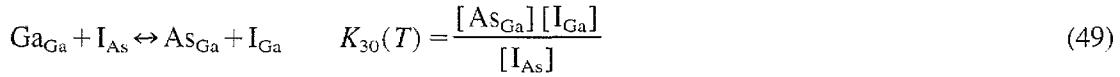
$$K_{23}(T_2) = [\text{V}]_1[\text{I}]_2 = \{[\text{V}]_1 + \delta[\text{V}]\}\{[\text{I}]_1 + \delta[\text{I}]\} \approx [\text{V}]_1[\text{I}]_1 + \delta[\text{V}]\{[\text{V}]_1 + [\text{I}]\} \quad (46)$$

because $\delta[\text{V}] \equiv \delta[\text{I}]$ when changes in the native defect concentration can occur only by the single chemical pathway, Eq. (42). Thus, the only possible way that the native defects could be at their equilibrium values at T_2 is if $[\text{V}]_2 = [\text{I}]_2$ at both temperatures. This is generally believed not to be the case, and Eq. (46) shows that the steady-state (nonequilibrium) native defect concentrations approached at T_2 must depend upon the thermal history of the crystal, i.e. the initial defect concentrations set during prior processing. Thus, even though the equilibrium native defect concentrations are well-defined by Eqs. (42)–(44) they are not attainable when the cap cuts off atomic exchange between the solid and vapor.

2.4.2. Ideal sealed cap covering GaAs

For the case of a GaAs wafer which is covered with an ideal capping material, we have $C=4$ (Ga, As, ideal cap material, and an inert gas which sets the total pressure), $P=3$ (wafer, cap, and gas), and thus $F=3$. Setting both the temperature and the total pressure with the inert gas leaves one degree of freedom undetermined. Indeed, without an external reservoir of Ga and As atoms in contact with the GaAs, the equilibrium state is uncontrolled and not even defined by the experimenter. The third degree of freedom which determines the final native defect concentrations is a composition variable: the precise *initial stoichiometry* of the GaAs. Since this is not a variable which is easily controlled, any properties of the crystal which depend upon native defect concentrations will be difficult to control.

Since an ideal capping material halts all reactions with the vapor, then only the four internal reactions,



exist to relate the six fundamental point defect concentrations. Two additional equations of constraint are needed to describe the final defect concentrations. These equations define the stoichiometry of the crystal through the constant density of group III and V atoms, N_{III} and N_{V} , respectively,

$$N_{\text{III}} = \sum_{\text{Unit vol. of crystal}} \{N_{\text{Ga sites}} - [\text{V}_{\text{Ga}}] + [\text{I}_{\text{Ga}}] + [\text{Ga}_{\text{As}}]\} \quad (51)$$

and

$$N_{\text{V}} = \sum_{\text{Unit vol. of crystal}} \{N_{\text{As sites}} - [\text{V}_{\text{As}}] + [\text{I}_{\text{As}}] + [\text{As}_{\text{Ga}}]\} \quad (52)$$

These equations express the obvious: at the beginning of an anneal, the native defect concentrations depend upon the conditions used during prior high temperature annealing. If the crystal were then annealed in contact with the vapor, a net flux of atoms between the two phases would occur until these concentrations changed to their equilibrium values, as determined by the current degrees of freedom, i.e. T and P_{As_4} . However, under an impervious cap, the six native defect concentrations will eventually reach well-defined steady-state (but not equilibrium) values which are strongly dependent upon the history of the crystal. It is worth considering a specific example.

Consider two undoped crystals of GaAs annealed in contact with the vapor. We shall assume that crystal #1 was annealed until it reached equilibrium at $T = 800$ °C with $P_{\text{As}_4} = 1$ atm, and that crystal #2 was annealed until it reached equilibrium at $T = 800$ °C with $P_{\text{As}_4} = 10^{-4}$ atm. Table 1 shows that all of the native defect concentrations in the two crystals will differ by at least one order of magnitude with this variation in P_{As_4} . We shall also assume that the defect concentrations do not change significantly when cooled down, during encapsulation with an ideal capping material, or during the subsequent heating for a second anneal. We shall consider the second anneal of these crystals, covered with an ideal sealed cap, to be performed simultaneously at $T = 800$ °C and $P_{\text{tot}} = 1$ atm. Since the K s in Eqs. (47)–(52) are a function only of T , and no chemical reactions occur with any other phase during the second anneal, the point defect concentrations in each crystal must remain unchanged at their original values, i.e. *at apparently different equilibrium values*.

The apparent paradox of having two equilibrium states exist in the ‘same’ experiment is resolved by appreciating why the solids are, in fact, in *two* different *equilibrium* states. The phase rule does not dictate which three variables must be controlled, and thus a choice of T , P_{tot} , and any one point defect concentration defines the crystal equilibrium state. An arbitrary choice of one defect concentration (set by the first anneal) in each crystal then defines all of the other defect concentrations, and the stoichiometry of each crystal after the second anneal is a result of the conditions chosen for the first anneal. In contrast, if the second anneal were at a temperature other than $T = 800$ °C, the defect concentrations in each crystal would approach *two* different steady-state, but *nonequilibrium*, values. Clearly, this could complicate the interpretation of diffusion mechanisms if one has not designed an experiment with a clear understanding of the final thermodynamic states.

Of course, a GaAs epitaxial layer is also in contact with a GaAs substrate which may act as a point defect reservoir. This could be important because modern epitaxial layers typically have a volume 2–3 orders of magnitude smaller than that of the substrate. If the epilayer top surface is open to the vapor, perhaps even through a leaky cap, then as the native defect concentrations near that surface approach well-defined equilibrium values, the remainder of the epilayer and substrate gradually approach those values as well. However, if an ideal cap is used, then the concentration of each native defect in the epilayer will be largely determined by interaction with the native defects in the substrate. Since the native defect concentrations in the substrate were originally set during ingot growth, and modified by any later annealing, those concentrations may vary significantly between wafers from different ingots, wafers from the seed and tail end of the same ingot, and even radially across the same wafer. We suspect that this accounts for at least some of the substantial scatter reported for diffusivities in capped GaAs, as summarized in Section 5.

2.5. Diffusion mechanisms and point defects

Diffusion may be discussed using a quasi-thermodynamic approach to model a system's relaxation toward an equilibrium state, or by a kinetic approach which models the specific mechanistic pathways. It is often overlooked that the assumptions contained in both approaches are intimately linked to specific experimental designs. We shall develop a few of the basic concepts behind each approach in this section, and then apply those concepts to real examples throughout the remainder of this paper. We also note here that there are subtle, and rigorous, differences between diffusion and interdiffusion. However, we shall often ignore those differences because they are of secondary importance compared with the very significant experimental problems which have caused large variations in the reported diffusivity and the inferred diffusion mechanisms in GaAs.

Common ways to calculate D include (a) use of the PL energy shift of partially interdiffused quantum wells, (b) application of the Boltzmann–Matano method to the concentration profile of atoms moving from one material to another [27], and (c) a Gaussian curve fit to a narrow concentration spike which is widening with time. Because of some confusion in the literature, it is worth discussing the latter approach. If a Gaussian describes an impurity concentration profile both before and after annealing, then one may calculate the diffusivity, $D = (L_f^2 - L_i^2)/4t$, where t is the time of the anneal and L_i , L_f are the initial and final characteristic diffusion lengths, i.e. the initial and final standard deviations, respectively. It is worth noting that $C(x) = C_{\max} \exp\{-(x/L)^2\}$ and that $C(x)$ changes very rapidly with position. Occasionally an erroneous description of atomic diffusion is used based upon the expression, $C(x) = C_{\max} \exp\{-(x/L)\}$, i.e. an expression describing a combination of *electronic carrier diffusion* and recombination. Thus, at a distance, x , from the atomic concentration peak, i.e. for $x = L$, $\sqrt{2}L$, $\sqrt{3}L$, or $2L$, one expects approximately $C(x) = 0.37$, 0.14 , 0.06 , or $0.02C_{\max}$, respectively.

2.5.1. Simple transport model

Equilibrium for an impurity is described by a constant chemical potential for that impurity throughout the system. It is important to keep in mind that equilibrium is dynamic, i.e. equal fluxes of a species continually move forward and backward across any arbitrary surface. A one-dimensional example of a generalized driving force for the transport of an impurity may be written [28] in terms of the gradient of its chemical potential energy,

$$F = \frac{\partial \mu}{\partial x} \quad \dots \dots \dots \quad (53)$$

where, locally, one may describe the chemical potential

$$\mu = \mu^{\text{std}} + kT \ln N \quad (54)$$

for a sufficiently dilute concentration, N .

An impurity flux, J , can be related in general to the product of N and an average velocity, v , for the impurities. For small velocities, it is reasonable to assume that J varies linearly in response to the generalized driving force, F ,

$$J = Nv = NMF = -MkT \frac{\partial N}{\partial x} = -D \frac{\partial N}{\partial x} \quad (55)$$

where v is related to the generalized force through a generalized mobility, M , which can in turn be related to the impurity diffusivity, D .

If there is a net flow of impurity atoms in or out of the volume of crystal being considered, then the concentration of those atoms in that volume must be changing with time. This is expressed formally as

$$\nabla \cdot J = - \frac{\partial N}{\partial t} \quad (56)$$

which, for one dimension, yields the simple transport equation,

$$\frac{\partial N}{\partial t} = \frac{\partial}{\partial x} \left\{ D \frac{\partial N}{\partial x} \right\} \quad (57)$$

When D is independent of position, then the transport equation may be solved analytically given appropriate boundary conditions. For example, a complementary error function is obtained if the concentration is held constant at a surface, or a gaussian function is obtained if a constant total number of impurity atoms diffuses from an originally narrow distribution.

The simple thermodynamic approach to describing diffusion does not entirely ignore point defects, but rather assumes that their properties do not change with position. Thus, the model can avoid discussing specific diffusion mechanisms. Indeed, Eq. (57) is quite useful if D does not vary with position. However, if D varies with position, then either the concentration or the mobility of the native defects have changed, and this needs to be accounted for in a model. Some physical causes of a change in D include (1) the composition of the host material changes, and this is likely to change both the equilibrium concentrations and jump rates of point defects, (2) any electric field present will change the local concentration of charged point defects, and (3) a nonequilibrium concentration of one point defect may lead to changes in the concentration of other point defects and have strong indirect effects on diffusion. We shall discuss specific examples of each of these in Section 5.

2.5.2. Diffusion mechanisms

Typically, a description of diffusion as described by Eq. (57) is of limited use when the impurity diffusivity changes with position. An atomistic model of the diffusion is generally required to relate the diffusion to position-dependent native defect concentrations. We shall introduce some of the models of diffusion mechanisms below.

2.5.2.1. Direct exchange

In principle, atoms might exchange position on nearest-neighbor sites or via a larger loop in a variation of what has been called a ring mechanism, a small version of which is shown in Fig. 7. If diffusion of an impurity via a direct exchange mechanism occurred in GaAs, the diffusivity would be

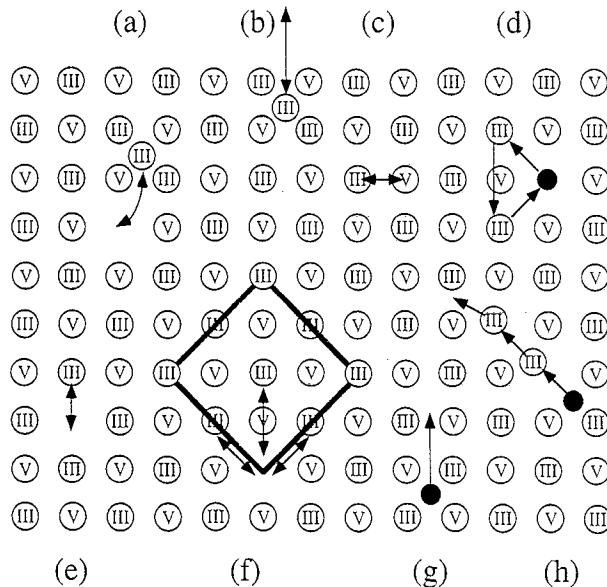


Fig. 7. Schematic illustration of several diffusion mechanisms in a III-V semiconductor. Impurity atoms are shown as filled circles. (a) Interstitial and vacancy generation-recombination via a Frenkel reaction. (b) Group III interstitial generation or annihilation via condensation or sublimation. (c) Direct exchange between nearest neighbors on opposite sublattice will generate two antisite defects. (d) Indirect, concerted, exchange through a simple ring on the same sublattice. (e) Exchange with nearest neighbor vacancy on opposite sublattice. (f) Exchange with next-nearest neighbor vacancy (or nearest neighbor on the same sublattice). The shaded diamond shows one face of the unit cell, and possible jumps to the vacancy from three (of the nearest twelve) group III sites are shown. (g) Impurity moving by pure interstitial mode. (h) Interstitialcy, or kick-out, mechanism showing an impurity replacing a group III atom, which may in turn replace other group III or dopant substitutionals nearby.

insensitive to P_{As} , and isoelectronic impurity markers on each sublattice should exhibit comparable diffusivities. However, diffusion measurements generally find a dependence upon P_{As} , and often sharply different diffusivities are associated with each sublattice. Thus, mechanisms involving the direct exchange of atoms are seldom considered in GaAs.

2.5.2.2. Interstitial defects

Atoms or ions moving solely through the interstices of the crystal are moving via a pure interstitial mechanism. Such motion may occur rapidly if little energy is required for these jumps to take place between interstitial sites. This is a fairly ‘ideal’ form of diffusion since the diffusivity, D , in Eq. (55) reflects essentially the jump frequency of the atom, and the interstitial sites do not change their concentration since they are essentially 100% empty. Thus, Eq. (57) should apply to neutral interstitial atoms, and to interstitial ions if no significant electric fields are present. We note that many transition metals diffuse interstitially and have a low substitutional solubility. Presumably this is related to their tendency to form electronic states deep in the bandgap, i.e. the effective mass approximation does not apply because the atoms do not form strong covalent bonds to the host semiconductor.

2.5.2.3. Vacancy defects

Hu [29] appears to be the first to suggest that the relatively fast diffusion of impurities, compared with isotopic self-diffusion, in Si could be explained by a reduced activation energy for diffusion caused by electrostatic attraction between a charged substitutional ion and a nearby vacancy of opposite charge. One could also imagine that such a model might apply as well to GaAs. However, we consider it appropriate to first consider the simplest possible mechanistic models (without electrostatic inter-

action), and to consider more complex models when simple models cannot explain the results of well designed and conducted experiments. Since we consider ion-defect pair interactions to be an unneeded complication in interpreting the available data for GaAs, we shall avoid further discussion of defect pairs in this section.

In the simplest view of a vacancy mechanism, a substitutional impurity atom may not move until a neighboring substitutional site becomes vacant. Diffusion occurs by a vacancy mechanism when a ‘direct exchange’ occurs between the atom and the vacancy. The vacancy mechanism is sometimes assumed to be an ‘ideal’ mechanism as well. In that case, it is implicitly assumed that the density of vacancies does not change, i.e. the probability of an impurity jump is constant over time and position. However, if the density of vacancies were to increase for some reason, then D of an impurity would increase simply because impurities could jump more often. Assuming that both the impurity and vacancy concentrations are low enough to be considered as dilute solutions, then D is expected to scale linearly with the vacancy concentration. Theoretical considerations to be discussed in Section 3 indicate that vacancies may exist in either charged or neutral states, and we shall find it useful to consider an example of vacancies which exist in three charge states: positive, negative, and neutral. The measured impurity diffusivity may be modeled as if impurity atoms moved via three independent parallel pathways, each pathway corresponding to vacancies in a single charge state. With this description, the measured diffusivity of an impurity in an intrinsic elemental crystal, $D_{m,i}$, is expected to be described by

$$D_{m,i}(T) = D_{V+,i}(T) + D_{V-,i}(T) + D_{V^0,i}(T) \quad (58)$$

which is a sum of independent impurity diffusivities associated with vacancies in the charge state shown, and where each diffusivity corresponds to a standard state (the intrinsic crystal) denoted by the subscript, i . Each of the three contributions to the diffusivity above are well-defined because the equilibrium concentration of each vacancy is well-defined, as discussed in Section 2.2.4, and the rate of attempted jumps is dependent only on temperature.

If the concentration of the vacancies changes for any reason, then the contribution from each term will scale up or down accordingly. One often rewrites Eq. (58) as

$$D_m(T) = D_{V+,i}(T) \frac{[V^+]}{[V^+]_i} + D_{V-,i}(T) \frac{[V^-]}{[V^-]_i} + D_{V^0,i}(T) \frac{[V^0]}{[V^0]_i} \quad (59)$$

to indicate how the contribution to D_m from vacancies in each charge state will change when their actual concentration, $[V]$, is different than the equilibrium concentration in the intrinsic crystal, $[V]_i$. In an elemental crystal, the equilibrium concentrations, $[V]_i$, of each defect and the three intrinsic diffusivities, $D_{V,i}$, are determined solely by the chosen temperature and P_{tot} . If the defect concentrations are kept close to their equilibrium values in one experiment, and somehow kept at different values in a second experiment, then a comparison of any resulting marker layer diffusion in the two experiments can indicate whether a vacancy mechanism controls the diffusion for the specific impurity.

In GaAs, there is evidence that impurities, which normally reside upon one sublattice, move via vacancies associated only with that sublattice. It is reasonable to generalize Eq. (59), originally used to describe vacancy-assisted diffusion in Si, to the diffusion of impurities which substitute on the group III sublattice in GaAs. This was originally accomplished by defining a vacancy mechanism associated solely with the Ga sublattice,

$$D_m(T) = D_{V_{Ga,i}^+}(T) \frac{[V_{Ga}^+]}{[V_{Ga}^+]_i} + D_{V_{Ga,i}^-}(T) \frac{[V_{Ga}^-]}{[V_{Ga}^-]_i} + D_{V_{Ga,i}^0}(T) \frac{[V_{Ga}^0]}{[V_{Ga}^0]_i} \quad (60)$$

and assuming that T defined the standard state for the vacancy concentrations [8]. This was later corrected to include the two degrees of freedom, T and P_{As} [6], required as discussed in Section 2.2.3 and Section 2.2.4. This means that measured data need to be taken at the same (or at least similar) values of P_{As} to be directly compared, or else adjustments must be made in the data to compare results obtained under very different P_{As} . Thus, any plot of D versus $1/T$ is meaningless unless it is defined at a standard P_{As} condition.

In this paper, we choose to compare diffusivities at a standard $P_{\text{As}_4} = 1 \text{ atm}$ because it is simple to scale the standard diffusivities to the actual P_{As_4} used in an experiment. Thus, we shall consider the measured impurity diffusivity in an intrinsic crystal annealed under $P_{\text{As}_4} = 1 \text{ atm}$ as a sum of contributions from independent diffusion pathways associated with vacancies in different charge states, i.e. a sum of the standard diffusivities $D_{V_{\text{Ga},i}^+}(T, P_{\text{As}_4} = 1)$, $D_{V_{\text{Ga},i}^-}(T, P_{\text{As}_4} = 1)$, etc. The measured impurity diffusivity in an intrinsic crystal annealed under $P_{\text{As}_4} \neq 1 \text{ atm}$ can then be related to these standard diffusivities, after they have been scaled by changes in the vacancy concentration associated with the different value of P_{As} , i.e.

$$D_m(T, P_{\text{As}_4}) = D_{V_{\text{Ga},i}^+}(T, P_{\text{As}_4} = 1) \frac{[V_{\text{Ga}}^+]_{i, P_{\text{As}_4}}}{[V_{\text{Ga}}^+]_{i, P_{\text{As}_4} = 1}} + D_{V_{\text{Ga},i}^-}(T, P_{\text{As}_4} = 1) \frac{[V_{\text{Ga}}^-]_{i, P_{\text{As}_4}}}{[V_{\text{Ga}}^-]_{i, P_{\text{As}_4} = 1}} + D_{V_{\text{Ga},i}^0}(T, P_{\text{As}_4} = 1) \frac{[V_{\text{Ga}}^0]_{i, P_{\text{As}_4}}}{[V_{\text{Ga}}^0]_{i, P_{\text{As}_4} = 1}} \quad (61)$$

Since the dependence of the equilibrium vacancy concentration upon P_{As_4} is independent of the vacancy charge state, then for any given charge state,

$$\frac{[V_{\text{Ga}}]_{i, P_{\text{As}_4}}}{[V_{\text{Ga}}]_{i, P_{\text{As}_4} = 1}} = P_{\text{As}_4}^{1/4} \quad (62)$$

and Eq. (61) simplifies to

$$D_m(T, P_{\text{As}_4}) = \{D_{V_{\text{Ga},i}^+}(T, P_{\text{As}_4} = 1) + D_{V_{\text{Ga},i}^-}(T, P_{\text{As}_4} = 1) + D_{V_{\text{Ga},i}^0}(T, P_{\text{As}_4} = 1)\} P_{\text{As}_4}^{1/4} \quad (63a)$$

or

$$D_m(T, P_{\text{As}_4}) = \sum_{Q = \text{All charge states}} D_{V_{\text{Ga},i}^Q}(T, P_{\text{As}_4} = 1) P_{\text{As}_4}^{1/4} \quad (63b)$$

if vacancies in any arbitrary charge state, Q , are to be considered. This result will be generalized further in Section 3.4 for the extrinsic crystal, and evidence from several experiments relating to this model will be provided in Section 5.

Clarification is required to appreciate why the model summarized by Eqs. (63) and (63) should be referred to as an equilibrium model of vacancy-controlled diffusion. If the native defects in the GaAs have equilibrated with the vapor, then the diffusivity of any grown-in marker element will be dependent upon a unique equilibrium concentration of vacancies. Of course, the model cannot be indiscriminately applied to the results of every experiment because the assumption that the native defect concentrations quickly approach their equilibrium values depends critically upon the experimental design and must be verified. It needs to be emphasized that the results obtained from equilibrium experiments, i.e. results which find that $D_{\text{III}} \propto P_{\text{As}_4}^{1/4}$, do not prove that a vacancy mechanism occurs. This is because there are other point defects (i.e. I_{As}) and defect complexes which have the same P_{As} dependence as V_{Ga} and they are associated with entirely different mechanisms which might also describe the observed equilibrium diffusion. However, we consider that the wisest interpretation is the simplest one until evidence from a well designed experiment shows that the interpretation fails. Thus,

we shall consider that measurements of $D_{\text{III}} \propto P_{\text{As}_4}^{1/4}$ do provide reasonable evidence for diffusion via a Ga vacancy.

Vacancies were once believed to control the diffusion of all dopants in metals and semiconductors. However, in GaAs strong experimental evidence that Zn moved by a mechanism other than a vacancy mechanism has been given by Kendall [3]. In a $t = 10$ min closed ampoule experiment, Zn was diffused into GaAs at $T = 900$ °C to form a pn junction a few microns below the surface. Pieces of the GaAs were then removed and placed in other ampoules without Zn present, and reannealed at the same temperature for various times up to 15 h. Further diffusion of Zn was nil in all cases. These simple experiments conclusively show that the rapid in-diffusion of Zn into GaAs from the vapor did not take place by a vacancy mechanism because rapid diffusion would have continued during the second group of anneals if vacancies had controlled the initial diffusion.

2.5.2.4. Interstitial–substitutional mechanism

One alternative to a vacancy mechanism is the so-called interstitial–substitutional (I–S) model. After Kendall's results, this was suggested as the Zn diffusion mechanism [2], i.e. a mechanism by which Zn atoms move interstitially in the crystal until they occupy group III vacancies, $I_{\text{Zn}} + V_{\text{Ga}} \rightarrow \text{Zn}_{\text{Ga}}^- + h^+$. For the sake of completeness, we note that different groups sometimes have made slightly different assumptions about the charge state of the defects. The mass-action expressions associated with this reaction have been used to describe the observed non-Fickian Zn concentration profiles, and this was taken as strong evidence that the I–S model actually applied. However, the results are based on an equilibrium reaction, and such evidence should be taken as suggestive rather than conclusive. Indeed, if the model did apply, then marker interdiffusion, i.e. Al in GaAs, would presumably slow down as P_{Zn} were raised because the increased I_{Zn} concentration would reduce the vacancy density. In fact, the opposite occurs [30], as D_{III} is increased by many orders of magnitude when Zn is diffusing into the crystal. This cannot be explained by an I–S model and it means that an I–S model applies to neither Zn in-diffusion from the vapor nor the accelerated interdiffusion associated with Zn. A model which appears to account for a wide range of observations is the interstitialcy mechanism.

2.5.2.5. Interstitialcy, or kick-out, mechanism

A model of transport which is consistent with several observations in GaAs is the interstitialcy model, which is outlined in Fig. 7. In the simplest version of this model, a highly mobile interstitial impurity becomes an immobile substitutional when it displaces, i.e. kicks out, a host group III atom from its substitutional site. This reaction can be summarized by $I_{\text{Zn}} + \text{Ga}_{\text{Ga}} \leftrightarrow \text{Zn}_{\text{Ga}}^- + I_{\text{Ga}}^\pm$, but various assumptions have been made by different groups regarding the specific charge states. The impurity is assumed to remain in a substitutional site until the presence of another group III atom causes the impurity to become interstitial and to get replaced by the group III atom, i.e. the reaction moves to the left. It is important to note that a number of subtle variations of this model have been used and that they are seldom distinguished from each other by different authors. In this paper, we adopt the simplest version of the model, i.e. we assume that the interstitials move independently of the substitutional atoms or ions. In contrast, Tan and Gosele have assumed that I–S pairs are formed, and apparently move together throughout the crystal. This appears to be analogous to Hu's vacancy–impurity pair formation model suggested for Si. Appropriate nonequilibrium experiments are required to distinguish between the two versions, and the evidence presented in Section 5 suggests that the simpler version applies adequately. The degree to which the interstitialcy mixes with the interstitial mechanism, i.e. the distance that interstitial atoms may move without displacing a substitutional atom, is unclear at this time.

Decades ago, it was noticed that smooth pn junctions could be formed by Zn diffusion from the vapor when using a high group V partial pressure. The smooth pn junctions were found to change to

rough, irregular, junctions in several III-V compounds [2,31,32] when microscopic liquid metal droplets formed on the crystal during annealing under very low group V partial pressure. These reports provide the first indication that the high concentration of group III interstitials below those droplets diffused into the nearby pn junction and caused the roughened junction by increasing the diffusion of Zn below those droplets. This result cannot be accounted for by a vacancy mechanism and it provides evidence that Ga interstitials push Zn off of substitutional sites.

Early evidence for the interstitialcy mechanism was obtained from observations of the sharply increased interdiffusion of GaAs-AlAs superlattices when Zn was diffused in from the vapor. As discussed above, this observation cannot be accounted for by an I-S mechanism. The dramatic reduction of both interdiffusion and Zn diffusion observed (in heavily Zn-doped crystals) when annealing with $P_{\text{Zn}} = 0$ is strong evidence that a vacancy mechanism cannot account for the observations. Gosele and Morehead [33] appear to be the first to suggest that the interstitialcy mechanism applies when Zn diffuses into a GaAs-AlGaAs superlattice. It is worth considering the implications of a few simple calculations. Under commonly used annealing conditions, Zn can diffuse from the vapor into $\sim 1000 \text{ \AA}$ of GaAs in a few minutes. SIMS shows that a superlattice will measurably interdiffuse during this time if $N_{\text{Zn}} \sim 10^{19} \text{ cm}^{-3}$, i.e. when an areal density of only $N'_{\text{Zn}} \sim 10^{14} \text{ cm}^{-2}$ exists. This is particularly striking because the number of group III atoms is $\sim 1000 \times$ higher than the number of Zn atoms in the heavily doped region of the crystal, and yet virtually all of the group III atoms must have moved by several lattice constants for their interdiffusion to be measured. Therefore, each Zn atom is associated with many thousands of group III atom jumps. Several modern experiments which offer further insight into the interstitialcy are discussed in Section 5.

Just as with vacancies, we may build a simple semi-empirical model to describe the interstitialcy by assuming that the substitutional impurity moves only when it is displaced by a host interstitial atom. Just as with vacancies, this approach avoids discussing any fundamental physics, and makes an effort to relate observations to a minimum number of parameters, i.e. the empirically obtained ‘standard diffusivities’. Once again, we shall define standard conditions to be at $P_{\text{As}_4} = 1 \text{ atm}$ for convenience, and we shall assume that the impurity diffusion in the intrinsic crystal can be separated into contributions from host atom interstitials in two different charge states, i.e. in neutral, $D_{I_{\text{Ga},i}^0}(T, P=1)$, and singly charged positive states, $D_{I_{\text{Ga},i}^+}(T, P=1)$. Thus, if the interstitialcy mechanism controls the diffusion of an impurity, then the measured diffusivity, D_m , of that impurity in intrinsic GaAs, for a chosen T and P_{As_4} , is related to the diffusivity at standard conditions, but adjusted for the actual concentrations of the host interstitials,

$$D_m(T, P_{\text{As}_4}) = D_{I_{\text{Ga},i}^+}(T, P_{\text{As}_4}=1) \frac{[I_{\text{Ga}}^+]_{i, P_{\text{As}_4}}}{[I_{\text{Ga}}^+]_{i, P_{\text{As}_4}=1}} + D_{I_{\text{Ga},i}^0}(T, P_{\text{As}_4}=1) \frac{[I_{\text{Ga}}^0]_{i, P_{\text{As}_4}}}{[I_{\text{Ga}}^0]_{i, P_{\text{As}_4}=1}} \quad (64)$$

From Section 2.2.3, the dependence of the Ga interstitial concentration upon P_{As} may be used to simplify Eq. (64) to

$$D_m(T, P_{\text{As}_4}) = \{D_{I_{\text{Ga},i}^+}(T, P_{\text{As}_4}=1) + D_{I_{\text{Ga},i}^0}(T, P_{\text{As}_4}=1)\} P_{\text{As}_4}^{-1/4} \quad (65a)$$

or

$$D_m(T, P_{\text{As}_4}) = \sum_{Q=\text{All charge states}} D_{I_{\text{Ga},i}^Q}(T, P_{\text{As}_4}=1) P_{\text{As}_4}^{-1/4} \quad (65b)$$

if interstitials in an arbitrary charge state, Q , are to be considered in the intrinsic crystal. Similar equations may be set up for impurity diffusion associated with the As sublattice. The results above will be generalized further in Section 3.4 for the extrinsic crystal. At present, there appears to be little reliable evidence for the obvious modifications of the vacancy and interstitialcy models, i.e. I_{As}

displacing an impurity on the group III sublattice and forming an As_{Ga} antisite defect, V_{As} providing a pathway for impurities normally associated with the group III sublattice, divacancy pairs, or other point defect complexes. Thus, we shall not consider those alternatives here.

There is an obvious difference between the predictions made by the vacancy ($D_{\text{III}} \propto P_{\text{As}_4}^{1/4}$) and interstitialcy ($D_{\text{III}} \propto P_{\text{As}_4}^{-1/4}$) models which is measurable. In practice, the measured diffusivity could be a sum of diffusion by both mechanisms. This suggests that there might be a crossover between mechanisms, i.e. D_{III} may increase either (1) with increasing P_{As} via vacancies at high P_{As} , or (2) with decreasing P_{As} via the interstitialcy at low P_{As} . Based upon interdiffusion data from two different groups taken over a ~ 1000 -fold variation in P_{As} , Cohen [34] suggested that a crossover might occur between the vacancy and interstitialcy mechanism near $P_{\text{As}_4} = 1$ atm in undoped GaAs. Tan and Gosele used similar reasoning to draw conclusions about the standard diffusivities in p-type GaAs. However, the more recent and much more complete data summarized in Section 5 show that both of the above conclusions are wrong because they were based upon noisy and unreliable data.

3. Point defect equilibria in extrinsic GaAs

3.1. Fundamental concepts applied to practice

3.1.1. Gibbs phase rule and phase diagrams

In this section, we shall discuss open flow versus closed ampoule systems. As discussed in Section 2.1.2, the thermodynamics describing GaAs and a proximity cap in an open system is essentially the same as the thermodynamic description of GaAs in a closed system. However, there are some critical differences and those will be discussed in this section. For either system containing solid GaAs, a vapor phase, and a dopant, $F = 3$. Thus, a choice of three independent thermodynamic variables, i.e. T , P_{As} , and P_{dop} , is a *necessary* condition to define the equilibrium values of the dependent thermodynamic parameters (i.e. point defect concentrations) without an unwanted change in the number of condensed phases. The *sufficient* condition for attaining solid–vapor equilibrium is satisfied when the experimental design makes it possible for the dependent thermodynamic parameters to actually attain those equilibrium values during the time of the experiment. There are at least two good reasons for having a clear understanding of the equilibrium state: (1) some useful diffusion experiments can actually be conducted when the native defect concentrations are at their equilibrium values, and (2) some useful interpretations of nonequilibrium results can be made when one knows which equilibrium state the system is relaxing toward.

3.1.1.1. Closed system

It is useful to recall that Casey and Panish were the first to understand the thermodynamics of the Ga–As–Zn ternary system and thus were the first to obtain reproducible diffusion of Zn into GaAs. They elegantly explained [35] a variety of diffusion results in the Ga–As–Zn system in terms of both the phase rule and the phase diagram. For the Ga–As–Zn system, the phase rule gives $F = 5 - P$, where P is the number of phases present in the ampoule. From a practical standpoint, it is desirable to have $F = 1$ so that the temperature, and no other parameter, needs to be controlled. Since the vapor phase is always present, this implies that the GaAs sample plus 2 condensed phases should be placed in the ampoule.

Casey and Panish showed [35] that an isothermal slice of the phase diagram near $T = 700$ °C can be approximated as shown in Fig. 8. Tie lines were shown to exist between GaAs and ZnAs₂, GaAs and Zn₃As₂, and GaAs and one Zn + Ga-rich liquid mixture. In a closed system, the average composition

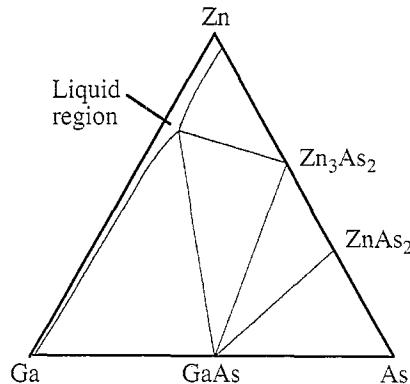


Fig. 8. $T = 700$ °C slice of Ga–As–Zn ternary phase diagram, adapted from Ref. [35]. Tie lines show the condensed phases with which GaAs may equilibrate. Reprinted with permission, from Ref. [35].

of the ternary system can be described by a single point. When the average composition is inside one of the triangles shown, then the three condensed phases defined by the corners of the triangle will coexist indefinitely at equilibrium. Thus, a thermodynamically well-determined closed system could be defined at the beginning of an experiment by combining GaAs with either (1) $\text{ZnAs}_2 + \text{As}$, (2) $\text{ZnAs}_2 + \text{Zn}_3\text{As}_2$, or (3) a $\text{Zn}_3\text{As}_2 + \text{liquid} (\text{Zn} + \text{Ga} + \text{As})$ mixture. The latter mixture eventually forms when only GaAs and Zn are placed together in an ampoule. An average composition lying between GaAs and Zn (or in any Ga-rich region) is undesirable because the GaAs wafer must decompose for the liquid phase to reach its equilibrium concentration. Since the time to reach equilibrium will depend upon the total mass chosen for the GaAs and Zn, and since the effects of the nonequilibrium ambient early in the diffusion cycle cannot be easily reproduced, the Zn concentration profile after diffusion was found to vary substantially from run to run in the early 1960s. In addition, dissolution of the GaAs into the liquid phase was often associated with undesirable damage to the GaAs surface.

Casey and Panish demonstrated that a very convenient and reproducible choice for a Zn diffusion source was a combination of solid Zn_3As_2 and ZnAs_2 (each saturated with Ga). When the average composition is inside the triangle defined by Zn_3As_2 , ZnAs_2 , and GaAs, then the vapor quickly equilibrates with the solid compounds, all partial pressures are insensitive to the mass of each phase, and no damage occurs to the GaAs surface. Thus, when the near-surface region of the GaAs wafer is close to equilibrium with the vapor for essentially all of the anneal period, the defect concentrations near the surface are reproducible and this gives rise to reproducible diffusion into the solid. Shih [36] has shown that an average composition in the GaAs– ZnAs_2 –As triangle, or on the tie line connecting GaAs and ZnAs_2 , will also give high quality reproducible results. When the average composition of the system lies in the narrow region defined by the GaAs– ZnAs_2 tie line, then P_{As} varies between a unique value defined in the GaAs– ZnAs_2 –As triangle and a different unique value defined in the GaAs– ZnAs_2 – Zn_3As_2 triangle. Of course, this occurs because there is one less condensed phase when the system composition lies on a tie line, and the extra degree of freedom which appears is satisfied by setting P_{As} . However, it needs to be emphasized that this is not a practical approach to diffusion because minor errors in weighing would readily push the average composition off of the tie line.

This brief discussion about the Ga–As–Zn ternary is intended to illustrate the need for knowledge about *both* the phase rule and the ternary phase diagram when working with dopants in a closed system. However, relatively few ternary phase diagrams have been mapped out. One is also forced to assume that no significant chemical reactions occur with the ampoule walls or with residual oxidants, and this

may be incorrect as discussed in Section 2.1.2. An additional practical consideration is that the entire process of ampoule diffusion is simply less convenient than open tube annealing.

3.1.1.2. Open system

When using a flowing gas in an open tube for annealing, one generally wants to avoid the formation of additional condensed phases. In addition, if one wants the grown-in dopant concentration in the epilayer to be at equilibrium with the vapor, then the correct partial pressure of that dopant must be added to the gas stream. These issues are effectively addressed by controlling the appropriate degrees of freedom. For a four component Ga–As–dopant–inert gas system composed of only a wafer and vapor, the phase rule gives $F = 4$. Control of T , P_{As} , P_{dop} , and P_{inert} satisfies the *necessary* condition to define the equilibrium state. The inert gas is, of course, optional and present only to bring the total pressure up to a convenient value, say 1 atm. An appropriate experimental design which makes it possible for the dependent thermodynamic variables to rapidly reach their equilibrium values satisfies the *sufficient* condition for a system to actually attain this well-defined equilibrium state.

As discussed in Section 2.1.2, one convenient and practical way to quickly approach equilibrium is to use a GaAs proximity cap. If the dopant is reasonably volatile, then it can be easily added to the gas stream. However, several dopants have extremely low volatility and there is no practical way to control their partial pressure in an open tube system. In such cases, the experimental design must be reconsidered, as discussed in Section 3.2.5.

3.1.2. Solubility versus solubility limit

3.1.2.1. Solubility

An interesting ‘technological’ problem from the days preceding modern semiconductor processing sheds some light on the concept of solubility, i.e. the dependence of the impurity concentration in solution upon the impurity partial pressure. At one time, it had been found difficult to reproducibly equilibrate HCl in the aqueous phase with the partial pressure of anhydrous HCl in the gas phase [37]. The reaction can be described as



and, at equilibrium the chemical potentials are related by

$$\mu_{\text{HCl}}^\circ + kT \ln P_{\text{HCl}} = \mu_{\text{H}^+}^\circ + kT \ln [\text{H}^+] + \mu_{\text{Cl}^-}^\circ + kT \ln [\text{Cl}^-] \quad (67)$$

Assuming the usual approximations of ideality, the reaction is thus described by

$$K_{68}(T) = \frac{[\text{H}^+][\text{Cl}^-]}{P_{\text{HCl}}} \quad (68)$$

If a buffering agent were used to keep the pH constant, it was found that the chloride ion solubility varied linearly with the partial pressure,

$$[\text{Cl}^-] \propto P_{\text{HCl}} \quad (69)$$

However, if pure water were used to begin with, then $[\text{H}^+] = [\text{Cl}^-]$ at equilibrium, and the chloride ion solubility varied parabolically with partial pressure, i.e.

$$[\text{Cl}^-] \propto P_{\text{HCl}}^{1/2} \quad (70)$$

Similarly, impurities which dissolve into semiconductors as charged pairs, i.e. an ion and an electronic carrier, follow a similar behavior. Thus, equilibrium between a neutral dopant in the vapor,

and the ion plus electronic carrier in the solid, must be considered together. Any changes in the solubility between a linear and parabolic dependence upon the dopant partial pressure may have significant implications, and this important issue is discussed further in Section 3.2.

3.1.2.2. Solubility limit

If the dopant partial pressure is increased sufficiently, then an additional condensed phase will form and this will define a solubility limit for the dopant in GaAs. In the case of the GaAs wafer + vapor system with a high partial pressure of Zn added to the vapor, this second condensed phase may be $ZnAs_2$, Zn_3As_2 , a liquid Zn + Ga + As mixture, or even a combination of two of these condensed phases. The solubility limit obtained is intimately related to the second condensed phase formed, and it will *not be* a function of temperature alone. Relevant examples are explicitly discussed in Section 3.2.3.

It is also worth noting that the solubility limit is routinely exceeded in many situations. For example, the dopant concentration in most electronic devices is in excess of the solubility limit at room temperature. Of course, the dopant concentration is metastable, i.e. diffusion is so slow that there is no practical effect on a time scale of interest. However, evidence is presented in Section 4 that dopant concentrations above the solubility limit, corresponding to the growth or annealing ambient, are often associated with so-called ‘anomalous’ rapid diffusion.

3.1.3. Electrochemical potential

When charged species can move between regions of different electric potential, then the electrochemical potential should be used to describe the system. Conceptually, the electrochemical potential, $\tilde{\mu}$, accounts for changes in the energy of charged particles by separating energy changes associated with the electric potential from those associated with the chemical potential. We shall consider a charged species i which is associated with arbitrary charge of a given magnitude and sign, $z_i = 0, \pm 1, \pm 2, \dots$. To the extent that it is possible to separate the chemical potential energy from the electrostatic potential energy associated with the particle in a phase at a potential, ϕ , one may write the electrochemical potential,

$$\tilde{\mu}_i = \mu_i + z_i q \phi \quad (71)$$

where q is the unit charge. It is worth considering a simple application of this equation and one of the critical assumptions underlying that application. We shall consider the simple step pn junction, i.e. the interfacial region separating one part of a semiconductor crystal assumed to be uniformly doped with donors, and the remaining part which is assumed to be uniformly doped with acceptors. A space charge region, or ‘band bending region’, results between the n- and p-type semiconductors, as shown in Fig. 9.

We shall apply Eq. (71) to the equilibration of electrons, for which $z_i = -1$. A common name for the electrochemical potential of electrons is the Fermi energy, i.e. $E_f \equiv \tilde{\mu}_e$. For convenience, we place $x = 0$ where E_f crosses E_i , the Fermi energy for an intrinsic crystal. This position is also where $n = n_i$, the carrier concentration for the intrinsic crystal. Thus, we have

$$\tilde{\mu}_e(x) = \mu_e^{\text{Std}} + kT \ln(n(x)/n_i) - q\phi(x) \quad (72)$$

Defining $\phi(0) = 0$, then the standard state for electrons is at $x = 0$, i.e. $\mu_e^{\text{Std}} = E_f(0) = E_i(0)$, where $n = n_i$. It is clear that the position dependence of the potential energy E_i is described by the relative value of the electric potential, i.e. $E_i(x) = \mu_e^{\text{Std}}(0) - q\phi(x)$. Thus, Eq. (72) may be rewritten as

$$E_f(x) = E_i(x) + kT \ln(n(x)/n_i) \quad (73)$$

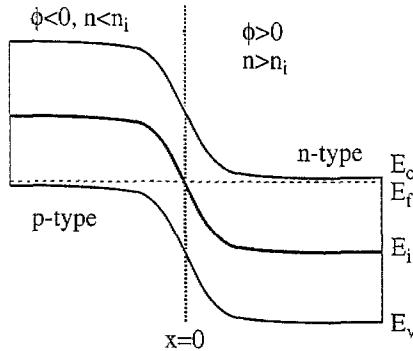


Fig. 9. Energy band diagram for a step pn junction, i.e. a junction in which the donor and acceptor concentrations are uniform on their respective sides of the metallurgical junction. At equilibrium, the Fermi energy is flat, and the intrinsic Fermi energy follows the conduction and valence bands, E_c and E_v , respectively. The electrochemical potential relates changes in the equilibrium concentration of charged point defects and electronic carriers to changes in electrical potential.

When equilibrium holds for the electrons throughout the crystal, then the Fermi energy has a constant value everywhere, and Eq. (73) gives the well known relationship,

$$n(x) = n_i \exp\left\{\frac{E_f - E_i(x)}{kT}\right\} \quad (74)$$

One critical assumption was made to obtain Eq. (74) which is *never* correct, i.e. the assumption that equilibrium holds throughout the entire crystal. Strict equilibrium demands that the concentration of all species: donors, acceptors, and carriers be at a uniform value in each material phase. However, in practice, we make the useful assumption that the donors and acceptors are frozen in place for an unmeasurably long time, and that the relatively fast species (electrons and holes) equilibrate around these immobile ions. This is an acceptable approximation to equilibrium over time scales of human interest.

As will be discussed in Sections 3.2.4 and 3.2.5, the same assumption can be made for native defects which move more rapidly than the ions which set the potential variation in the semiconductor. Thus, rapidly diffusing native defects may quickly approach their equilibrium values in adjacent layers containing different dopant concentrations. Of course, slower native defects may not approach their equilibrium values in any practical time. When considering the joint effect of electric fields, partial pressures, and the rate of defect diffusion, several limiting cases should be considered and these will be discussed throughout Section 3. We note that a description of any point defect concentration versus position becomes significantly more complicated when n_i and μ^{Std} become a function of position, i.e. when the composition of the material changes with position, and we shall not pursue such a description in this paper. However, this complication clearly suggests that impurity markers should be kept at concentrations which are as low as practical.

3.1.4. Counting statistics

Fermi statistics apply to the occupation of states by electrons. However, there are subtle differences between the electron states associated with impurity atoms and those associated with native defects. In particular, the total density of impurity atoms is generally independent of the Fermi energy, but the total density of native defects depends strongly on the Fermi energy. This leads to subtle, but important, differences in the description of electron states associated with atoms versus native defects. For simplicity, we shall (1) ignore the effects of spin degeneracy in the following discussion, (2) assume an elemental doped semiconductor which is in contact with the vapor, i.e. $F = 2$, and (3) assume that

T and P_{dop} have been set, and thus the value of all the equilibrium defect concentrations are well defined.

We shall first consider electron states associated with shallow acceptor atoms. The ionized acceptor concentration, N_A^- , is related to the total acceptor concentration, N_A , multiplied by the Fermi-Dirac probability of that state being occupied,

$$N_A^- = N_A \frac{1}{1 + \exp\{(E_A - E_f)/kT\}} \quad (75)$$

The total concentration of the ionized plus neutral states is independent of the Fermi energy, and it is independent of temperature, to first order.

The situation is entirely different with point defects because there is no mass conservation requirement to restrict the total defect density. To illustrate this point, we shall consider the example of a vacancy which can be in only one of two charge states: neutral or negative. The two charge states are related via

$$V^0 + e^- \leftrightarrow V^- \quad K_{76} = \frac{[V^-]_e}{[V^0]_n} = \frac{[V^-]_i}{[V^0]_n} \quad (76)$$

where the subscripts, e and i, refer to crystals which are doped either extrinsic or intrinsic, respectively. The extrinsic and intrinsic cases are both shown in Eq. (76) because this equation is often used to relate changes in vacancy concentration to changes in carrier concentration. In Section 2.2.2, it was shown that the equilibrium $[V^0]$ was determined by temperature alone. This and Eq. (76) clearly show that the equilibrium $[V^-]$ varies with the electron concentration, n , as well as temperature. For simplicity, we shall assume that the carrier concentration is not degenerate, and consider the special case when $[V^-] = [V^0]$, i.e.

$$K_{76}(T) = \frac{1}{n_{-/0}} = \frac{1}{n_i \exp\{(E_{-/0} - E_i)/kT\}} \quad (77)$$

where $n_{-/0}$ is the electron concentration when the Fermi energy just equals the energy of the electronic state, $E_f = E_{-/0}$ at which $[V^-] = [V^0]$. For other values of n , we can use Eq. (77) to eliminate K_{76} from Eq. (76),

$$[V^-] = K_{76}[V^0]n = \frac{[V^0]n_i \exp\{(E_f - E_i)/kT\}}{n_i \exp\{(E_{-/0} - E_i)/kT\}} \quad (78)$$

Straightforward algebra can be used to show that Eq. (78) is equivalent to

$$[V^-] = [V]^{\text{Total}} \frac{1}{1 + \exp\{(E_{-/0} - E_f)/kT\}} \quad (79)$$

where the equilibrium $[V]^{\text{Total}} = [V^-] + [V^0]$ clearly varies with E_f because the equilibrium $[V^0]$ stays constant at the chosen temperature, as shown by Eq. (25). This is in sharp contrast to the equilibrium total $N_A = N_A^- + N_A^0$ which remains constant as N_A^0 and N_A^- each vary with E_f . Both cases, of course, are functions of the two degrees of freedom chosen, i.e. T and P_{dop} (or equivalently, E_f , if the solid and vapor have equilibrated). We note that some published simulations have been based on the incorrect assumption that the total density of each native defect is independent of Fermi energy, and that the interested reader needs to exercise caution in this area.

3.2. Equilibrium concentration of charged point defects

The purpose of this section is to relate the equilibrium concentration of charged impurities and native defects in GaAs to the degrees of freedom which are actually controlled. In recent years, diffusivity and native defect concentrations often have been assumed to be set by the carrier concentration. This approach implicitly assumes that the carrier concentration has equilibrated with the independent variables chosen. However, if the carrier concentration is not in equilibrium with the controlled variables, then the defect concentrations will drift with time until the system eventually relaxes to the equilibrium state. Several of the equilibrium relationships developed in this section will be useful in understanding the results of experiments discussed in Sections 4 and 5.

3.2.1. Impurity solubility

3.2.1.1. Solubility of single substitutional dopant

Dopant atoms are generally used to control the concentration of either electrons or holes, and thus set the Fermi energy in an extrinsic crystal. Atoms from the group VIA column of the periodic chart substitute on the As sublattice, and act as electron donors while atoms from the group IIA column substitute on the Ga sublattice and act as electron acceptors. Atoms from the group IVA column of the periodic chart are amphoteric and often can reside on the either the Ga or As sublattice and act as donors or acceptors, respectively.

We shall discuss a generic acceptor species, A. A thermodynamic description of solid-vapor equilibrium for this species depends upon which sublattice it is associated with. For example, if this acceptor sits on the Ga sublattice, then



or, with the aid of Table 1,

$$K_{80} = \frac{N_A^- p}{P_A [V_{\text{Ga}}^0]} \propto \frac{N_A^- p}{P_A P_{\text{As}_4}^{1/4}} \quad (80)$$

and one concludes that

$$p = N_A^- \propto P_{\text{II}}^{1/2} P_{\text{As}_4}^{1/8} \quad \text{for group II acceptor} \quad (81)$$

assuming that the Fermi energy is set by the fully ionized acceptor.

However, if A resides on the As sublattice, we must describe the situation via



or, with the aid of Table 1,

$$K_{82} = \frac{N_A^- p}{P_A [V_{\text{As}}^0]} \propto \frac{N_A^- p}{P_A P_{\text{As}_4}^{-1/4}} \quad (82)$$

and one concludes that

$$p = N_A^- \propto P_{\text{IV}}^{1/2} P_{\text{As}_4}^{-1/8} \quad \text{for a group IV acceptor} \quad (83)$$

assuming that the Fermi energy is set by the fully ionized acceptor.

It is worth pointing out that the results of Eqs. (81) and (83) are written in terms of the degrees of freedom which are actually controlled, T , P_{As} , and P_A . The equilibrium hole density varies with the

1/2 power of the acceptor partial pressure in both cases, but the dependence upon As partial pressure differs because P_{As} affects dopant solubility on the two sublattices differently.

It is worth emphasizing that our use of the neutral vacancy in Eqs. (80) and (82) is for our mental convenience only. The choice of a reaction to relate the solid and vapor is arbitrary and has no effect on the resulting expression. Even though Weisberg [38] pointed this out decades ago, there are several examples in the recent literature which incorrectly infer the defect charge state from the behavior of impurity solubility. Thermodynamic expressions alone cannot be used to deduce the mechanistic pathway taken by atoms between two phases. This can be appreciated by simply adding an electron to each side of the reaction of Eq. (80). Superficially, this appears to lead to an entirely different reaction,



or

$$K_{84} = \frac{N_{\text{A}}^-}{P_{\text{A}}[\text{V}_{\text{Ga}}^-]} \quad (84)$$

This mass action expression appears to predict $N_{\text{A}} \propto P_{\text{A}}$ in the extrinsic crystal, but it must make the same prediction as Eq. (81) because it is exactly the same reaction. Indeed, this ‘paradox’ can be resolved by realizing that

$$\mu_{\text{V}_{\text{Ga}}^-} = \mu_{\text{V}_{\text{Ga}}^0} + \mu_{\text{e}^-}$$

or equivalently Eq. (32), shows that Eq. (84) is precisely the same as Eq. (80). Indeed, the vacancy itself is an unnecessary artifice for predicting the solubility of a dopant, and one may use interstitials or even no defects at all to mediate the solid–vapor interaction and the same description of the equilibrium dopant solubility must always be obtained. However, charged point defect models are needed to interpret the measured results of nonequilibrium processes, and thus it is worth developing and understanding point defect models for the simplest possible situation, i.e. at equilibrium.

Using a closed ampoule, it is difficult to demonstrate the dependence of the solubility upon $P_{\text{dop}}^{1/2}$ as calculated above because it is difficult to change only a single partial pressure at a time without the formation of unexpected condensed phases. Indeed, Casey and Panish [35,39] derived the relationship Eq. (81) and elegantly discussed how these experimental difficulties often arised because of the phase diagram. Using a closed ampoule, they found it necessary to vary both P_{Zn} and P_{As} simultaneously to obtain a wide range of data before concluding that Eq. (81) described the equilibrium Zn solubility in GaAs.

Using an open tube system, Reichert et al. [21] appear to be the first to report the results for Zn solubility, measured after Zn diffused into undoped GaAs, when using different values of P_{Zn} , but the same T and P_{As} . Test samples were covered by GaAs proximity caps to assure that all equilibrium partial pressures were approached. The solubilities were obtained from measurements of N_{Zn} , taken by SIMS 20 nm below the surface (where the Zn concentration was changing only slowly). The measured N_{Zn} are shown in Fig. 10. The solid line has a slope of 1/2, and it is in close agreement with the data.

The solubility was confirmed by comparing N_{Zn} obtained from in-diffusion to out-diffusion [22]. Epitaxial layers were grown containing either $N_{\text{Zn}} = 8 \times 10^{19} \text{ cm}^{-3}$, $N_{\text{Zn}} = 3 \times 10^{19} \text{ cm}^{-3}$, or no added Zn were annealed together under ambient partial pressures conditions believed to correspond to equilibrium with $N_{\text{Zn}} \approx 3 \times 10^{19} \text{ cm}^{-3}$ at $T = 700^\circ\text{C}$. The Zn concentration profile for the sample grown with high Zn concentration is shown before and after annealing in Fig. 11(a). It is clear that the Zn concentration drops to $N_{\text{Zn}} \approx 3 \times 10^{19} \text{ cm}^{-3}$ after annealing. Fig. 11(b) makes it clear that this is the

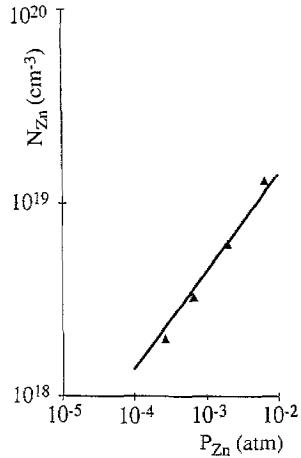


Fig. 10. Solubility, N_{Zn} , measured at various values of P_{Zn} , at the same values of T and P_{As} . N_{Zn} was measured by SIMS and these values were taken approximately 20 nm below the surface after in-diffusion of Zn from the vapor. Adapted from Ref. [21].

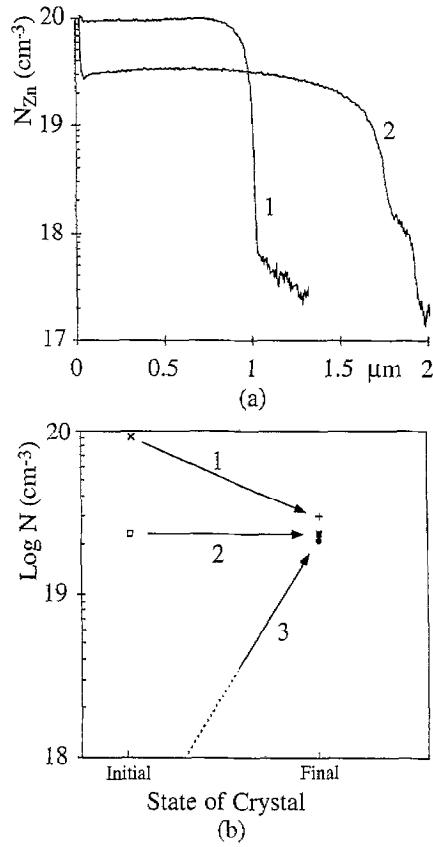
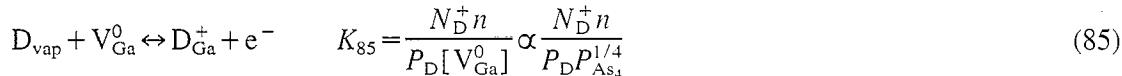


Fig. 11. (a) SIMS concentration profiles showing that out-diffusion from a supersaturated as-grown concentration of Zn (profile 1) quickly approaches the equilibrium solubility, N_{Zn} , (curve 2) as set by the chosen T , P_{As} , and P_{Zn} . Adapted from Ref. [22]. (b) Convergence to the same equilibrium state, i.e. the same N_{Zn} under the same ambient P_{As} and P_{Zn} at $T = 700$ °C, for three crystals with very different initial states. These include (1) substantial out-diffusion from the epilayer shown in (a), which began with an as-grown $N_{\text{Zn}} \approx 9 \times 10^{19} \text{ cm}^{-3}$, (2) minimal change for an epilayer with an initial as-grown $N_{\text{Zn}} \approx 2.5 \times 10^{19} \text{ cm}^{-3}$, and (3) in-diffusion into an initially undoped epilayer from the vapor. This result is conclusive evidence for the equilibrium state originally proposed by Reichert et al. [21].

equilibrium value because the same result is achieved after annealing independent of whether the initial N_{Zn} in the GaAs was higher, lower, or the same as this value.

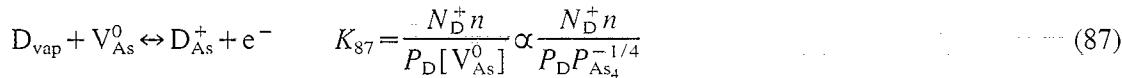
For a generic shallow donor species, D, residing on either sublattice and setting the Fermi energy, we also obtain two different results depending upon which sublattice the donor generally sits. If a fully ionized group IV donor resides on the Ga sublattice, then



which implies

$$n = N_D^+ \propto P_{\text{IV}}^{1/2} P_{\text{As}_4}^{1/8} \quad \text{for a group IV donor} \quad (86)$$

If a fully ionized group VI donor resides on the As sublattice, then



implies that

$$n = N_D^+ \propto P_{\text{VI}}^{1/2} P_{\text{As}_4}^{-1/8} \quad \text{for a group VI donor} \quad (88)$$

Given the relatively low diffusivity of donors, it is unlikely that an effort to experimentally confirm either Eq. (86) or Eq. (88) will ever be made in GaAs. The behavior of the donor and acceptor solubilities have been summarized in Table 2. It should be emphasized that a linear $N_{\text{dop}} \propto P_{\text{dop}}$ relationship is never expected to be observed when the impurity is ionized, and this will lead to important conclusions in Section 4.

Group IV elements often exhibit amphoteric behavior, i.e. they can reside on both sublattices and thus act as either a donor or acceptor. Assuming low concentrations of charged native defects, then electrical neutrality can be used to determine the relative fraction of a given group IV species which acts as donors and acceptors,

$$N_D^+ + p = N_A^- + n \quad (89)$$

We may apply Eqs. (82) and (85) to Eq. (89) to obtain

$$\frac{n^2}{n_i^2} = \frac{K_{85} P_{\text{IV}} P_{\text{As}_4}^{1/4} + n_i^2}{K_{82} P_{\text{IV}} P_{\text{As}_4}^{-1/4} + n_i^2} \quad (90)$$

If we assume that the dopant acts primarily as a donor, i.e. $K_{85} > K_{82}$, then at relatively low values of P_{IV} , the numerator of Eq. (90) increases with P_{IV} while the denominator remains essentially

Table 2

Predicted dependence of equilibrium substitutional dopant concentration in extrinsic GaAs upon ambient conditions. Only one ionized impurity is assumed to be present

| Parameter | $\propto f(T, P_{\text{Dop}}, P_{\text{As}})$ |
|------------------------------------|---|
| N_{II}^- | $P_{\text{II}}^{1/2} P_{\text{As}_4}^{1/8}$ |
| N_{IV}^- | $P_{\text{IV}}^{1/2} P_{\text{As}_4}^{-1/8}$ |
| N_{IV}^+ | $P_{\text{IV}}^{1/2} P_{\text{As}_4}^{1/8}$ |
| N_{VI}^+ | $P_{\text{VI}}^{1/2} P_{\text{As}_4}^{-1/8}$ |
| $R = N_A^- / N_D^+$ | $n^2 P_{\text{As}_4}^{-1/2}$ |
| $n_{\text{max}} (R \rightarrow 1)$ | $P_{\text{As}_4}^{1/4}$ |

constant. However, when P_{IV} is increased sufficiently, then the expression approaches a limiting value. In this case, the dopant may substantially compensate itself. This is sometimes described by a compensation ratio, R ,

$$R = \frac{N_{\text{A}}^-}{N_{\text{D}}^+} = \frac{K_{82} n^2}{K_{85} n_i^2} P_{\text{As}_4}^{-1/2} \propto n^2 P_{\text{As}_4}^{-1/2} \quad (91)$$

which has a value that varies between 0 (no impurity atoms occupying the group V sublattice) and 1 (50% of the impurity atoms on each sublattice). In principle, the As pressure dependence could be used to distinguish between the limit imposed on carrier concentration by amphoteric versus chemical solubility limit. In practice, however, the two are not simple to distinguish given the noise typically present in solubility measurements.

3.2.1.2. Solubility of a second substitutional charged impurity

Intentionally controlling the concentration of a minor dopant, i.e. a charged impurity at a low concentration in an otherwise heavily doped semiconductor, is not commonly done. However, dual doping can provide a useful tool for studying point defects, and an important experiment based on this is discussed in Section 4. With the addition of two dopants to a vapor-solid GaAs system, there are 4 degrees of freedom requiring control, and in practice a typical choice of variables to control would be T , P_{As} , P_{dop_1} , and P_{dop_2} . We shall consider the major dopant, dop_1 , as the dopant which sets the Fermi energy, and the minor impurity, dop_2 , as a charged impurity with a concentration too low to affect the Fermi energy. The solubility of the major dopant has been described in Table 2. The solubility of the minor impurity is still described by either Eqs. (80), (82), (85) or (87) but now the carrier concentration has been set by the major dopant. The results are summarized in Table 3, and they show a linear dependence upon the minor impurity partial pressure.

The effect of the major dopant upon the minor impurity can be seen in a simple example. If Zn were the only dopant present in system #1, then the equilibrium hole concentration would depend upon three degrees of freedom, $p_1 = N_{\text{Zn},1}(T, P_{\text{As}}, P_{\text{Zn}})$. If, in system #2, the same T , P_{As} , and P_{Zn} were used, but a larger hole concentration, p_2 , were set by a second group II dopant (at equilibrium with the vapor), then Eq. (80) shows that four degrees of freedom determine the Zn solubility, $N_{\text{Zn},2}$, and that

$$N_{\text{Zn},2}\{T, P_{\text{As}}, P_{\text{Zn}}, p_2\} = N_{\text{Zn},1}\{T, P_{\text{As}}, P_{\text{Zn}}\} \frac{p_1}{p_2} \quad (92)$$

Table 3

Predicted dependence of equilibrium substitutional charged impurity concentration upon ambient conditions, for a second (minor) impurity which is at a sufficiently low concentration that it does not affect the Fermi energy. The second dopant concentration is shown as a function of its own partial pressure, P_{dop_2} , P_{As} , and the carrier concentration set by the major dopant, which is assumed to be in equilibrium with its own vapor at P_{dop_1} . Note that the P_{As} dependence shown must be modified further by one of the terms in Table 2, depending upon the sublattice on which the major dopant resides

| Minor dopant concentration (dopant #2) | $\propto f(T, P_{\text{As}}, n(P_{\text{Dop}_1}), P_{\text{Dop}_2})$ |
|---|--|
| N_{II}^- | $P_{\text{As}_4}^{1/4} n P_{\text{II}}$ |
| N_{IV}^- | $P_{\text{As}_4}^{-1/4} n P_{\text{IV}}$ |
| N_{IV}^+ | $P_{\text{As}_4}^{1/4} p P_{\text{IV}}$ |
| N_{VI}^+ | $P_{\text{As}_4}^{-1/4} p P_{\text{VI}}$ |

Thus, the equilibrium solubility of N_{Zn} incorporated in the GaAs will be reduced if $p_2 > p_1$, and increased if an n-type dopant sets the Fermi energy. Experimental results which demonstrate this behavior will be discussed in Section 5.1. Indeed, the utility of using an ionized dopant as a marker species will be appreciated throughout the remainder of this paper.

3.2.1.3. Solubility of interstitial impurity

The concentration of neutral dopant interstitial atoms was discussed in Section 2.3. Charged interstitial dopant atoms will equilibrate with the same three degrees of freedom that substitutional dopant atoms equilibrate with. If only one dopant is present, and if its interstitial concentration is much lower than its substitutional concentration (which determines the Fermi energy), then Table 4 summarizes the relationship between the equilibrium dopant interstitial concentrations and the three degrees of freedom which may be taken as $(T, P_{\text{As}}, P_{\text{dop}})$ or (T, P_{As}, n) , provided solid-vapor equilibrium can be approximated. The leftmost column of Table 4 specifies the type of the substitutional dopant, i.e. donor or acceptor, which generally characterizes the impurity. Pairs of rows are used to describe each type of impurity: the first row describes an interstitial in an arbitrary charge state, ι , and the second row illustrates the special case of $\iota = +1$. The two columns at the right of Table 4 describe the concentration dependence of the interstitial upon the controlled degrees of freedom. Superficially, those two columns appear to be in conflict, but they can easily be shown to be identical when n and P_{dop} are in equilibrium. When n and P_{dop} are not at equilibrium, then one must cautiously consider which (if either) column applies to a specific experimental situation.

Most of the entries in Table 4 can be easily derived and the results are more or less intuitive. However, it is worth briefly exploring the curious result that the relationship between the group IV charged interstitial concentration and the group IV partial pressure depends upon the sublattice where the substitutional group IV atoms reside. This can be appreciated by describing equilibration between the group IV interstitials and the group IV atoms in the vapor,



or

$$K_{92} = \frac{[\text{I}_{\text{IV}}^{\iota+}]n^{\iota}}{P_{\text{IV}}} \Rightarrow [\text{I}_{\text{IV}}^{\iota+}] \propto P_{\text{IV}} p^{\iota} \quad (93)$$

Table 4

Predicted dependence of the equilibrium interstitial dopant concentration upon ambient conditions, assuming that the majority of the dopant atoms are substitutional and set the Fermi energy. The superscript ι denotes a charge of magnitude ι associated with the interstitial. Negative or zero values for ι may also be used. Examples are given for the case $\iota = +1$

| Substitutional dopant | Corresponding interstitial | $\alpha f(n, P_{\text{As}})$ | $\alpha f(P_{\text{Dop}}, P_{\text{As}})$ |
|-----------------------|----------------------------------|-------------------------------------|---|
| N_{II}^- | $[\text{I}_{\text{II}}^{\iota}]$ | $p^{2+\iota} P_{\text{As4}}^{-1/4}$ | $P_{\text{II}}^{(\iota+2)/2} P_{\text{As4}}^{\iota/8}$ |
| N_{II}^- | $[\text{I}_{\text{II}}^{+1}]$ | $p^3 P_{\text{As4}}^{-1/4}$ | $P_{\text{II}}^{3/2} P_{\text{As4}}^{1/8}$ |
| N_{IV}^- | $[\text{I}_{\text{IV}}^{\iota}]$ | $p^{2+\iota} P_{\text{As4}}^{1/4}$ | $P_{\text{IV}}^{(\iota+2)/2} P_{\text{As4}}^{-\iota/8}$ |
| N_{IV}^- | $[\text{I}_{\text{IV}}^{+1}]$ | $p^3 P_{\text{As4}}^{1/4}$ | $P_{\text{IV}}^{3/2} P_{\text{As4}}^{-1/8}$ |
| N_{IV}^+ | $[\text{I}_{\text{IV}}^{\iota}]$ | $n^{2-\iota} P_{\text{As4}}^{-1/4}$ | $P_{\text{IV}}^{(2-\iota)/2} P_{\text{As4}}^{-\iota/8}$ |
| N_{IV}^+ | $[\text{I}_{\text{IV}}^{+1}]$ | $n P_{\text{As4}}^{-1/4}$ | $P_{\text{IV}}^{1/2} P_{\text{As4}}^{-1/8}$ |
| N_{VI}^+ | $[\text{I}_{\text{VI}}^{\iota}]$ | $n^{2-\iota} P_{\text{As4}}^{1/4}$ | $P_{\text{VI}}^{(2-\iota)/2} P_{\text{As4}}^{\iota/8}$ |
| N_{VI}^+ | $[\text{I}_{\text{VI}}^{+1}]$ | $n P_{\text{As4}}^{1/4}$ | $P_{\text{VI}}^{1/2} P_{\text{As4}}^{1/8}$ |

The equilibrium hole concentration clearly depends upon which sublattice the group IV element sits on. Replacing P_{IV} in Eq. (93) by either Eq. (83) or Eq. (86) results in the two different forms shown in the third column of Table 4. Similarly, we may rewrite Eq. (93)

$$[I_{\text{IV}}^+] \propto P_{\text{IV}} P_{\text{IV}}^{-\nu/2} P_{\text{As}_4}^{-\nu/8} \propto P_{\text{IV}}^{(2-\nu)/2} P_{\text{As}_4}^{-\nu/8} \quad \text{for a group IV donor} \quad (94)$$

or

$$[I_{\text{IV}}^+] \propto P_{\text{IV}} P_{\text{IV}}^{+\nu/2} P_{\text{As}_4}^{\nu/8} \propto P_{\text{IV}}^{(2+\nu)/2} P_{\text{As}_4}^{-\nu/8} \quad \text{for a group IV acceptor} \quad (95)$$

to obtain the forms shown in the fourth column of Table 4.

For an uncharged interstitial, the concentration is simply a linear function of P_{IV} , just as described in Section 2.3. However, reactions such as $\text{I}^0 + \text{p}^+ \leftrightarrow \text{I}^+$ are driven to the right in p-type GaAs, but to left in n-type GaAs. Thus, the concentration of charged interstitials are dependent upon carrier concentration as described in Eq. (93). The identical dependence upon P_{As} shown in Eqs. (94) and (95) can be understood by appreciating that in n-type GaAs, an increased P_{As} leads to a higher group IV donor concentration, and the higher electron concentration causes the reduced $[I_{\text{IV}}^+]$, but in p-type GaAs, an increased P_{As} causes a lower group IV acceptor and hole concentration, and this also gives a reduced $[I_{\text{IV}}^+]$.

3.2.2. Native point defect solubilities

Kroger has described [7] the concentration of the six native point defects in GaAs as a function of T , P_{As} , and the carrier concentration, n or p . With minor modifications, this description has been used to relate diffusion to either the carrier concentration [8], or both [6] the carrier concentration and P_{As} . Such treatments are correct to the extent that the carrier concentration has equilibrated with any external phases, i.e. the vapor. However, the model of defect equilibration with carrier concentration may fail if solid–vapor equilibrium does not hold. For example, in Section 4 we shall show that the grown-in dopant concentration is often far from equilibrium with the partial pressures used for growth, and this result appears to be intimately associated with very rapid diffusion as a heavily doped crystal relaxes back toward equilibrium. To appreciate how the native defect concentrations may change during an anneal, and thus affect diffusion, requires a clear understanding of their dependence upon the experimental conditions chosen. Thus, this section describes how the equilibrium native defect concentrations are affected by the degrees of freedom that one typically chooses, i.e. the temperature and partial pressures.

We shall follow the present custom of assuming that the calculations of Baraff and Schluter [40] are qualitatively correct, i.e. the six point defects may exist in multiply charged states which depend upon the Fermi energy, and a large fraction of two defect concentrations, $[\text{V}_{\text{Ga}}]$ and $[\text{Ga}_{\text{As}}]$, will be negatively charged while a large fraction of the other four defect concentrations will be positively charged. This custom will be reflected in our notation which indicates the sign and magnitude of the anticipated charge with a defect. For example, interstitial Ga atoms will be denoted $I_{\text{Ga}}^{+\alpha}$, to describe the number of positive units of charge, $\alpha = 1, 2, \dots$ associated with the interstitial, and Ga vacancies will be denoted by $V_{\text{Ga}}^{-\beta}$, to describe the number of negative units of charge, $\beta = 1, 2, \dots$ associated with the vacancy. If an experiment determines that the assumed sign of the charge is different, then the value obtained for the superscript will have a negative (or zero) value.

We begin with a discussion of the Ga interstitial. In Table 1, it was shown that the equilibrium concentration of the neutral, $[I_{\text{Ga}}^0]$, is defined solely by two degrees of freedom, i.e. T and P_{As} . At equilibrium, the charged interstitial is simply related to the neutral interstitial via the reaction

$$I_{Ga}^0 \leftrightarrow I_{Ga}^{+\alpha} + \alpha e^- \Rightarrow K_{96} = \frac{[I_{Ga}^{+\alpha}] n^\alpha}{[I_{Ga}^0]} \quad (96)$$

which relates the equilibrium concentration of a Ga interstitial of charge, α , to the carrier density,

$$[I_{Ga}^{+\alpha}] \propto [I_{Ga}^0] / n^\alpha \propto P_{As_4}^{-1/4} p^\alpha \quad (97)$$

This equation is a correct and complete description of $[I_{Ga}^{+\alpha}]$ when the solid has equilibrated with the actual three degrees of freedom controlled. However, it should be clear that the hole concentration is generally not a degree of freedom because it is not directly controlled, but rather it is a function of the three degrees of freedom actually controlled, as shown in Section 3.2.1. If the three degrees of freedom controlled are T , P_{As} , and P_{dop} , we may rewrite Eq. (97) in terms of the variables actually controlled,

$$[I_{Ga}^{+\alpha}] \propto P_{As_4}^{-1/4} P_{II}^{\alpha/2} P_{As_4}^{\alpha/8} \propto P_{II}^{\alpha/2} P_{As_4}^{-(2-\alpha)/8} \quad \text{for group II acceptors} \quad (98)$$

or

$$[I_{Ga}^{+\alpha}] \propto P_{As_4}^{-1/4} P_{IV}^{\alpha/2} P_{As_4}^{-\alpha/8} \propto P_{IV}^{\alpha/2} P_{As_4}^{-(2+\alpha)/8} \quad \text{for group IV acceptors} \quad (99)$$

These expressions are also correct if $\alpha \leq 0$. The different dependence of $[I_{Ga}^{+\alpha}]$ upon P_{As} in the above expressions is a result of how the substitutional group II and IV solubilities depend upon P_{As} , as described by Eqs. (81) and (83).

If a donor rather than an acceptor controls the carrier concentration, then one obtains

$$[I_{Ga}^{+\alpha}] \propto P_{IV}^{-\alpha/2} P_{As_4}^{-(2+\alpha)/8} \quad \text{for group IV donors} \quad (100)$$

and

$$[I_{Ga}^{+\alpha}] \propto P_{IV}^{-\alpha/2} P_{As_4}^{-(2-\alpha)/8} \quad \text{for group VI donors} \quad (101)$$

Similarly, it is worth briefly analyzing the situation for the charged Ga vacancy, $V_{Ga}^{-\beta}$, by considering the reaction

$$V_{Ga}^0 + \beta e^- \leftrightarrow V_{Ga}^{-\beta} \Rightarrow K_{102} = \frac{[V_{Ga}^{-\beta}]}{[V_{Ga}^0] n^\beta} \quad (102)$$

which describes the equilibrium concentration of the negatively charged Ga vacancy in terms of T , P_{As} , and carrier concentration,

$$[V_{Ga}^{-\beta}] \propto [V_{Ga}^0] n^\beta \propto P_{As_4}^{1/4} n^\beta \quad (103)$$

This concentration can also be related to a specific group IV or group VI partial pressure with the aid of Table 2, i.e.

$$[V_{Ga}^{-\beta}] \propto P_{IV}^{\beta/2} P_{As_4}^{(2+\beta)/8} \quad \text{for group IV donors} \quad (104)$$

or

$$[V_{Ga}^{-\beta}] \propto P_{VI}^{\beta/2} P_{As_4}^{(2-\beta)/8} \quad \text{for group VI donors} \quad (105)$$

Using the approach presented above, it is straightforward to relate all of the equilibrium charged native defect concentrations to the chosen thermodynamic degrees of freedom. At a given temperature, the variation of the six native defect concentrations upon the typically controlled variables is summarized for p-type GaAs in Table 5. Since this appears to be the first tabulation of the defect concentrations in terms of the actually controlled variables, the first column shows each point defect in an

Table 5

Predicted dependence of equilibrium native defect concentrations upon ambient conditions, in p-type GaAs. The Fermi energy is assumed to be set by either a single group II or group IV acceptor. Examples are given for some commonly assumed defect charge states

| Defect | $\alpha f(P_{\text{As}}, p)$ | $\alpha f(P_{\text{As}}, P_{\text{II}})$ | $\alpha f(P_{\text{As}}, P_{\text{IV}})$ |
|------------------------------------|-------------------------------------|--|--|
| $[\text{I}_{\text{Ga}}^{+\alpha}]$ | $P_{\text{As}_4}^{-1/4} p^{\alpha}$ | $P_{\text{As}_4}^{-(2-\alpha)/8} P_{\text{II}}^{\alpha/2}$ | $P_{\text{As}_4}^{-(2+\alpha)/8} P_{\text{IV}}^{\alpha/2}$ |
| $\alpha=0$ | $P_{\text{As}_4}^{-1/4}$ | $P_{\text{As}_4}^{-1/4}$ | $P_{\text{As}_4}^{-1/4}$ |
| $\alpha=1$ | $P_{\text{As}_4}^{-1/4} p$ | $P_{\text{As}_4}^{-1/8} P_{\text{II}}^{1/2}$ | $P_{\text{As}_4}^{-3/8} P_{\text{IV}}^{1/2}$ |
| $\alpha=2$ | $P_{\text{As}_4}^{-1/4} p^2$ | P_{II} | $P_{\text{As}_4}^{-1/2} P_{\text{IV}}$ |
| $[\text{V}_{\text{Ga}}^{-\beta}]$ | $P_{\text{As}_4}^{1/4} p^{-\beta}$ | $P_{\text{As}_4}^{(2-\beta)/8} P_{\text{II}}^{-\beta/2}$ | $P_{\text{As}_4}^{(2+\beta)/8} P_{\text{IV}}^{-\beta/2}$ |
| $\beta=0$ | $P_{\text{As}_4}^{1/4}$ | $P_{\text{As}_4}^{1/4}$ | $P_{\text{As}_4}^{1/4}$ |
| $\beta=1$ | $P_{\text{As}_4}^{1/4} p^{-1}$ | $P_{\text{As}_4}^{1/8} P_{\text{II}}^{-1/2}$ | $P_{\text{As}_4}^{3/8} P_{\text{IV}}^{-1/2}$ |
| $\beta=2$ | $P_{\text{As}_4}^{1/4} p^{-2}$ | P_{II}^{-1} | $P_{\text{As}_4}^{1/2} P_{\text{IV}}^{-1}$ |
| $[\text{I}_{\text{As}}^{+\gamma}]$ | $P_{\text{As}_4}^{1/4} p^{\gamma}$ | $P_{\text{As}_4}^{(2+\gamma)/8} P_{\text{II}}^{\gamma/2}$ | $P_{\text{As}_4}^{(2-\gamma)/8} P_{\text{IV}}^{\gamma/2}$ |
| $\gamma=0$ | $P_{\text{As}_4}^{1/4}$ | $P_{\text{As}_4}^{1/2}$ | $P_{\text{As}_4}^{1/4}$ |
| $\gamma=1$ | $P_{\text{As}_4}^{1/4} p$ | $P_{\text{As}_4}^{3/8} P_{\text{II}}^{1/2}$ | $P_{\text{As}_4}^{1/8} P_{\text{IV}}^{1/2}$ |
| $\gamma=2$ | $P_{\text{As}_4}^{1/4} p^2$ | $P_{\text{As}_4}^{1/4} P_{\text{II}}$ | P_{IV} |
| $[\text{V}_{\text{As}}^{+\delta}]$ | $P_{\text{As}_4}^{-1/4} p^{\delta}$ | $P_{\text{As}_4}^{-(2-\delta)/8} P_{\text{II}}^{\delta/2}$ | $P_{\text{As}_4}^{-(2+\delta)/8} P_{\text{IV}}^{\delta/2}$ |
| $\delta=0$ | $P_{\text{As}_4}^{-1/4}$ | $P_{\text{As}_4}^{-1/4}$ | $P_{\text{As}_4}^{-1/4}$ |
| $\delta=1$ | $P_{\text{As}_4}^{-1/4} p$ | $P_{\text{As}_4}^{-1/8} P_{\text{II}}^{1/2}$ | $P_{\text{As}_4}^{-3/8} P_{\text{IV}}^{1/2}$ |
| $\delta=2$ | $P_{\text{As}_4}^{-1/4} p^2$ | P_{II} | $P_{\text{As}_4}^{-1/2} P_{\text{IV}}$ |
| $[\text{As}_{\text{Ga}}^{+\mu}]$ | $P_{\text{As}_4}^{1/2} p^{\mu}$ | $P_{\text{As}_4}^{(4+\mu)/8} P_{\text{II}}^{\mu/2}$ | $P_{\text{As}_4}^{(4-\mu)/8} P_{\text{IV}}^{\mu/2}$ |
| $\mu=0$ | $P_{\text{As}_4}^{1/2}$ | $P_{\text{As}_4}^{1/2}$ | $P_{\text{As}_4}^{1/2}$ |
| $\mu=1$ | $P_{\text{As}_4}^{1/2} p$ | $P_{\text{As}_4}^{5/8} P_{\text{II}}^{1/2}$ | $P_{\text{As}_4}^{3/8} P_{\text{IV}}^{1/2}$ |
| $\mu=2$ | $P_{\text{As}_4}^{1/2} p^2$ | $P_{\text{As}_4}^{3/4} P_{\text{II}}$ | $P_{\text{As}_4}^{1/4} P_{\text{IV}}$ |
| $[\text{Ga}_{\text{As}}^{-\nu}]$ | $P_{\text{As}_4}^{-1/2} p^{-\nu}$ | $P_{\text{As}_4}^{-(4+\nu)/8} P_{\text{II}}^{-\nu/2}$ | $P_{\text{As}_4}^{-(4-\nu)/8} P_{\text{IV}}^{-\nu/2}$ |
| $\nu=0$ | $P_{\text{As}_4}^{-1/2}$ | $P_{\text{As}_4}^{-1/2}$ | $P_{\text{As}_4}^{-1/2}$ |
| $\nu=1$ | $P_{\text{As}_4}^{-1/2} p^{-1}$ | $P_{\text{As}_4}^{-5/8} P_{\text{II}}^{-1/2}$ | $P_{\text{As}_4}^{-3/8} P_{\text{IV}}^{-1/2}$ |
| $\nu=2$ | $P_{\text{As}_4}^{-1/2} p^{-2}$ | $P_{\text{As}_4}^{-3/4} P_{\text{II}}^{-1}$ | $P_{\text{As}_4}^{-1/4} P_{\text{IV}}^{-1}$ |

arbitrary charge state, followed by three examples of charge states for that point defect. Although we have assumed four of the defects to be positive and two to be negative, the table does apply to defects in any charge state. The second column of Table 5 shows the dependence of the equilibrium defect concentration upon arsenic partial pressure and hole concentration, i.e. the form which is commonly used [6,41].

The third and fourth columns describe the defect concentration in terms of P_{As_4} and the acceptor partial pressure, P_{II} or P_{IV} , respectively. When equilibrium holds between the vapor and the extrinsic crystal, then the second and third (or second and fourth) columns are equivalent. The different exponents on the As partial pressure reflects the dependence of the ionized acceptor solubility upon both the dopant and arsenic partial pressures. The experimental design determines which column applies to an experiment and this will become apparent as we discuss examples throughout the remainder of this paper. Other questions arise about the application of Tables 5 and 6 in practical situations. For example, if $P_{\text{dop}}=0$, it is clear that the solid will be in equilibrium with the vapor only when $N_{\text{dop}}=0$. Diffusion seldom proceeds this far in practice, but the tables may be applicable given an appropriate experimental design, as discussed in Section 3.2.5.

Table 6

Predicted dependence of equilibrium native defect concentrations upon ambient conditions, in n-type GaAs. The Fermi energy is assumed to be set by either a single group VI or IV donor

| Defect | $\alpha f(P_{\text{As}}, n)$ | $\alpha f(P_{\text{As}}, P_{\text{VI}})$ | $\alpha f(P_{\text{As}}, P_{\text{IV}})$ |
|------------------------------------|--------------------------------------|---|---|
| $[\text{I}_{\text{Ga}}^{+\alpha}]$ | $P_{\text{As}_4}^{-1/4} n^{-\alpha}$ | $P_{\text{As}_4}^{-(2-\alpha)/8} P_{\text{VI}}^{-\alpha/2}$ | $P_{\text{As}_4}^{-(2+\alpha)/8} P_{\text{IV}}^{-\alpha/2}$ |
| $[\text{V}_{\text{Ga}}^{-\beta}]$ | $P_{\text{As}_4}^{1/4} n^{\beta}$ | $P_{\text{As}_4}^{(2-\beta)/8} P_{\text{VI}}^{\beta/2}$ | $P_{\text{As}_4}^{(2+\beta)/8} P_{\text{IV}}^{\beta/2}$ |
| $[\text{I}_{\text{As}}^{+\gamma}]$ | $P_{\text{As}_4}^{1/4} n^{-\gamma}$ | $P_{\text{As}_4}^{(2+\gamma)/8} P_{\text{VI}}^{-\gamma/2}$ | $P_{\text{As}_4}^{(2-\gamma)/8} P_{\text{IV}}^{-\gamma/2}$ |
| $[\text{V}_{\text{As}}^{+\delta}]$ | $P_{\text{As}_4}^{-1/4} n^{-\delta}$ | $P_{\text{As}_4}^{-(2-\delta)/8} P_{\text{VI}}^{-\delta/2}$ | $P_{\text{As}_4}^{-(2+\delta)/8} P_{\text{IV}}^{-\delta/2}$ |
| $[\text{As}_{\text{Ga}}^{+\mu}]$ | $P_{\text{As}_4}^{1/2} n^{-\mu}$ | $P_{\text{As}_4}^{(4+\mu)/8} P_{\text{VI}}^{-\mu/2}$ | $P_{\text{As}_4}^{(4-\mu)/8} P_{\text{IV}}^{-\mu/2}$ |
| $[\text{Ga}_{\text{As}}^{-\nu}]$ | $P_{\text{As}_4}^{-1/2} n^{\nu}$ | $P_{\text{As}_4}^{-(4+\nu)/8} P_{\text{VI}}^{\nu/2}$ | $P_{\text{As}_4}^{-(4-\nu)/8} P_{\text{IV}}^{\nu/2}$ |

For n-type GaAs, one may develop equilibrium native defect expressions similar to those for p-type GaAs. We have concisely tabulated the results for the six equilibrium native defect concentrations in n-type GaAs in Table 6.

We note for completeness that a variation in the concentration of native defect complexes with the ambient can be easily obtained by using the appropriate tables. For example, one could show, by the reaction $\text{V}_{\text{Ga}}^- + \text{V}_{\text{As}}^+ \leftrightarrow \text{V}_{\text{Ga}}^- \cdot \text{V}_{\text{As}}^+$, that the concentration of neutral divacancies will have an equilibrium concentration which is determined by temperature but is insensitive to the Fermi energy or P_{As} . However, the analytical complexity rises enormously when point defect complexes are considered. The measurements discussed in Section 5 do not present any compelling evidence for divacancies or for other defect complexes, and thus we shall not discuss them further.

3.2.3. Impurity and native point defect solubility limits

3.2.3.1. Impurities

The solubility limit for an impurity in a material is determined by the formation of a new condensed phase. This is easily appreciated for an elemental semiconductor such as Si. For the Si wafer-vapor system which includes a dopant, an equilibrium state is defined by setting T and P_{dop} . The dopant concentration in the solid will increase with P_{dop} until a second condensed phase appears. The second condensed phase may be a liquid or solid solution, or a solid compound, and it will form when P_{dop} is less than the vapor pressure of the pure elemental dopant. With the formation of the second condensed phase, the number of degrees of freedom is reduced by one to $F=1$, i.e. temperature uniquely determines all of the other thermodynamic variables, including the solubility limit. However, the same is not true for a dopant in GaAs because of the additional degree of freedom coming with a binary compound. We shall first discuss the different equilibrium states which can be defined for GaAs and a dopant in a closed ampoule, and then we shall compare those states with the states obtained in a practical open tube system.

3.2.3.1.1. Impurities in closed system

In a closed ampoule containing Ga, As, and dopant, we have $F=5-P$. We would like to bring the system to $F=1$, i.e. we want the equilibrium state well defined when we choose the temperature. Although it is possible to leave $F>1$ and to control partial pressures in addition to T , this approach requires careful weighing of the starting materials, and it is typically avoided because it is inconvenient and prone to errors. To obtain the desired $F=1$, we must have the GaAs, vapor, and two additional condensed phases present in the ampoule. As discussed in Section 3.1.1, an *appropriate choice* of these additional condensed phases can be made only after considering the ternary phase diagram.

Unfortunately, this is more than inconvenient because very few ternary phase diagrams have been mapped out. However, if the appropriate additional phases have been chosen, they determine P_{dop} , P_{As} , and the dopant concentration, N_{dop} , in the GaAs. Of course, this value of N_{dop} is the solubility limit simply because it cannot be increased. For example, if the elemental vapor pressures in the ampoule were higher than their equilibrium values, then growth of the condensed phases would occur and as the excess pressures relaxed back to their equilibrium values, N_{dop} would be at its equilibrium value, i.e. at the solubility limit. However, in a closed system, changing the source compounds leads to a different N_{dop} . It is worth considering the example of Zn in GaAs.

The first correct determination of more than one solubility limit for Zn in GaAs was made by Casey and Panish [35]. After determining much of the ternary phase diagram, they used closed ampoule experiments with an average composition in the triangle region defined by GaAs, Zn_3As_2 , and ZnAs_2 of Fig. 8. Thus, with $F = 1$, a choice of temperature determined a thermodynamically well-defined system. In particular, in the near-surface region of the GaAs, N_{Zn} quickly approached a limiting value, i.e. its solubility limit. Casey and Panish compared the equilibrium N_{Zn} obtained from this experiment with the significantly lower value of equilibrium N_{Zn} measured for experiments performed in the composition triangle defined by the GaAs– Zn_3As_2 –Ga-rich region of the phase diagram. These classic closed ampoule experiments nicely demonstrate that the solubility limit of Zn in GaAs depends upon the region of the phase diagram which defines the system. For comparison below, it is worth noting that with the Zn_3As_2 – ZnAs_2 –GaAs system, Casey and Panish reported that a modest rise in N_{Zn} occurred when going from $T = 650$ °C to $T = 700$ °C.

3.2.3.1.2. Impurities in open system

For an open tube system consisting of vapor, GaAs wafer and a dopant, there are three degrees of freedom, and in practice the parameters often controlled are T , P_{dop} , and P_{As} . As discussed in Section 2.1.1, the addition of an inert gas to control the total pressure does not change this situation because the extra degree of freedom obtained is used up by setting the partial pressure of the inert gas. In practice, H_2 is often added as a carrier gas. Although atomic hydrogen apparently can enter the crystal and interact with point defects, there is no evidence which indicates that molecular hydrogen leads to measurable effects at typical temperatures, and thus we shall treat it as an inert gas.

Using an open tube experiment, Chen et al. [42] measured N_{Zn} diffused into the GaAs over a wide range of Zn partial pressure. At sufficiently high partial pressure, a second condensed phase, Zn_3As_2 , was formed. However, N_{Zn} was found to decrease slightly as temperature increased, opposite to the result reported by Casey and Panish. The major difference between the two experiments was that one value of P_{As} was used in the open tube work at all temperatures, while P_{Zn} was varied. In the closed system, both partial pressures were known to increase with temperature. The phase rule provides insight on how this subtle difference affects the results.

When the two condensed phases, GaAs and Zn_3As_2 , equilibrate with the vapor, then $F = 2$ rather than $F = 1$ (for the three condensed phases of Casey and Panish), and thus the experimental designs are significantly different. In the open tube, the average solid composition lies along a narrow region in the phase diagram defined by the tie line between GaAs and Zn_3As_2 . As the Zn_3As_2 second phase approaches equilibrium with GaAs, i.e. as Zn diffuses sufficiently into the GaAs, the N_{Zn} obtained is a Zn solubility limit which depends upon the two degrees of freedom chosen, i.e. T and P_{As} . With increasing P_{As} , one expects N_{Zn} to increase until ZnAs_2 forms. When ZnAs_2 coexists with the Zn_3As_2 and GaAs, then the open tube experiment is in essentially the same thermodynamic state as that for the closed ampoule of Casey and Panish. Unfortunately, a direct comparison did not get made because the required high P_{As} exceeded the capability of the equipment. Similarly, we note that if P_{As} had been reduced sufficiently in the open system, then a Zn and Ga-rich liquid mixture (determined by the phase

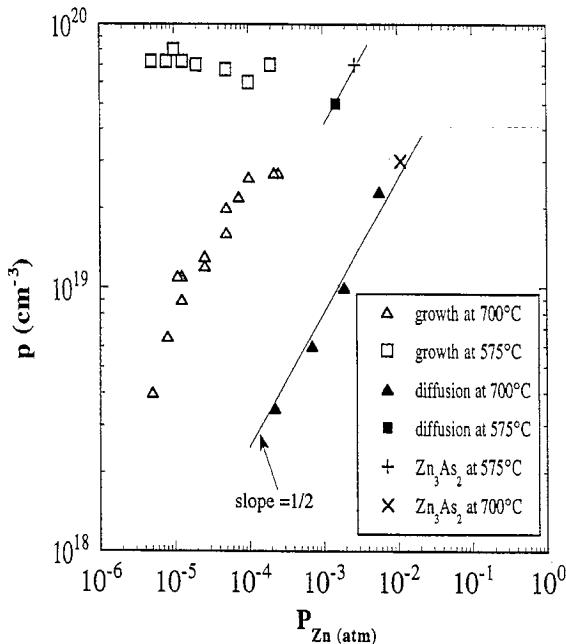


Fig. 12. N_{Zn} from diffusion, which gives the equilibrium N_{Zn} , at 575 °C (filled square) or 700 °C (filled triangles). At a higher value of P_{Zn} , second phase formation, i.e. growth of Zn_3As_2 , occurs at 575 °C (+) or 700 °C (x). The formation of the additional phase defines the solubility limit at the T and P_{As} used. N_{Zn} obtained from OMVPE growth, at $T=575$ °C (open squares) and 700 °C (open triangles), and at the same P_{As} , shows that the solubility limit for growth is essentially the same as that obtained at equilibrium. Adapted from Ref. [42].

diagram) would eventually have equilibrated with the Zn_3As_2 and GaAs, and a much lower solubility limit, presumably that observed by Casey and Panish for these condensed phases, would have been set. Chen et al. did not explore the highest and lowest Zn solubility limits set by P_{As} , but some groups have inadvertently explored them by using poorly conceived experiments.

Chen et al. [42] also compared the solubility limit in their open tube system to the maximum N_{Zn} obtained from OMVPE growth. They reported values for the maximum grown-in N_{Zn} which were nearly identical to the equilibrium solubility limit (for the same T and P_{As}). The results of growth at $T=575$ and 700 °C are shown in Fig. 12. The solubility limit at each temperature is shown by the '+' or 'x' symbols. A gradual change between the two temperatures was confirmed by taking measurements at 25 °C intervals (not shown). The key result of this figure shows that the maximum N_{Zn} obtained during these growths was set by the thermodynamic solubility limit. The very different N_{Zn} versus P_{Zn} dependence observed between growth and diffusion below the solubility limit appears to be related to a pinned Fermi energy at the surface, as discussed in Section 4.

3.2.3.1.3. Implications for diffusion

Given the modern practice of annealing at temperatures above the growth temperature and with a P_{As} which is no higher than the P_{As} used for growth, Fig. 12 suggests that the Zn solubility limit, and presumably that of any other group II atom, is routinely exceeded during annealing. In Sections 4 and 5, minor processing variations will be shown to sharply change the rate at which such a supersaturated Zn concentration drops to its equilibrium value. A rapid reduction of N_{Zn} toward its equilibrium value may occur as either (1) Zn diffuses toward the surface and into the crystal, generally changing the concentration profile of any pn junction or heterojunction present, or (2) Zn forms precipitates of an appropriate condensed phase inside the GaAs. Indeed, Jager et al. [43] have reported the formation of Zn_3As_2 precipitates during annealing of Zn-doped GaAs at $T=900$ °C. However, at $T=700$ °C,

Zn_3As_2 precipitation was looked for but not observed [21]. Several important effects associated with Zn out-diffusion from supersaturated layers will be discussed in detail in Section 5.

3.2.3.2. Native defects

The solubility limit of the native defects can easily be understood for the case of undoped GaAs. The T - X slice of the phase diagram shown in Fig. 2(b) is shown again in Fig. 13(a). Table 1 makes it clear that the maximum equilibrium value of defect concentration, i.e. the solubility limit, is reached when the crystal has equilibrated either with the Ga-rich liquidus or with the As-rich condensed phase. It should also be clear that the equilibrium native defect concentrations will change by many orders of magnitude as the ratio of Ga:As changes very slightly in the crystal. This is what can cause large variations in impurity diffusivity.

Of course, the total concentration of a particular native defect is a sum of its neutral and charged defect concentrations. As already shown, the equilibrium concentration of neutrals are defined by a choice of T and P_{As} , but the equilibrium concentration of the charged defects depends upon T , P_{As} , and P_{dop} . Thus, the total concentration of a native defect in GaAs is expected to vary with the dopant concentration. The solubility limit for each native defect in an extrinsic crystal will thus depend upon the specific condensed phase or phases which limit the dopant concentration, and this is determined by the applicable Ga–As–dopant ternary phase diagram. A schematic isothermal slice of a generic ternary phase diagram is shown in Fig. 13(b). Along the Ga–As axis, i.e. without dopant present, this figure shows the same range of equilibrium stoichiometry illustrated in Fig. 13(a) at a given T . With a dopant present, the average composition of the system moves away from the Ga–As line, and the range of solid composition may vary as illustrated by the small region shown encompassing the pure GaAs composition. Fig. 13(b) shows this as an asymmetric solid composition, but the actual shape of this region will be defined by the energies of formation of specific condensed phases of the dopant with As or Ga. To the extent that the charged native defect concentrations increase with dopant concentration, the crystal can become more As-rich or Ga-rich than in the undoped crystal. Thus, the dopant concentration affects an impurity diffusivity by increasing the native defect solubility limits. The native defect solubility limits in GaAs depend upon the dopant which is present, i.e. upon the formation of the same condensed phases which limit the impurity solubility.

3.2.4. Variation of point defect solubility with electric potential

3.2.4.1. Native point defects

In this section, we shall discuss a pn junction with its p-type layer next to the GaAs surface. We shall assume that the bulk of the p-region is at equilibrium with the vapor, i.e. in an equilibrium state defined by T , P_{As} , and $P_{acceptor}$. The energy band diagram for the pn junction is shown in Fig. 14 which also shows the crystal surface which has the Fermi energy pinned near the center of the bandgap. In this section, we shall show that the equilibrium *native defect concentrations* in both the p-type and n-type regions have well defined equilibrium values with respect to the three external variables, plus the built-in voltage of the pn junction. In Section 3.2.5, we shall explicitly discuss effects associated with the pinned Fermi energy.

We shall begin by considering the example of a singly charged Ga interstitial, I_{Ga}^+ . If equilibrium can be approximated between Ga in the vapor, Ga_{vap} , and I_{Ga}^+ plus its associated electron in the p-type layer, then we may equate the electrochemical potentials,



or, equivalently,

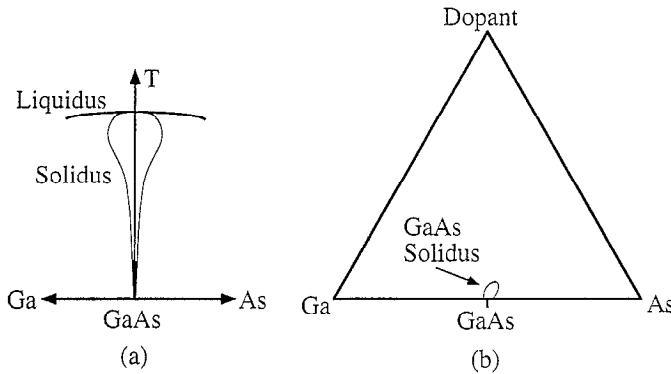


Fig. 13. (a) Schematic illustration of the solidus illustrating the variation in the stoichiometry of the binary 'line compound' GaAs. (b) Schematic slice of the ternary phase diagram for a dopant plus GaAs. The solidus line is exaggerated to illustrate how the range of stoichiometry may change. Such changes correspond to changes in the native defect concentrations and they are expected to strongly affect any impurity diffusion mediated by native defects.

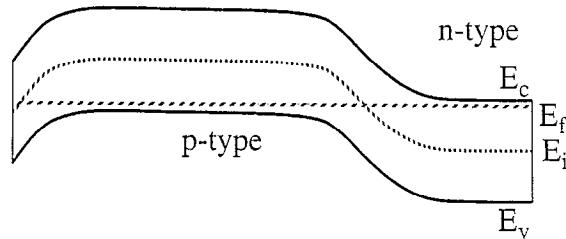


Fig. 14. A p-type epitaxial layer is shown forming a pn junction with an n-type substrate. At the surface, the Fermi energy is pinned near the center of the bandgap. The equilibrium concentration of charged point defects is determined by their electrochemical potential, and thus they are different in the different regions of this structure.

$$\mu_{\text{Ga,vap}}^{\text{Std}} + kT \ln \frac{P_{\text{Ga}}}{P_{\text{Ga}}^{\text{Std}}} = \mu_{\text{I}_{\text{Ga}}^+}^{\text{Std}} + kT \ln \frac{[\text{I}_{\text{Ga}}^+]}{[\text{I}_{\text{Ga}}^+]^{\text{Std}}} + q\phi_p + \mu_e^{\text{Std}} + kT \ln \frac{n}{n_i} - q\phi_p$$

where the bulk of the p-type layer is taken to be at an arbitrary potential, ϕ_p . Thus one obtains

$$K_{106} = \frac{[\text{I}_{\text{Ga}}^+]n}{P_{\text{Ga}}} \quad (106)$$

With the help of Eq. (13) this leads to

$$[\text{I}_{\text{Ga}}^+] \propto p P_{\text{As}_4}^{-1/4} \quad (107)$$

in agreement with Table 5. The potential, ϕ_p , of the layer does not appear in Eq. (107) because a neutral atom in the vapor equilibrates with *both* an ion *and* its associated electron, and the electrostatic work on the ion and electron entering the solid sums to zero. Indeed, this is in agreement with the phase rule which implies that the bulk equilibrium concentration of any charged defect must be unaffected by a potential drop at the surface (although the kinetics of mass transfer may be affected).

The situation is different across the pn junction because a fourth degree of freedom is required to define the equilibrium state of the n-type layer. If we assume Ga interstitials diffuse much more rapidly than the substitutional atoms which define the pn junction, then it is meaningful to approximate the substitutional atoms as frozen in place and to discuss the equilibration of interstitials across the junction. This is analogous to the equilibration of rapidly moving electrons and holes across the pn junction, and it can be described by equating

$$\tilde{\mu}_{\text{I}_{\text{Ga}}^+}(\text{p-side}) = \tilde{\mu}_{\text{I}_{\text{Ga}}^+}(\text{n-side})$$

or

$$\mu_{I_{Ga}^+}^{Std} + kT \ln \frac{[I_{Ga}^+]}{[I_{Ga}^+]_p^{Std}} + q\phi_p = \mu_{I_{Ga}^+}^{Std} + kT \ln \frac{[I_{Ga}^+]_n}{[I_{Ga}^+]_n^{Std}} + q\phi_n$$

where ϕ_n is the potential of the n-type GaAs. Setting $\Delta\phi = \phi_n - \phi_p$, one obtains

$$K_{108} = \frac{[I_{Ga}^+]_n}{[I_{Ga}^+]_p} \exp\left(\frac{q\Delta\phi}{kT}\right) \quad (108)$$

When $\Delta\phi = 0$, it is clear that we must have $K_{108} = 1$. Thus, it is possible to define the equilibrium concentration of I_{Ga}^+ in the buried n-type GaAs, in terms of the ambient conditions (the three degrees of freedom defining the equilibrium concentration of I_{Ga}^+ in the p-type layer) plus the potential drop across the junction (the fourth degree of freedom set by the donor concentration). Equivalently, one may relate the equilibrium defect concentrations through the majority carrier concentrations, n_n and p_p , at the n- and p-type edges of the junction depletion region, respectively, i.e.

$$[I_{Ga}^+]_n = [I_{Ga}^+]_p \exp\left\{-\frac{q\Delta\phi}{kT}\right\} = [I_{Ga}^+]_p \frac{n_i}{n_n p_p} \quad (109)$$

The implications of Eq. (109) are not widely appreciated at present. This expression shows that the *equilibrium concentration* of I_{Ga}^+ (or any other charged native defect) is expected to vary by orders of magnitude across a pn junction. Thus, a relatively *high concentration* of a charged native defect in a substrate will not necessarily diffuse into an epilayer which contains a lower concentration of the same defect. Indeed, the electric field at a junction is so important that it may cause charged native defects and impurities to move against a concentration gradient.

If equilibrium is assumed to exist between the vapor and the acceptors in the p-type layer, and between Ga interstitials across the pn junction, then Eqs. (109) and (107) combine to give

$$[I_{Ga}^+]_n \propto \frac{P_{As_4}^{-1/4}}{n_n} \quad (110)$$

Of course, this is exactly what any thermodynamic model will predict for an n-type region of GaAs located anywhere in the crystal. Some of the results discussed in Sections 4 and 5 indicate that interstitials can rapidly equilibrate across pn junctions, and this will be seen to have important consequences for diffusion in practical structures.

A similar analysis may be applied to a negatively charged native defect. If we consider a negative, singly charged Ga vacancy, the equilibrium concentration in the buried n-type layer will be larger than that in the p-type region, as seen by

$$[V_{Ga}^-]_n = [V_{Ga}^-]_p \exp\left\{+\frac{q\Delta\phi}{kT}\right\} = [V_{Ga}^-]_p \frac{n_n p_p}{n_i n_i} \propto n_n P_{As_4}^{1/4} \quad (111)$$

This is the same behavior predicted for an n-type layer not covered by p-type GaAs, i.e. as described by the second column of Table 6. However, without direct contact with the vapor, the rate at which a native defect approaches equilibrium in a buried layer is expected to depend upon the magnitude of the defect flux which can be supported through the adjacent layers. Since the flux of a given defect depends upon the concentration of that defect, which in turn depends upon the background doping level, fluxes may vary significantly between nominally similar structures. Indeed, Section 5 will provide examples of buried layers which do not approach equilibrium in any practical annealing time, and one well known example will be shown to be that of the ‘emitter push’ effect.

3.2.4.2. Impurity point defects

Clearly, the donor atoms in the n-type GaAs layer of Fig. 14 can never equilibrate with a vapor which does not contain those atoms. However, it will be useful to consider how the acceptor interstitial concentration approaches equilibrium across a pn junction. We anticipate that this case will apply to Zn interstitials, and thus we shall relate the electrochemical potential for interstitial Zn on each side of the pn junction to the electrochemical potential for Zn in the vapor. In contrast to the common assumption that I_{Zn} equilibrates with substitutional Zn, we explicitly assume that I_{Zn} equilibrates primarily with P_{Zn} . Evidence that this assumption often applies will be discussed in Section 5. Thus, for the pn junction considered in Fig. 14,

$$\tilde{\mu}_{Zn,vap} = \tilde{\mu}_{I_{Zn}^+} + \tilde{\mu}_{e^-}|_{p\text{-type}} = \tilde{\mu}_{I_{Zn}^+} + \tilde{\mu}_{e^-}|_{n\text{-type}}$$

gives

$$\begin{aligned} \mu_{Zn,vap}^{std} + kT \ln \frac{P_{Zn}}{P_{Zn}^{std}} &= \mu_{I_{Zn}^+}^{std} + kT \ln \frac{[I_{Zn}^+]_p}{[I_{Zn}^+]_{std}} + \mu_{e^-}^{std} + kT \ln \frac{n_p}{n_i} \\ &= \mu_{I_{Zn}^+}^{std} + kT \ln \frac{[I_{Zn}^+]_n}{[I_{Zn}^+]_{std}} + q \Delta \phi + \mu_{e^-}^{std} + kT \ln \frac{n_n}{n_i} - q \Delta \phi \end{aligned}$$

which is equivalent to

$$K_{112} = \frac{[I_{Zn}^+]_p n_p}{P_{Zn}} = \frac{[I_{Zn}^+]_n n_n}{P_{Zn}} \quad (112)$$

Thus, the equilibrium concentration of a singly charged Zn interstitial is significantly lower in n-type GaAs compared with p-type GaAs, and those concentrations may be related to the built-in junction potential,

$$[I_{Zn}^+]_n = [I_{Zn}^+]_p \frac{n_p}{n_n} = [I_{Zn}^+]_p \exp \left\{ -\frac{q \Delta \phi}{kT} \right\} \quad (113)$$

In a similar fashion, one may show that the solubility of substitutional Zn in the buried n-type layer will be enhanced by

$$[Zn_{Ga}]_n = [Zn_{Ga}]_p \exp \left\{ +\frac{q \Delta \phi}{kT} \right\} \quad (114)$$

Of course, this equation applies *only* if the equilibrium Zn concentration in the n-type GaAs remains below the donor concentration. Such a condition may be met in practice at relatively high temperatures and relatively low Zn partial pressures.

3.2.5. Fermi level pinning at the surface

In Section 4, we shall discuss evidence that the Fermi energy at the surface is pinned near the center of the bandgap during growth of either n- or p-type GaAs. Such pinning has profound implications for the concentration of both the dopant and charged native defects and any associated diffusion. The energy band diagram of Fig. 14 illustrates Fermi energy pinning at the surface, and the bending of the bands next to the surface indicates that a potential drop is expected between the surface and the p-type region of the semiconductor. In this section, we shall consider the implications of pinning for the concentration of species *inside* the space charge region adjacent to the surface.

3.2.5.1. Defect equilibrium in bulk versus surface

As shown in Section 3.2.4, an electrostatic potential drop cannot change the equilibrium relations between the neutral species on either side of a charged interface. Even if the Fermi energy is pinned, the same equilibrium occurs between a neutral atom in the vapor and the corresponding neutral combination of ion *plus* its associated electronic carrier(s) in the bulk. Insight may be gained by considering that as the electrostatic energy for one charged species goes up, the electrostatic energy for the other charged species goes down by exactly the same amount as both cross a charged interface. One major implication of this is that the *ion concentration* in the *bulk* can equilibrate with an entirely different ion concentration at the *surface*, and that *both can equilibrate simultaneously with the same partial pressure*.

For simplicity, we shall show that this is true for Zn, and then discuss the more general results for other dopants and for native defects. When Zn in the vapor equilibrates with substitutional Zn in both the p-type bulk and in the space charge region adjacent to the surface, then

$$\tilde{\mu}_{Zn,vap} + \frac{1}{4}\tilde{\mu}_{As_4,vap} = \tilde{\mu}_{Zn_{GaAs}} + \tilde{\mu}_{h^+}|_{bulk} = \tilde{\mu}_{Zn_{GaAs}} + \tilde{\mu}_{h^+}|_{surf}$$

or

$$\begin{aligned} \mu_{Zn,vap}^{Std} + kT \ln \frac{P_{Zn}}{P_{Zn}^{Std}} + \frac{1}{4}\mu_{As_4,vap}^{Std} + kT \ln \left[\frac{P_{As_4}}{P_{As_4}^{Std}} \right]^{1/4} &= \mu_{Zn_{GaAs}}^{Std} + kT \ln \frac{[Zn_{GaAs}]_{bulk}}{[Zn_{GaAs}]_{Std}} \\ + \mu_{h^+}^{Std} + kT \ln \frac{p_{bulk}}{n_i} &= \mu_{Zn_{GaAs}}^{Std} + kT \ln \frac{[Zn_{GaAs}]_{surf}}{[Zn_{GaAs}]_{Std}} + q\Delta\phi + \mu_{h^+}^{Std} + kT \ln \frac{p_{surf}}{n_i} - q\Delta\phi \end{aligned}$$

which is equivalent to

$$K_{115} = \frac{[Zn_{Gal}]_{bulk} p_{bulk}}{P_{Zn}} = \frac{[Zn_{Gal}]_{surf} p_{surf}}{P_{Zn}} \quad (115)$$

If the Fermi energy, and thus the carrier concentration, is pinned at the surface, then

$$[Zn_{Gal}]_{surf} = [Zn_{Gal}]_{bulk} \frac{p_{bulk}}{p_{surf}} \quad (116)$$

shows that the equilibrium Zn concentration very close to the surface will be significantly larger than the equilibrium Zn concentration in the bulk of the epilayer if $p_{bulk} > p_{surf}$.

From Section 3.2.2, it is clear that $N_{dop,bulk} (\equiv [Zn_{Gal}]) \propto P_{dop}^{1/2}$ for any dopant which sets the Fermi energy. Conversely, if the Fermi energy is not affected by the dopant, then the discussion associated with Eq. (92) shows that $N_{dop} \propto P_{dop}$. Although the same three degrees of freedom, T , P_{As} , and P_{dop} apply, the solubility relation is different at the surface than in the bulk. It is simple to show that the equilibrium substitutional concentration of a dopant at the surface, $N_{dop,surf}$, is related to its equilibrium substitutional concentration in the bulk, $N_{dop,bulk}$, by

$$N_{dop,surf} = N_{dop,bulk} \exp \left\{ \pm \frac{q(\phi_{surf} - \phi_{bulk})}{kT} \right\} \quad (117)$$

where the '+' sign applies for a p-type bulk, and the '-' sign applies for an n-type bulk, and $q\phi \equiv (E_f - E_i)$. When the Fermi energy is pinned near the center of the bandgap, this expression shows that the equilibrium concentration of *either donors or acceptors is enhanced* at the surface relative to the bulk.

Similarly, the concentrations of the charged native defects at the surface are different than the corresponding bulk concentrations. For example, consider positively charged Ga interstitials, I_{Ga}^+ ,

$$I_{Ga,surf}^+ \leftrightarrow I + Ga_{bulk} \Rightarrow \tilde{\mu}_{I_{Ga,surf}^+} = \tilde{\mu}_{I_{Ga,bulk}^+} \quad (118)$$

which leads immediately to

$$[I_{Ga}^+]_{surf} = [I_{Ga}^+]_{bulk} \exp\left\{ \frac{q(\phi_{bulk} - \phi_{surf})}{kT} \right\} = [I_{Ga}^+]_{bulk} \frac{p_{surf}}{p_{bulk}} \quad (119)$$

Similarly, for negatively charged Ga vacancies,

$$[V_{Ga}^-]_{surf} = [V_{Ga}^-]_{bulk} \exp\left\{ \frac{q(\phi_{surf} - \phi_{bulk})}{kT} \right\} = [V_{Ga}^-]_{bulk} \frac{p_{bulk}}{p_{surf}} \quad (120)$$

One of the more important effects of Fermi energy pinning occurs during the growth process. For example, the dopant and defect concentrations in the bulk of an epitaxial layer appear to have values which correspond to equilibrium between vapor and a pinned surface, i.e. these concentrations appear to have been frozen into place during growth. Sections 4 and 5 will present evidence that this is common behavior and that it often has large effects on diffusion.

3.2.5.2. Design of an 'equilibrium' diffusion experiment: Fermi-level model

In Section 5, we shall provide evidence that a class of experiments exist in which the native defect concentrations actually approach their equilibrium values in an epitaxial layer. In such experiments, the diffusion of an isoelectronic marker, such as Al or In, can be interpreted within a simple equilibrium model to determine some of the fundamental properties of the native defects involved in diffusion. Since some of that topic relates to the pinned surface during annealing, we shall develop here the idea of marker diffusion when native defect equilibrium holds.

In a modern open tube system consisting of wafer and vapor, there are two types of experiment which may be performed to bring the native defect concentrations in an epilayer close to their equilibrium values: type (1) in which T , P_{As} , and P_{dop} are set to values which correspond to equilibrium with the carrier concentration previously grown into the GaAs, or type (2) in which a directly controlled T , P_{As} , and $(\phi_{surf} - \phi_{bulk})$ define the equilibrium concentrations of the defects in the bulk, i.e. when $P_{dop} = 0$ is used. Type (1) is appealing because it corresponds to true equilibrium between the solid and vapor. However, it will be shown in Section 5.2 that type (1) experiments often become too complicated in practice to be useful. Experiments similar to type (2) are routinely performed in practice. However, since $N_{dop} = 0$ is the only equilibrium concentration in equilibrium with $P_{dop} = 0$, the relevant question is: how can this possibly approximate a type (1) equilibrium?

The answer is simple if the Fermi energy at the surface is constant throughout the anneal. This is obviously the case if the Fermi energy is pinned at the surface. However, it also occurs even if pinning is not present because only a small amount of dopant diffusion out of the crystal will bring the dopant concentration at the surface to sufficiently small values that the carrier concentration at the surface approaches a constant value, n_i . Thus, in practice a constant potential drop across the surface space charge region is created more or less quickly during annealing without any dopant added to the gas phase. The Ga interstitial and vacancy concentrations next to the surface are well defined by Eqs. (97) and (103), respectively. Presumably, those concentrations will be attained quickly because of the close proximity of the surface to the vapor. The bulk defect concentrations are well-defined, relative to the surface, via Eqs. (119) and (120) and similar equations, and the equilibrium native defect concentrations in the epilayer are exactly the same as expected in a type (1) experiment.

If the native defect concentrations in the bulk approach those equilibrium values in a time which is short compared with the time of the entire anneal, then the marker diffusion in the epilayer proceeds in the presence of an equilibrium native defect concentration. Clearly, in a type (2) experiment, the

length of time used for the annealing is critical to its success. A sufficiently long anneal time must be used for the native defects below the space charge layer to approach their equilibrium values early in the anneal. However, if the anneal time is too long, then N_{dop} and the defect concentrations will gradually decrease throughout the epilayer and one cannot interpret the results by a simple equilibrium model. The appropriate time can only be determined by experiment, and Section 5 will discuss experiments in which a native defect equilibrium approximation appears to apply.

We also note that this logic underlies application of the so-called ‘Fermi-level model’ to marker diffusion without any dopant added to the vapor. If an appropriate anneal time is chosen, then the Fermi-level model applies because Eqs. (97) and (103) and similar equations apply. To the best of our knowledge, these underlying assumptions have never been stated explicitly before. This discussion also provides an understanding for some of the failures of the Fermi-level model, particularly when solid–vapor (and native defect) equilibration is not approached or is affected by experimental design. Particularly dramatic examples of this can be observed when GaAs covered by a dielectric cap layer, as illustrated in Section 5.4.

3.3. Alternative descriptions of the native defect concentrations

Shockley and Last [44] provided an early description of how charged native defect concentrations vary in Si with the Fermi energy, E_f . Kroger [7] has broadened this description to include compounds. Both approaches relate changes in the concentration of charged defects to changes in carrier concentration. Kroger made use of the fact that the equilibrium concentration of neutrals was independent of the Fermi energy, and considered reactions such as Eq. (32)



Making the reasonable assumption that the concentration of multiply charged vacancies will become significant for a sufficiently high electron concentration, one may also consider



and perhaps even more highly charged vacancies. It is also plausible to assume that similar reactions may occur with holes, and thus write



Kroger used a useful graphical format to show the relative defect concentrations as a function of electron concentration, as shown for V_{Ga} in Fig. 15 for a constant T and P_{As} . Such plots implicitly take the carrier concentration as the third degree of freedom, and they assume ideality, i.e. the Fermi energy is not close enough to either band to cause degeneracy effects.

Results from these, or other simple, electrochemical reactions show that the concentration of any negative (positive) defect is expected to increase (decrease) as the electron concentration increases. Although these calculations cannot show that such reactions occur to any significant extent for a Fermi energy within the bandgap, they do provide a reasonable motivation for believing that native defects may exist in charged states. Unfortunately, no calculation of the absolute defect densities can be made.

Baraff and Schluter attempted to improve on this picture by performing a priori calculations of the formation energy for several native defects. Their results, for the free energy of defect formation versus E_f , yield plots which are qualitatively consistent with Fig. 15. However, their calculation explicitly predicted the Fermi energy at which each electrochemical reaction shifts to the right or to the left, i.e. the Fermi energy at which equal densities of a defect in two different charge states coexist. These calculations appear to provide strong motivation for believing that the native defects exist in

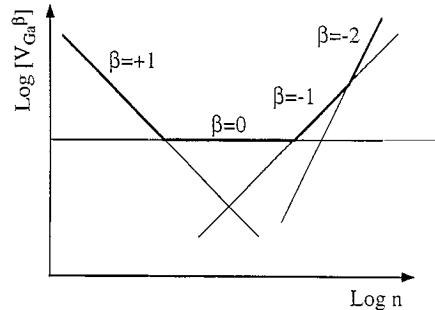


Fig. 15. Log plot of the relative values of the equilibrium concentration of Ga vacancies, $[V_{Ga}^{\beta}]$, in different charge states vs. electron concentration, n , based upon an analysis originally presented by Kroger. This clearly illustrates that the total number of vacancies depends upon n , and that the neutral vacancy concentration is insensitive to carrier concentration. In addition, defects in all charge states exist at some concentration for all electron concentrations. If impurities moved only by vacancies in a crystal with a given electron concentration, then the vacancy species with the highest concentration (shown by the heavy line) is presumed to dominate the impurity diffusion (if the jump rate of vacancies in different charge states are similar). However, the electron concentration corresponding to the crossover points has yet to be successfully predicted, and must be inferred from experiments.

several charge states in nondegenerate GaAs, and that the charged defect densities depend strongly upon the Fermi energy.

However, there are several weaknesses in the Baraff and Schluter calculations. There is no explicit dependence upon any real degree of freedom, i.e. the concentration of vacancies, interstitials, and antisites cannot be compared because these defects have a different dependence upon temperature and pressure. In addition, the lack of a temperature dependence in the model appears to be equivalent to assuming that the entropies of defect formation are zero [45]. Indeed, their results appear to be consistent with nonzero equilibrium defect concentrations at all temperatures, including $T = 0$ K. However, statistical mechanics shows this is the one temperature at which the equilibrium defect concentrations must have a value of zero. Thus, it is unclear what experiment corresponds to these predictions.

We consider both the Kroger and Baraff approaches to provide strong plausibility arguments for variable concentrations of charged native defects in a nondegenerate semiconductor. However, the specific calculations of formation energy presented by Baraff and Schluter are probably not correct, and they should be used with caution. However, their calculations have stimulated other efforts to reconsider this problem. With the more recent Car–Parinello approach to simulating molecular dynamics, it now may be possible to relate the native defect concentrations to a reservoir of atoms, i.e. a reservoir which simulates the external phase which defines the equilibrium state in a real experiment. We simply note that several other efforts [46–53] have been made to calculate the relative concentrations of multiply charged native defects, and we make no attempt to evaluate any here.

3.4. Diffusion models and mechanisms

3.4.1. General transport model

The electrochemical potential should be used to describe equilibrium for a charged species. Given some simple approximations, i.e. a dilute solution approximation, many simple but useful expressions can be developed. The electrochemical potential energy, $\tilde{\mu}$, is generally described in terms of separate chemical and electrical contributions, as discussed in Section 3.1.3,

$$\tilde{\mu}_i = \tilde{\mu}_i + z_i q \phi \quad (124)$$

For a positive, singly charged, species,

$$\tilde{\mu} = \mu^{\text{std}} + kT \ln N + q\phi \quad (125)$$

we may write a generalized force acting on those particles analogous to that written in Section 2.5,

$$-F = \frac{\partial \tilde{\mu}}{\partial x} = \frac{kT}{N} \frac{\partial N}{\partial x} + q \frac{\partial \phi}{\partial x} \quad (126)$$

and relate the net flux to this generalized force. If the generalized force is not too large, then one may assume that a linear relation exists between the flux, J , and the force through an average velocity, v , and generalized mobility, M . These may be related to the more commonly used diffusivity, D , mobility, μ , and electric field, \mathcal{E} ,

$$J = Nv = NMF = -MkT \frac{\partial N}{\partial x} - NMq \frac{\partial \phi}{\partial x} = -D \frac{\partial N}{\partial x} + N\mu\mathcal{E} \quad (127)$$

In Eq. (127) the flux has been related to the familiar diffusive and drift components. The divergence of the impurity flux can be related to the time rate of change of its concentration,

$$\frac{\partial N}{\partial t} = -\frac{\partial}{\partial x} \left\{ -D \frac{\partial N}{\partial x} + N\mu\mathcal{E} \right\} \quad (128)$$

or

$$\frac{\partial N}{\partial t} = D \frac{\partial^2 N}{\partial x^2} + \frac{\partial D}{\partial x} \frac{\partial N}{\partial x} - \frac{\partial N}{\partial x} \mu\mathcal{E} - N \frac{\partial \mu}{\partial x} \mathcal{E} - N\mu \frac{\partial \mathcal{E}}{\partial x} \quad (129)$$

If an impurity moves solely by a pure interstitial or direct exchange mechanism, then D and μ will be independent of native defect concentrations, but may be sensitive to electric fields. If impurity diffusion is mediated by a native defect, then the impurity D and μ will be affected by any position dependence in the concentration of that native defect. If either species is charged, then an electric field can affect the impurity diffusion. Simulating the diffusion process often involves coupled equations analogous to Eq. (129) for substitutional and interstitial impurities and all relevant native defects. Several attempts have been made to model these interactions to produce working engineering models. A detailed discussion of the models used for simulation is beyond the scope of this paper, and we note that a good and readable summary has been given recently by Haddara et al. [54].

3.4.2. Diffusion mechanisms with charged point defects

Shaw [4] appears to have been the first to relate the equilibrium native defect concentrations, carrier concentration, and arsenic partial pressure to the diffusivity in GaAs. In more recent times, Tan and Gosele [8] attempted to introduce a similar model, called the ‘Fermi level model’, which was based upon their work with Si. Cohen [6] pointed out that the Fermi level model for an elemental solid could not carry over directly to GaAs because of the extra degree of freedom, P_{As} . Since the data analyzed by Tan and Gosele were taken at several different values of P_{As} , it was pointed out that their proposed D_{III} versus $1/T$ relationship should not apply. However, Tan and Gosele later re-interpreted their original equation describing D_{III} versus $1/T$ as valid at $P_{\text{As}_4} = 1$ atm. Given the more careful experimental work done in recent years which is discussed in Section 5, this conclusion appears to be in error.

We note that several physical effects may complicate any description of the mass transport in diffusion experiments. For example, (1) the electric field associated with Fermi energy pinning at the

surface (or buried junction interface) may speed up or slow down the fluxes of ionized point defects, (2) substrates come from GaAs ingots grown and annealed under conditions quite different than those used for the epitaxial growth of thin films, and thus some substrates may act as point defect reservoirs during annealing, and (3) the temperature and ambient partial pressures used for epitaxial growth (and thus the growth technique) determine the initial grown-in defect concentrations and this may affect the diffusion rates as the system approaches equilibrium during annealing. These kinetic complications will be largely ignored in this paper since little is known about them at present.

3.4.3. Near-equilibrium diffusion with native defect isoconcentration

We shall consider a solid-vapor system consisting of GaAs, a dopant which sets the Fermi energy, and an isoelectronic impurity (Al, In, etc.) which is used as a diffusion marker on the group III sublattice. If an epilayer containing these species had been grown from the vapor, then typically the four degrees of freedom chosen would have been T , P_{As} , P_{dop} , and P_{marker} . We shall assume that the dopant has a uniform concentration in the epilayer, and that the concentration of the isoelectronic marker is nonuniform, perhaps spike shaped. We must also assume that the isoelectronic impurity exists at a sufficiently low concentration that the solid can still be considered to be GaAs, i.e. the equilibrium native defect concentrations are unaffected by the presence of the marker. Although this system has four degrees of freedom during annealing, only three (T , P_{As} , and P_{dop}) need to be controlled to set the equilibrium native defect concentrations. The fourth degree of freedom is a composition variable: the final density of the marker atoms after equilibrating throughout the GaAs. During the course of an actual experiment, the marker atoms remain out of equilibrium, but the rate at which they approach equilibrium provides information about the native defect concentration controlling that marker diffusion. To interpret changes in marker diffusivity with partial pressure or carrier concentration, the expressions for the vacancy and interstitialcy in Section 2.5.2 are particularly useful.

The semi-empirical expressions of Section 2.5.2 buried all of the difficult physics of diffusion into different standard (or intrinsic) diffusivities, each of which corresponded to impurity diffusion via a different kinetic pathway at some standard conditions. For example, if marker atoms move solely by a vacancy mechanism, and if they jump at a comparable rate via neutral and singly charged vacancies, then it is reasonable to separate the measured rate of marker diffusion into a rate of diffusion via uncharged vacancies plus a rate via charged vacancies. If these vacancy concentrations equilibrate sufficiently rapidly with the external ambient during annealing, then it becomes reasonable to relate changes in the measured marker diffusivity to the predicted changes for diffusion via each of the charged vacancies. Thus, if we choose to discuss the measured diffusion, D_m , solely in terms of neutral and negatively charged vacancies, then we may scale the value of the standard diffusivities up or down with the expected changes in vacancy concentration similar to the way Eq. (61) was developed, i.e.

$$D_m(T, n, P_{\text{As}_4}) = D_{V_{\text{Ga}}^0}(T, n_i, P_{\text{As}_4} = 1) \frac{[V_{\text{Ga}}^0]_{i, P_{\text{As}_4}}}{[V_{\text{Ga}}^0]_{i, P_{\text{As}_4} = 1}} + D_{V_{\text{Ga}}^-}(T, n_i, P_{\text{As}_4} = 1) \frac{[V_{\text{Ga}}^-]_{i, P_{\text{As}_4}}}{[V_{\text{Ga}}^-]_{i, P_{\text{As}_4} = 1}} \frac{[V_{\text{Ga}}^-]_{e, P_{\text{As}_4}}}{[V_{\text{Ga}}^-]_{i, P_{\text{As}_4}}} \quad (130)$$

where 'e' refers to the equilibrium concentrations in the extrinsic crystal, and 'i' refers to equilibrium values in the intrinsic crystal, and the standard state has been chosen to be the convenient values $P_{\text{As}_4} = 1$ atm and $n = n_i$. With the aid of Eqs. (32) and (62), and analogous expressions, D may be generalized to include Ga vacancies in any arbitrary charge state, Q , to give

$$D_m(T, n, P_{\text{As}_4}) = \sum_{Q = \text{All charge states}} D_{V_{\text{Ga}}^Q}(T, n_i, P_{\text{As}_4} = 1) P_{\text{As}_4}^{1/4} \left(\frac{n}{n_i} \right)^{-Q} \quad (131)$$

If negative and positive vacancies exist in sufficient concentration, and have sufficient diffusivity, then Eq. (131) predicts that the measured diffusivity may increase at large values of n (via negatively charged Ga vacancies) and also at large values of p (via positively charged vacancies). We also note that Eq. (131) applies to the diffusion of either neutral markers or charged ions on the group III sublattice. This formalism does not distinguish between other kinetic details, i.e. between a mechanism where vacancies tend to pair up with charged impurities and a mechanism where vacancies move without any significant attractive interaction with impurities. Thus, we shall consider the simplest interpretation of this equation to apply, that of independent unpaired vacancies moving in the crystal, until clear evidence to the contrary is obtained.

A similar relationship can be developed for diffusion via the interstitialcy mechanism. For any charge state, Q , associated with Ga interstitials, one obtains

$$D_m(T, n, P_{\text{As}_4}) = \sum_{Q=\text{All charge states}} D_{iQ}(T, n_i, P_{\text{As}_4} = 1) P_{\text{As}_4}^{-1/4} \left(\frac{p}{n_i} \right)^Q \quad (132)$$

Analogous to the discussion about vacancy controlled diffusion, Ga interstitials may, in principle, be associated with an increase in D_m in either n- or p-type crystals. The equation does not distinguish between I-S defect pair diffusion and non-pair diffusion. We shall consider the simplest interpretation to hold, that of independent defects, until evidence demonstrates that pairs control diffusion. Like its vacancy counterpart, Eq. (132) applies only when the Ga interstitial concentration is close to its equilibrium value.

Evidence will be presented in Section 5 that a native defect isoconcentration approximation is reasonable during the annealing of both n-type and p-type GaAs provided that common group II dopants are *not* present. When group II elements are present in typical experimental designs, it will be shown that the crystal is usually far from equilibrium and the expressions developed in this section may not apply. Some limited interpretations of the results obtained from nonequilibrium experiments can be made and these will be discussed in Section 5.

4. Effects caused by a pinned Fermi energy at a growing surface

Evidence will be presented in this section that several experimental results are consistent with a Fermi energy which is pinned at the GaAs surface at elevated temperatures. Such a description is inherently a thermodynamic one, and thus no mechanistic information about gas-surface-bulk reactions or surface reconstruction can be obtained. However, this thermodynamic interpretation appears to be useful for understanding how nonequilibrium concentrations of both dopants and native defects are grown into GaAs epitaxial layers, and how those concentrations affect impurity diffusion.

4.1. Impurity solubility during OMVPE growth of GaAs

4.1.1. Evidence for the same pinning during n- and p-type growth

Perhaps the simplest way to appreciate the effects of Fermi energy pinning is to assume that no pinning occurs at the surface. Depending upon the dopant supplied to the vapor, the surface of the GaAs should become either p- or n-type during growth. The solubility of an ionized species would be that summarized in Table 2, i.e. after growth one expects to measure $N_{\text{dop}} \propto P_{\text{dop}}^{1/2}$. However, as seen in Fig. 12 for Zn, the Zn dopant concentration after growth varies approximately linearly with P_{Zn} and is also far from the equilibrium concentration of Zn determined at the same T , P_{As} , and P_{Zn} . This linear dependence upon dopant partial pressure is commonly seen in VPE. One might be tempted to conclude

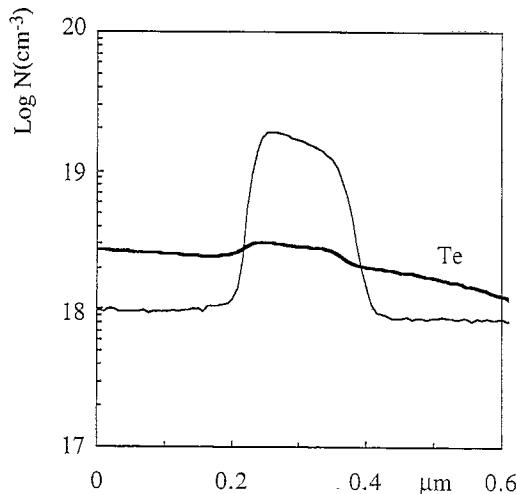


Fig. 16. N_{Te} and N_C vs. x , measured by SIMS, for an npn structure grown using a constant P_{Te} for the entire growth time. The lack of a significant change in N_{Te} in the p-type base layer is strong evidence that the Fermi energy was unchanged, i.e. pinned, at the surface during growth. Adapted from Ref. [55].

that dopant incorporation during growth is a highly nonequilibrium process. However, if a simple equilibrium model explains a wide range of results, then it is often preferred because it is fundamental and simple. Such a model might assume that a region of the crystal very close to the surface approaches equilibrium with the vapor and has a pinned Fermi energy. If this applies, then the dopant concentration near the surface will be described by an expression in Table 3, i.e. $N_{\text{dop}} \propto P_{\text{dop}}$. If this also occurs during growth, then the growth process will trap the dopant at this concentration in the epilayer, and $N_{\text{dop}} \propto P_{\text{dop}}$ will be measured after the growth is finished instead of the bulk equilibrium solubility relation expected from Table 2, i.e. $N_{\text{dop}} \propto P_{\text{dop}}^{1/2}$.

To determine whether the Fermi energy was pinned at the surface, Cohen et al. [55] measured the variations in dopant solubility for an npn structure grown by OMVPE. Te was used both as (1) a dopant for the n-type region, and (2) a *solubility marker* in the p-type region. From Table 3, the solubility of a group VI ion is described by

$$N_{\text{VI}}^+ \propto P_{\text{VI}} P_{\text{As}_4}^{-1/4} p \quad (133)$$

when that impurity ion does not set the Fermi energy. If the Fermi energy were not pinned at the surface, large variations in N_{Te} should have been observed because the Te source was kept at a constant P_{Te} during growth of the *entire* npn structure. In other words, the buried p-type layer was grown by making only a *single change*: turning the C source on, and then off. The resulting Te and C profiles, taken by SIMS, are shown in Fig. 16, and electrochemical C-V profiling was used to confirm that the corresponding electrical npn structure existed.

At the growth temperature, the hole concentration is expected to vary across the pn junction from $p \approx 2 \times 10^{19} \text{ cm}^{-3}$ in the p-type region to $p = n_i^2/n \approx 2 \times 10^{14} \text{ cm}^{-3}$ in the n-type region, i.e. a variation of $\approx 10^5$ for the carrier concentration. For an unpinned surface, Eq. (133) shows that a very large change in N_{Te} was expected during growth of the C-doped layer. However, the lack of change in N_{Te} in the npn structure is strong evidence that the carrier concentration was essentially constant at the growing surface. We can relate the small change in the incorporated value of N_{Te} to p_{surf} by using Eq. (133) i.e.

$$\frac{p_{\text{surf}}(\text{during growth of p-side})}{p_{\text{surf}}(\text{during growth of n-side})} = \frac{N_{\text{VI}}^+(\text{p-side of jctn})}{N_{\text{VI}}^+(\text{n-side of jctn})} \quad (134)$$

This expression shows that the carrier concentration at the surface varied by a factor less than 2 during growth, and therefore

$$\Delta E_{f,surf} = kT \ln \left\{ \frac{p_{p\text{-side of jctn}}}{p_{n\text{-side of jctn}}} \right\}_{\text{at the growing surface}} < 20 \text{ meV} \quad (135)$$

A similar result was obtained when Zn was used instead of C. The simplest explanation for these results is that the Fermi energy was pinned at essentially the same energy during growth of both the p- and n-type GaAs. A successful nonequilibrium model of dopant incorporation at an unpinned surface would require the use of questionable ad hoc assumptions; for example, an assumption that dopant atoms incorporate as neutral species at the growth temperature would predict the same result. Since the thermodynamic interpretation is simpler, and since it also fits with several other results to be discussed, we consider it to be the most reasonable model to use.

In brief, impurity atoms appear to incorporate into the growing surface at concentrations close to the values expected for solid–vapor equilibrium at a pinned surface. As growth continues, the grown-in impurity atom concentration at the surface becomes the grown-in impurity atom concentration throughout the entire epilayer. Thus, although the impurity concentration at the surface may be at equilibrium with the vapor, the same impurity concentration in the epilayer bulk must be far from equilibrium with the vapor.

Fermi level pinning at the growing surface has profound implications for the grown-in concentration of *all charged point defects*. If the charged native defect concentrations take on their equilibrium values at the growing surface, then those frozen-in concentrations will become nonequilibrium concentrations once they are in the neutral bulk (unless the defects can diffuse so rapidly that they equilibrate during the growth process). The discussion in Section 3.2.5 allows one to calculate how far from equilibrium the defect concentrations can be. In the remainder of Section 4, we shall discuss evidence for both pinning and for nonequilibrium impurity and native defect concentrations which have been grown into GaAs. Section 5 will discuss subsequent processing at elevated temperatures, and provide some examples of the rapid changes which may occur as highly nonequilibrium defect concentrations relax toward their equilibrium values.

4.1.2. Calculation of the pinned Fermi energy

Reichert et al. [21] originally compared the N_{Zn} – P_{Zn} relationship found for OMVPE growth on (100)-oriented GaAs surfaces with the equilibrium N_{Zn} – P_{Zn} relationship, and the results are shown in Fig. 17. The line passing through the growth results has a slope of unity, and it adequately describes the data until the solubility limit is reached. The equilibrium line was originally obtained by diffusing Zn into undoped GaAs from the vapor, and taking N_{Zn} just below the surface space charge region (as measured by SIMS) as the equilibrium concentration. Chen et al. [22] later confirmed that this was the equilibrium curve by approaching it from above, i.e. annealing very heavily Zn-doped GaAs under the same conditions used by Reichert et al., and obtaining the same N_{Zn} , as already shown in Fig. 11. The line shown passing through the near-equilibrium results has a slope of 1/2, illustrating the slope expected for solid–vapor equilibrium.

Reichert et al. extended the lines shown in Fig. 17 until they intersected at low concentrations, and used that intersection to calculate the pinned Fermi energy. This intersection gives the same value of N_{Zn} expected for either growth or equilibrium experiments, i.e. the same value of p_{surf} as for p_{bulk} . Since the intersection occurs at approximately $N_{Zn} \approx 3 \times 10^{17} \text{ cm}^{-3}$, then $p_{bulk} = p_{surf} \approx 3 \times 10^{17} \text{ cm}^{-3}$ or $p_{surf} \approx 10n_i$. This allows an immediate calculation of the pinned Fermi energy, relative to the intrinsic Fermi energy, E_i , at $T = 700^\circ\text{C}$, via

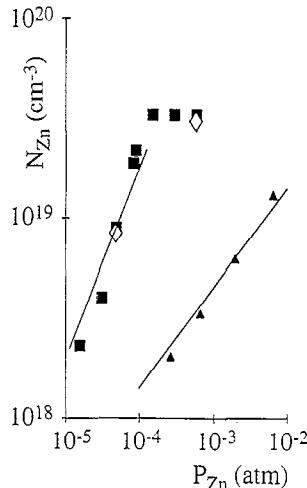


Fig. 17. Incorporated N_{Zn} vs. P_{Zn} measured after OMVPE growth compared with the equilibrium results, all at the same T and P_{As} . The equilibrium solubility data (triangles) were measured by both SIMS and C-V profiling. Nonequilibrium (growth) results were measured by either C-V (squares) or SIMS (diamonds). Adapted from Ref. [21].

$$(E_i - E_f)_{\text{surf}} = kT \ln \left\{ \frac{p_{\text{surf}}}{n_i} \right\} \approx 200 \text{ meV} \quad (136)$$

Recent results reported by Chen et al. [42] confirm the Fermi level pinning, but estimate a slightly smaller $(E_i - E_f)_{\text{surf}} \approx 100 \text{ meV}$. Given the uncertainty involved in extrapolating the curves to their intersection, the best estimate of the pinning energy is probably the average of these two estimates, or $(E_i - E_f)_{\text{surf}} = 150 \text{ meV}$. Improved accuracy using this experimental approach is unlikely to be obtained because diffusion is so slow that the equilibrium curve must be extrapolated from the higher partial pressures. The results summarized by Fig. 16 make it clear that the Fermi level is pinned at virtually the same energy during growth of both n- and p-type GaAs. Thus, we conclude that $(E_i - E_f)_{\text{surf}} \approx 150 \text{ meV}$ during the growth of either n- or p-type GaAs on a vicinal (100) surface.

Another way to view the implications of equilibrium between vapor and a pinned solid surface is to consider that the reaction



implies that the product, $p \cdot N_{\text{Zn}} = \text{constant}$ for a given P_{Zn} and T . In the bulk, one expects $p = N_{\text{Zn}}$. If the Fermi energy is pinned near the center of the bandgap, the relatively low hole concentration near the surface requires that the *equilibrium* N_{Zn} in the space charge region next to the surface be higher than the *equilibrium* N_{Zn} in the neutral bulk. Indeed, using the results of Section 3.2.4 it is simple to show that the *equilibrium* ion concentration at the *surface* is increased over the *equilibrium* ion concentration in the *bulk* by a factor which depends upon the magnitude of the built-in potential, $\Delta\phi$, between the surface and bulk,

$$[\text{Zn}_{\text{Gal}}]_{\text{surf}} = [\text{Zn}_{\text{Gal}}]_{\text{bulk}} \exp \left\{ \frac{q \Delta\phi}{kT} \right\} \quad (138)$$

During growth, of course, the Zn concentration incorporated at the surface eventually becomes the Zn concentration grown into the epilayer. In all practical cases, this gives a Zn concentration in the epilayer which is supersaturated relative to the T , P_{As} , and P_{Zn} used. We shall encounter several examples where this has significant consequences for diffusion.

4.1.3. Relaxation of metastable dopant concentration during growth

Chen et al. [42] reported that the linear $N_{\text{Zn}} \propto P_{\text{Zn}}$ observed during growth ended when the solubility limit of Zn in GaAs was reached. Superficially, this appears to be a reasonable result, but actually it is rather odd since a solubility limit is generally set by the formation of another condensed phase (which was not formed in these cases at the low P_{Zn} used). Indeed, if Fermi energy pinning occurs near midgap, then the discussion presented in the early paragraphs of this section strongly argue that one can exceed the traditional ‘solubility limit’ during growth. Chen et al. suspected that they could not exceed the solubility limit because D_{Zn} was so large, for the ambient conditions used, that any supersaturated Zn concentration was reduced to the solubility limit by diffusion. They tested this idea by reducing the growth rate, and showed that one could obtain a transition between the $N_{\text{Zn}} \propto P_{\text{Zn}}$ and a $N_{\text{Zn}} \propto P_{\text{Zn}}^{1/2}$ dependence simply by changing the growth rate. Fig. 18 shows that $N_{\text{Zn}} \propto P_{\text{Zn}}$ from the lowest Zn concentrations up to the solubility limit for growth rates of $r_g > 1.6 \mu\text{m h}^{-1}$. However, at $r_g = 1.6 \mu\text{m h}^{-1}$, the solubility approximately follows $N_{\text{Zn}} \propto P_{\text{Zn}}^{1/2}$ at the highest values of N_{Zn} . At an even slower $r_g = 0.5 \mu\text{m h}^{-1}$, the Zn solubility follows $N_{\text{Zn}} \propto P_{\text{Zn}}^{1/2}$ over a wide range, and it is clearly approaching the equilibrium ($r_g = 0$) solubility curve.

There appear to be a number of interesting possibilities for using this knowledge to exceed the ‘solubility limit’ of a dopant. As discussed in Section 3.2.3, the solubility limit is generally imposed by the formation of additional condensed phases. For example, if the Fermi energy is pinned near midgap during growth, then one should be able to use a sufficiently rapid growth rate to obtain a metastable, but supersaturated, dopant concentration in the epilayer before it can diffuse or precipitate out. Some examples of this appear to have been observed: (1) At high growth rates, Okamoto et al. [56] obtained an approximately linear $N_{\text{Zn}} - P_{\text{Zn}}$ solubility line which peaked at a value of N_{Zn} higher than the solubility limit reported by Chen et al. Okamoto et al. also reported that N_{Zn} dropped from this peak at still higher values of P_{Zn} . It appears that they were able to exceed the solubility limit because of the high growth rates used, $r_g > 6 \mu\text{m h}^{-1}$, and that the supersaturated Zn concentration dropped to the solubility limit when D_{Zn} increased at the higher values of P_{Zn} used. (2) Using Te as

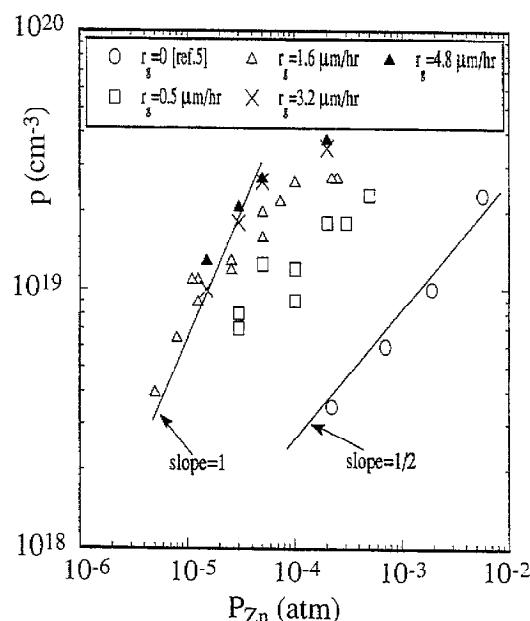


Fig. 18. Incorporated N_{Zn} vs. P_{Zn} obtained from OMVPE growth, using different growth rates, compared to the equilibrium N_{Zn} vs. P_{Zn} line, all at the same T and P_{As} . It is clear that the growth results change from a linear dependence to a parabolic dependence and converge on the equilibrium results at very low growth rates. Adapted from Ref. [42].

a dopant, Li et al. [57] have reported unusually high electron concentrations, $n = 2\text{--}3 \times 10^{19} \text{ cm}^{-3}$, in GaAs. Since these values of n were attainable (with good morphology) only at high growth rates, it is likely that the Te concentration exceeded the solubility limit. However, given the very low Te diffusivity, the n-type GaAs remained in a metastable state during growth. (3) Some of the Zn δ -doping results of Li et al. [58] also appear to correspond to a nonequilibrium supersaturated dopant concentration. This is easy to appreciate by considering the way the spikes were formed: growth of GaAs was stopped, and Zn was added to the vapor and diffused a few monolayers deep in a time of a few seconds. As predicted by Eq. (137), this should produce an extremely large Zn ion concentration in the space charge region. When GaAs growth was resumed without P_{Zn} in the vapor, the supersaturated Zn concentration spikes did not diffuse quickly. Section 5 will discuss how this experimental design can sharply reduce the Zn diffusivity.

4.1.4. Growth techniques other than OMVPE

Two decades ago, Casey and Panish concluded [59] that the Fermi energy must be pinned at the surface during growth of GaAs by LPE at $T \approx 1000^\circ\text{C}$. They reached this conclusion after observing that the resulting dopant solubility for several dopants in GaAs was linearly dependent upon the dopant concentration in the liquid metal phase, but that the dopant Zn exhibited a square root dependence. They also concluded that the different behavior was related to the rapid diffusion of Zn. Their conclusion appears to be justified based upon a theoretical prediction of the same phenomena by Zschauer and Vogel [60].

Some evidence for Fermi energy pinning during growth by molecular beam epitaxy (MBE) has been presented [61], but the energy of pinning appears not to have been determined. However, MBE is usually conducted at relatively low temperatures, using dopants which have very low vapor pressures, and it may be inappropriate to apply a thermodynamic model to dopant incorporation during MBE. However, it should be clear by now that if the ionized dopant concentration does not reflect solid-vapor equilibrium, then the ionized native defect concentrations are not expected to attain their equilibrium values either. At a minimum, this implies that during an anneal at elevated temperature, the initial native defect concentrations in MBE-grown GaAs will be significantly different than those in OMVPE-grown GaAs.

4.2. Pinning and interstitials

Early observations of the so-called ‘emitter push effect’ in GaAs have been discussed elsewhere [62–64]. The emitter push effect has traditionally been associated with the npn (emitter–base–collector) bipolar transistor structure. When a highly doped emitter is formed, by either growth or diffusion, over a relatively lightly doped base region, the diffusion of the base dopant is often observed to increase significantly.

Deppe [65] suggested that the enhanced group II dopant diffusion out of the GaAs base region during the growth of an n-type emitter could be explained if the Fermi energy were pinned near the center of the bandgap at the surface of the growing emitter. In particular, Deppe proposed that Fermi level pinning during growth of the n-type emitter would cause (1) negatively charged Ga vacancies in the emitter to be undersaturated (consistent with Section 3.2.5), and thus (2) Frenkel generation ($e^- + \text{Ga}_{\text{Ga}} \leftrightarrow V_{\text{Ga}}^- + V_{\text{Ga}}$) in the emitter would produce more vacancies and interstitials, and thus (3) the resulting neutral interstitials would diffuse into the p-type base, and enhance the group II element diffusion via a kick-out mechanism. Unfortunately, the computer simulation appeared to contain an error which allows the equilibrium neutral defect concentrations to change. Nevertheless, the basic

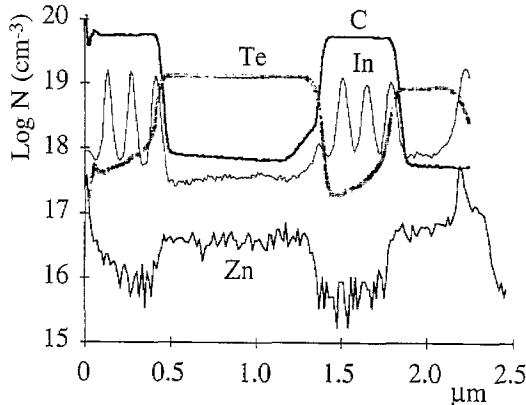


Fig. 19. As-grown pnnp structure, using C and Te, which shows no significant emitter push effect after growth. Trace contamination from Zn appears to have accumulated in the n-type regions. Diffusion of In markers will be discussed in Section 5. Adapted from Ref. [66].

concept of a pinned Fermi energy which induces a perturbation in a thermodynamic state is both insightful and qualitatively consistent with the observations.

Chen et al. [66] reported on the emitter push effect during the OMVPE growth of pnnp structures. The extra p-type layer on top of the npn structure allows significant insight to be gained into the point defects driving diffusion during both growth and annealing. Fig. 19 shows the concentration profiles, taken by SIMS, of a pnnp structure which was grown using C and Te. Narrow spikes of In were also grown in to act as a group III diffusion marker. No Zn was intentionally added during the growth, but Zn entering as a trace impurity appears to be present in the n-type regions (regions of expected high substitutional Zn solubility). No dramatic diffusion occurs, and only a small amount of In diffusion out of the In spikes in the buried layer appears to be present.

A dramatic emitter push effect was observed in similar pnnp structures which were grown containing Zn, in place of C, under otherwise identical conditions. As seen in Fig. 20, the different profile of the buried Zn-doped layer is striking. Chen et al. anticipated the rapid diffusion of the buried Zn-doped layer, and thus (1) the Zn source was not even opened until growth of the first Te-doped layer was nearly complete, and (2) the system was pumped to remove any residual Zn source prior to starting growth of the second Te-doped layer. It was separately confirmed that the Zn diffusion out of the buried Zn-doped layer was a direct result of the growth of the second Te-doped layer. If diffusion

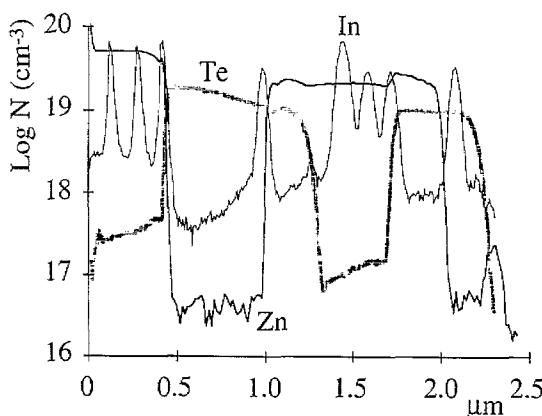


Fig. 20. As-grown pnnp structure, using Zn and Te, which shows a substantial emitter push effect, i.e. the buried p-type layer has widened significantly during growth. The widening is directly associated with the growth of the top n-type layer. The results of annealing will be discussed in Section 5. Adapted from Ref. [66].

had not occurred during growth, then the concentration profile of the buried Zn-doped layer would have been essentially the same as that of the top layer, i.e. $N_{\text{Zn}} = 5 \times 10^{19} \text{ cm}^{-3}$ and $x = 0.4 \mu\text{m}$ wide rather than the observed $N_{\text{Zn}} \approx 2 \times 10^{19} \text{ cm}^{-3}$ and $x \approx 1.0 \mu\text{m}$ wide.

Chen et al. considered it unlikely that Frenkel equilibration occurred rapidly during growth, as Deppe suggested, because Frenkel pairs are generally believed to have a relatively high energy of formation. They proposed that excess Ga interstitials in the n-type layer came directly from the pinned surface as a result of equilibration between Ga atoms in the vapor and Ga interstitial atoms near the surface,

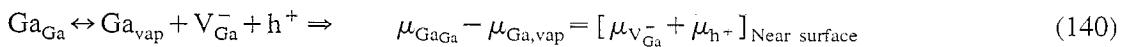


Close to a surface which has the Fermi level pinned near midgap, the concentration of positive Ga interstitials is expected to be much higher than its equilibrium concentration in the underlying n-type GaAs. Thus, Chen et al. proposed that some of the I_{Ga}^+ at relatively high concentrations in the space charge region adjacent to the surface would become temporarily trapped in the n-type region during the growth process. With the expected Ga interstitial concentration in the n-type layer above its equilibrium value, it was proposed that the interstitials left the n-type layer by diffusing either back to the surface or deeper into the crystal, i.e. into the buried p-type layer. The excess I_{Ga} entering the buried p-layer was proposed to enhance the Zn diffusion via a kick-out mechanism.

It is worth noting that both of the models proposed by Deppe and by Chen et al. use a logical and consistent reaction pathway to explain the growth results. Since the results discussed in Section 5 show that the Frenkel reaction does not proceed rapidly during annealing, we consider it likely that the model of Chen et al. provides a better explanation of the emitter push effect. Distinguishing between diffusion mechanisms is generally not simple, and some useful strategies for the design of experiments which help distinguish between the point defects driving diffusion will be discussed further in Section 5.

4.3. Pinning and vacancies

The effect of Fermi energy pinning near the center of the bandgap upon the concentration of negatively charged vacancies can be appreciated by considering the reaction between Ga atoms in the solid and vapor via singly charged Ga vacancies,



The left hand side of the equation for the chemical potentials is independent of Fermi energy pinning, but the right hand side is strongly dependent upon the pinning. Since Fermi energy pinning near midgap forces the hole concentration to be relatively high at the surface during growth of an n-type region, Eq. (140) implies that the equilibrium concentration of negatively charged Ga vacancies near the surface is reduced relative to its equilibrium value in the n-type bulk. For a sufficiently rapid growth rate, one expects that a nonequilibrium low concentration of negatively charged Ga vacancies will be grown into the n-type GaAs, as suggested by Deppe. The vacancies will presumably relax toward their equilibrium concentration through either (1) the Frenkel reaction, or (2) vacancy diffusion from the surface and substrate. At present, it is unclear how the vacancy concentration relaxes back to its equilibrium state and we shall not pursue it further here.

5. Effects of annealing on diffusion and point defects

Some of the important conclusions that we shall reach in Section 5 include the following. (1) Equilibrium native defect concentrations can be quickly approached in some well-designed and well-conducted annealing experiments. (2) Diffusion studies conducted using equilibrium native defect concentrations indicate that the Ga vacancy controls diffusion on the group III sublattice in both n- and p-type GaAs. (3) The charge on the Ga vacancy is -1 over a wide range of n- and p-type doping. (4) Rapidly moving interstitials are usually the first defects to approach their equilibrium concentrations and those concentrations are determined by the degrees of freedom actually chosen, i.e. the partial pressures rather than the carrier concentration. (5) Rapid group II diffusion is often associated with a metastable, but supersaturated, group II substitutional concentration which becomes unstable as Ga interstitials trigger a rapid relaxation of the substitutional concentration toward an equilibrium state.

We will limit this section to a discussion of results obtained from what we consider to be well-designed (and thus easily interpretable) experiments. Thus, this section will be limited to a fraction of the GaAs diffusion literature. In Section 5.1, we shall discuss the results of experiments which appear to justify the assumption of an equilibrium native defect concentration during annealing. In Section 5.2, we shall discuss selected nonequilibrium experiments for which it appears possible to interpret the results in terms of perturbations upon an equilibrium state, i.e. to interpret kinetic mechanisms. Interdiffusion measured after closed ampoule annealing will be compared with that measured after open tube annealing in Section 5.3, and the similarities and differences will be used to interpret native defect mechanisms and to evaluate the experimental techniques. Section 5.4 will briefly summarize the nonequilibrium diffusion results obtained after annealing under dielectric caps. Section 6 will summarize the current model of point defect interactions associated with impurity diffusion in GaAs. Although diffusion associated with annealing after ion implantation is of substantial technological importance, it is beyond the scope of this paper and will not be discussed.

5.1. Quasi-equilibrium results from native defect isoconcentration

A conceptually simple experiment, and one which appears possible to attain occasionally in practice, is that of native defect isoconcentration. To appreciate this experiment, consider uniformly doped GaAs which has equilibrated with the vapor, i.e. a system defined by the three degrees of freedom, T , P_{As} , and P_{dop} . In such an experiment, the native defect concentrations are at their equilibrium values throughout the GaAs. If a small quantity of an isoelectronic impurity were somehow added to the GaAs, then this impurity would diffuse at a rate determined by the equilibrium native defect concentration and jump rates. Studying the changes in marker diffusion versus changes in the equilibrium native defect concentrations, i.e. by changing one degree of freedom at a time, would provide significant (although not complete) characterization of the dominant native defect(s).

As discussed in Sections 2 and 3, control of the correct number of degrees of freedom is a necessary condition if one is to define the equilibrium state of the system. We shall assume that a nonuniform concentration profile of a diffusion marker has been grown into doped GaAs which is to be annealed in contact with the vapor. With the presence of a fourth component (the marker impurity), a fourth degree of freedom appears in the anneal system. The final equilibrium state will be defined by T , P_{As} , P_{dop} , and x_{marker} , the mole fraction of the marker element when it is eventually distributed throughout the solid at equilibrium. Of course, we have assumed that the impurity is always at a sufficiently low concentration that it does not significantly affect the solid composition, i.e. it does not change the equilibrium native defect concentrations. If this assumption holds, then the native defect concentrations are uniquely defined by the three degrees of freedom, T , P_{As} , and P_{dop} . If the impurity

diffuses by a mechanism involving native defects, then the strategy of using impurity diffusion to probe changes in well-defined equilibrium native defect concentrations makes logical sense.

Of course, the *sufficient* condition for native defect equilibration is satisfied only when it is possible for the dependent thermodynamic variables, i.e. the native defect concentrations, to actually reach their equilibrium values in a reasonably short time, i.e. a time which is significantly shorter than the time used for the experiment. Attaining the sufficient condition in a practical time will depend upon the experimental design, and success must be verified experimentally. For example, a value of marker diffusivity which varies with the independent variables, and does not depend upon the anneal time, is strong evidence that the concentration of the native defect controlling diffusion has approached an equilibrium value.

There are very few results reported in the literature using low marker concentrations. However, within the (rather substantial) experimental noise typically observed, the evidence to be discussed indicates that the interdiffusion of the marker Al, $D_{\text{Al-Ga}}$, is roughly independent of the Al fraction. In addition $D_{\text{Al-Ga}}$ appears to be comparable both to the interdiffusion measured using the marker In, $D_{\text{In-Ga}}$, with x_{In} of up to several %, and to the isotopic self-diffusion, D_{Ga} . Although we recognize the subtle differences between interdiffusion and self-diffusion, those differences are minor compared with the experimental noise associated with marker diffusion in GaAs. Thus, we shall routinely compare interdiffusion and self-diffusion without distinguishing between their differences.

In Section 5.1.1, we shall begin by comparing the results for self-diffusion and marker diffusion in intrinsic GaAs. This will be followed with a discussion of several key isoconcentration experiments, in both n-type and p-type GaAs, which provide substantial insight on the native defects. We shall also compare some of the limitations of real experimental techniques.

5.1.1. Isotopic self-diffusion results: intrinsic GaAs

Measurements of the self-diffusion of Ga in GaAs have been reported by four groups. Goldstein [67] made the first measurements after electroplating a radioactive isotope, ^{72}Ga , onto a piece of GaAs, placing the GaAs in an evacuated ampoule and annealing at temperatures of $1100 \leq T \leq 1230$ °C. After annealing, thin layers of the GaAs were removed, the rate of radioactive decay was measured, and a plot of the ^{72}Ga profile versus depth was made. From this profile, self-diffusion coefficients were calculated.

Unfortunately, this report serves as an example of how *not* to do an experiment with a compound semiconductor. Even at $T \approx 1200$ °C, Ga has a small vapor pressure and virtually all of the Ga metal is in a liquid phase. The liquid Ga will equilibrate with the solid GaAs only when the liquid metal has an arsenic composition of approximately $x_{\text{As}} \approx 30\%$ [11]. Since no As was added to the ampoule to begin with, a portion of the GaAs sample must have decomposed during annealing. Presumably, Ga-rich liquid droplets formed on the solid surface. If this happened, then the some of the solid under those droplets would have dissolved rapidly and left a pitted surface, and any sectioning technique would be unreliable and would overestimate the diffusivity. On the other hand, if all of the liquid Ga somehow rolled off of the GaAs surface, then the solid would equilibrate with the liquid through mass transport via the vapor. In this case, the partial pressures in the ampoule must have drifted for period of time and the measured diffusion would presumably reflect the time dependent native defect concentrations caused by a Ga-rich nonequilibrium ambient as it changed with time. The time required for the system to reach its equilibrium state, i.e. GaAs in equilibrium with a Ga-rich liquid, would depend upon the original mass of both the GaAs and the Ga, and the ampoule volume. None of this was discussed by Goldstein, and it is not possible to determine the kinetic details associated with his work. However, several examples will be discussed in this section which illustrate the importance and magnitude of the nonequilibrium effects associated with a Ga-rich ambient. Fig. 21 shows Goldstein's

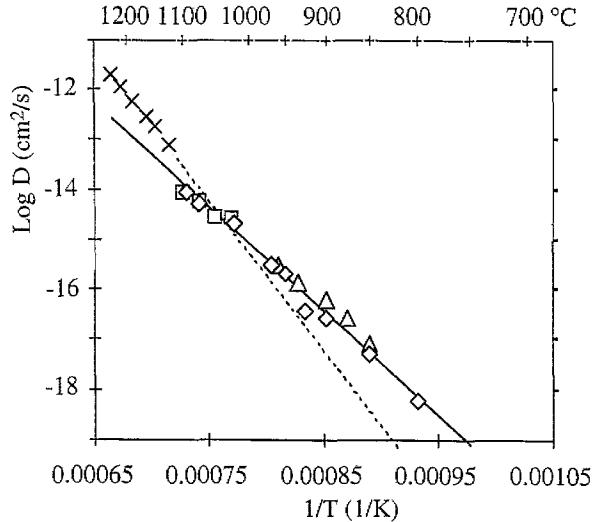


Fig. 21. D_{II} vs. $1/T$ for isotopic Ga self-diffusion in intrinsic GaAs. The Goldstein data (\times) were taken under conditions far from equilibrium. The data of Palfrey et al. (squares), Tan et al. (triangles), and Wang et al. (diamonds) were all taken near $P_{\text{As}} = 1 \text{ atm}$. The line of Wang et al. (solid) fits the $P_{\text{As}} = 1 \text{ atm}$ experiments better than the line of Tan and Gosele (dashed).

data (\times) compared with other self-diffusion results. It should be emphasized that the difference cannot be simply accounted for by the difference in P_{As} .

Fig. 21 shows the data reported by Wang et al. [68] (diamonds) using diffusion couples made of epitaxial isotopic heterojunctions of $^{69}\text{GaAs}-^{71}\text{GaAs}$. Except for their two highest temperature measurements (not shown here), the anneals were performed in quartz ampoules containing sufficient As to give $P_{\text{As}} \approx 1 \text{ atm}$. The two high temperature measurements were made in an open rapid thermal anneal system using a GaAs proximity cap, N_2 , and no added As, and will be discussed in Section 5.3. Rather than pick a constant diffusion time, Wang et al. chose diffusion times which were expected to give a characteristic diffusion length, $L = 2\sqrt{Dt} \approx 0.2 \mu\text{m}$, to assure that the concentration profiles taken by SIMS could be reliably compared. The solid line shown in Fig. 21 fits $D = 43 \exp\{-4.24 \text{ eV}/kT\} \text{ cm}^2 \text{ s}^{-1}$, as suggested by Wang et al. For comparison, the dashed line fits $D = 2.9 \times 10^8 \exp\{-6 \text{ eV}/kT\} \text{ cm}^2 \text{ s}^{-1}$, as recommended by Tan and Gosele [8].

Palfrey et al. [17] diffused ^{72}Ga into GaAs using quartz ampoules containing $P_{\text{As}} \approx 1 \text{ atm}$. Their experiments were conducted over a temperature range, $1025 \leq T \leq 1100 \text{ }^\circ\text{C}$, and the results (squares) are consistent with the Wang line shown in Fig. 21. Palfrey et al. considered it necessary to go to extensive lengths to reduce contamination and to bake out their quartz before use.

Over the temperature range, $850 \leq T \leq 960 \text{ }^\circ\text{C}$, Tan et al. [69] conducted closed ampoule interdiffusion of $^{69}\text{GaAs}-^{71}\text{GaAs}$ isotopic superlattices. These experiments were conducted at a nominal $P_{\text{As}} \approx 1 \text{ atm}$. These results (triangles) also fall consistently along the Wang line. However, we note that Tan and Gosele considered their results to be in error because the results did not fall on their previously proposed line, i.e. the dashed line shown in Fig. 21. Unfortunately, as shown in Section 5.1.3, it is the line of Tan and Gosele which is the result of both experimental and interpretive errors.

The close fit of the data from three groups indicates that the dominant native defect concentrations have approached consistent values. Superficially, this suggests that the three groups met both the necessary and sufficient conditions to bring the native defect concentrations to their equilibrium values in a short time. However, as discussed below, several results strongly suggest that other subtle issues

may have affected the diffusion in these ampoules. Thus we consider the extent to which these data reflect the results of equilibrium native defect concentrations for GaAs to remain uncertain.

5.1.2. Interdiffusion versus P_{As} : defect isoconcentration in intrinsic GaAs

The discussion of Section 3.2.2 has made it clear that if $D_{III} \propto P_{As}^x$, then we can eliminate half of the six native defects because only three defect concentrations increase with P_{As} . Although a measured $D_{III} \propto P_{As}^x$ is qualitatively consistent with an increasing Ga vacancy concentration, it is also consistent with changes in the As interstitial concentration and one might wonder if this could induce local stresses (or antisite defects) which cause atoms on the group III sublattice to move to interstitial sites and diffuse more quickly. Indeed, the closer an experiment is to equilibrium, the less kinetic information is going to be learned. One must perturb particular defect concentrations in different ways in order to test an atomistic model. Of course, this is easier said than done, and at this time there appear to be no experimental results available which can directly distinguish between group III diffusion driven by V_{Ga} versus I_{As} . However, if I_{As} affected D_{III} , then one would expect that D_V would vary similarly because the bonds holding Ga in substitutional sites are the same as those holding the As in substitutional sites. The lack of evidence for rapid group V interdiffusion at higher P_{As} and higher carrier concentration (i.e. when both D_{III} and I_{As} are expected to be large) suggests that there is probably no significant effect on D_{III} associated with an increased I_{As} concentration. Thus, we consider it reasonable to conclude that the Ga vacancy is the most likely defect responsible for the transport when $D_{III} \propto P_{As}^x$ is observed, and we shall interpret such observations in terms of a Ga vacancy.

The P_{As} dependence which first suggested that a vacancy mechanism controlled the interdiffusion of quantum wells appears to have originated with the results of Kaliski et al. [30]. They used TEM micrographs to show that interdiffusion was faster in n-type films when $P_{As} \approx 4 \text{ atm}$ than when $P_{As} \sim 10^{-8} \text{ atm}$ (i.e. during sublimation of the sample) in a closed ampoule. Not long afterward, measurements of D_{Al-Ga} from SL interdiffusion were reported by two groups [70,71] as a function of P_{As_4} . Using a hundredfold variation of P_{As_4} , both results reported a minimum in D_{Al-Ga} near $P_{As_4} = 1 \text{ atm}$. Based on these results, it was suggested [34] that the minimum might correspond to a crossover between diffusion controlled by the Ga vacancy and diffusion controlled by a Ga interstitialcy. However, when the results of more comprehensive studies were reported later by Olmsted et al., [72,73] it became clear that the minimum in D near $P_{As} = 1$ was simply experimental noise. Indeed, experimental noise is sufficiently severe that one is forced to critically examine interdiffusion results from several groups over as wide a range of P_{As} and T as possible.

Several groups have conducted three kinds of closed quartz ampoule experiments: (1) with elemental As added, (2) with neither As nor Ga added, and (3) with elemental Ga added to the ampoule containing the GaAs. For (1), a calculation of P_{As_4} depends upon the accuracy of the measurements made of the As weight and ampoule volume, and this calculation is not subject to large errors. However, Section 2.1 has shown that experiments (2) and (3) must (eventually) correspond to exactly the same *equilibrium value* of P_{As_4} and thus the same equilibrium state for the entire system. Experiment (2) will approach equilibrium quicker than experiment (3) will simply because a shorter period of time is required for a small amount of GaAs decomposition to take place and bring the GaAs to equilibrium with the vapor and a small amount of Ga-rich liquid, saturated with a few % of As. For example, at a typical annealing temperature of $T = 850^\circ\text{C}$, Fig. 1 shows that $\bar{P}_{As} \approx \bar{P}_{As_2} \approx 2 \times 10^{-6} \text{ atm}$ over the Ga-rich edge of the solidus curve. Thus, for experiment (2), only about 4 ng of As, or approximately 0.03 ML of GaAs needs to decompose from a sample with 1 cm^2 surface area in an ampoule of 1 cm^3 volume to bring the vapor pressures to their equilibrium values (assuming that no chemical reactions occur with the ampoule). Experiment (3) was already discussed in Section 2.1.2.1 for an ampoule of 1 cm^3 volume containing 3 mg of Ga and a sample with surface area of 1 cm^2 , and

it was calculated that approximately 120 μg of As left the GaAs, or approximately a 0.6 μm thick layer will decompose from the GaAs. Thus, it should be clear that experiments (2) and (3) will eventually attain the same equilibrium state, but that experiment (2) will reach equilibrium first. Any difference observed in the marker diffusion in the two experiments is expected to be a result of these different initial periods of time in which the native defect concentrations gradually drift toward their equilibrium values. Thus, when we discuss *equilibrium data* taken at the *Ga-rich* edge of the GaAs solidus line, we shall only discuss data *measured without any added Ga*, independent of how the original authors interpreted this equilibrium state. In Section 5.3, we shall discuss the very different interdiffusion results for nonequilibrium experiments conducted with Ga intentionally added to the ampoule.

Fig. 22 plots D_{III} versus P_{As_4} , for data measured at several temperatures. Al was used as the interdiffusion marker in nominally intrinsic GaAs. Results presented in Section 5.3 will show that a differences in the Al fraction are not responsible for the variations seen in this figure. At any temperature, the minimum value of P_{As_4} is taken to be that corresponding to a Ga-rich liquid–solid–vapor equilibrium, and the corresponding values of D at low P_{As} have been taken from diffusions performed without any intentionally added Ga or As. Values of the minimum P_{As_4} over the Ga-rich solidus at $T = 700$, 850, and 1000 °C are illustrated by the vertical dashed lines. Higher pressure data, using roughly $P_{\text{As}} \sim 1$ atm, can be seen on the right side of the figure. Many of those data have been offset along the P_{As} axis by $\pm 20\%$ for ease of view.

In 1992, Olmsted et al. [72] reported their initial measurements of $D_{\text{Al-Ga}}$ versus P_{As_4} for the closed ampoule, $T = 855$ °C, interdiffusion of multiple quantum wells. Soon afterwards, this group reported [73] the most complete interdiffusion study to date for GaAs. They used SIMS measurements to confirm [10] the accuracy of their diffusivities calculated from the photoluminescence energy shift

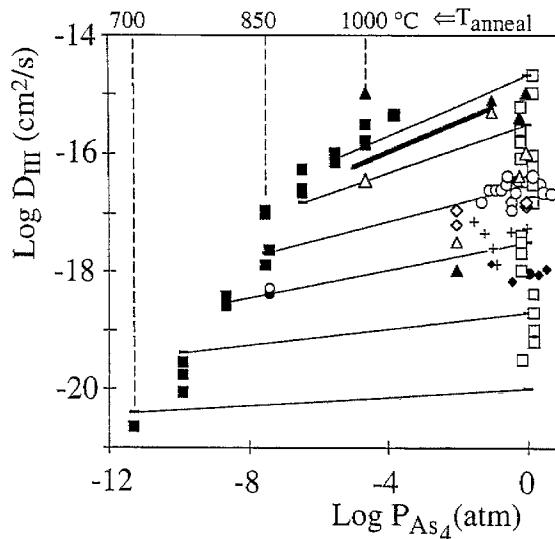


Fig. 22. D_{III} vs. P_{As_4} measured in nominally intrinsic GaAs. Using the best-fit equation for D_{III} vs. $1/T$ from Olmsted et al. [73], the light lines connect D_{III} (bars) at high and low P_{As_4} in 50 °C increments over $T = 700$ –950 °C. Their corresponding data is also shown at the equilibrium P_{As_4} above the Ga-rich solidus (filled squares) and at $P_{\text{As}} \approx 1$ atm (open squares), where the data have been offset for ease of viewing. The heavily shaded line illustrates a slope of 1/4, i.e. the slope expected for the vacancy mechanism when native defect isoconcentration holds. Data from many groups illustrate the typically large measurement noise seen even within individual groups. These data come from Baba-Ali et al. [74] (filled and open triangles, top and bottom QWs, respectively) at $T = 1000$ °C, Seshadri et al. [75] (open diamonds) at $T = 888$ °C, Furuya et al. [70] (+) at $T = 850$ °C, Hsieh et al. [76] (filled circles) at $T = 850$ °C, and Guido et al. [77] (solid diamonds) at $T = 825$ °C, and the initial 1992 Olmsted et al. [72] data (open circles) at $T = 855$ °C.

of the interdiffused quantum wells. In general, this group appears to have made a more painstaking effort to use clean experimental techniques than those used (or reported) by several other groups in recent years, and their data represents a significant contribution to the literature. D_{III} was measured in 50 °C increments over $T = 700\text{--}1050$ °C, in a quartz ampoule containing either $P_{\text{As}} \approx 1$ atm, added Ga, or no added Ga or As. Because they assumed that the first two ambients were more important than the third, they only published a summary line (and no data) for the more useful Ga-rich equilibrium case (i.e. no added Ga). We thank Drs. Olmsted and Houde-Walter for kindly making those previously unpublished data available for use in this review [78]. These data are tabulated in Section 5.3, and they are also shown in Fig. 22 (filled squares) and one may see that they are clearly lower than the data from the same group taken at $P_{\text{As}} \approx 1$ atm (open squares).

At six different temperatures, $700 \leq T \leq 950$ °C, lines connect values of D_{III} obtained from the best fit equations of Olmsted et al. [73] for D_{III} versus $1/T$ at high and low P_{As} . A heavily shaded line with a slope of $1/4$ is provided for comparison. At any of the highest temperatures, D_{III} versus P_{As} clearly approaches a slope close to $1/4$, and this is the behavior expected if V_{Ga} controls the group III interdiffusion. At lower temperatures, the slope of the lines are tending toward zero, and this is strong evidence that the native defect concentrations have not equilibrated in a time short compared with the anneal time. Indeed, it appears likely that the concentration of V_{Ga} has approached its equilibrium value relatively quickly at higher temperatures. The measurements of Olmsted et al. are the best evidence to date that a Ga vacancy mechanism controls interdiffusion in intrinsic GaAs. In Fig. 22, it can be seen that the magnitude of the measurement noise of individual groups and between different groups is significantly larger than that reported either in the isotopic measurements summarized by Fig. 21 or in the measurements reported for many metallurgical systems. This suggests that an important variable may have been uncontrolled, and this will be discussed further in Section 5.3.

Several other reports of D_{III} versus P_{As} have been summarized in Fig. 22, when at least a 100-fold variation in P_{As} was used. Data from the initial report of Olmsted et al. [72] at $T = 855$ °C are also shown (open circles) and they appear consistent with the broader study from this group. Early reports of D_{III} versus P_{As} from Furuya et al. [70] (+) at $T = 850$ °C, Guido et al. [77] (filled diamonds) at $T = 825$ °C, and Hsieh et al. [76] (filled circles) at $T = 850$ °C are qualitatively consistent with the expected pressure dependence, given their large measurement noise. More recent reports also provide qualitative confirmation of the P_{As} dependence. Baba-Ali et al. [74] reported annealing nominally identical samples containing two quantum wells at different P_{As} at $T = 1000$ °C (open and filled triangles, bottom and top QWs, respectively). The quantum well near the surface was reported to diffuse faster than the deep quantum well in all cases, and the reported differences appear to be relatively large. Only a modest depth dependence was reported by Seshadri et al. [75] (open diamonds) at $T = 888$ °C.

Fig. 22 also highlights significant differences in data reported by different groups at supposedly the same temperature. This variation strongly suggests that the temperature has not always been measured accurately. Because of the very large heat loss through radiation at these temperatures, overestimation of the correct temperature is the most likely error. Considering how little discussion is given to the control or calibration of temperature, we suspect that this may be a widespread error.

Allowing for the sizable measurement noise, we conclude that there is a qualitative fit to $D \propto P_{\text{As}}^{1/4}$. At higher temperatures, it appears that the concentration of native defects, or at least that of V_{Ga} , has approached equilibrium with the ambient in several experiments. This is the best evidence available to date that group III interdiffusion proceeds via a Ga vacancy mechanism in intrinsic GaAs, across the entire allowed As pressure range, and under conditions which approximate well defined native defect equilibrium at $T > 800$ °C. The lack of evidence for any upturn in diffusivity at low P_{As} shows that the interstitialcy, or kick-out, mechanism does not occur under quasi-equilibrium conditions.

5.1.3. Evidence for the diffusion mechanism in n-type GaAs

As discussed in Section 5.1.2, the D_{III} versus P_{As} dependence provided evidence that Ga vacancies control interdiffusion in intrinsic GaAs. The photomicrographs reported by Kaliski et al., [30] taken after quantum well interdiffusion under high and low P_{As} , were the earliest indication that interdiffusion is also controlled by the V_{Ga} in n-type GaAs. Not long afterward, Mei et al. [79] published an interdiffusion study of AlAs–GaAs multilayers, using SIMS to characterize the interdiffusion. Selected results from this study led Tan and Gosele [8] to conclude that interdiffusion was controlled by a negative, triply charged, Ga vacancy.

5.1.3.1. The triply charged Ga vacancy model

Tan and Gosele noted that a line extrapolated from the high temperature self-diffusion data of Goldstein passed through the low temperature data of Mei et al., [79] after appropriate adjustment of Mei's data. The adjustment was made after *explicitly assuming* that the interdiffusion was controlled by the V_{Ga}^{-3} and *implicitly assuming* that an equilibrium native defect concentration existed. Based upon these and other results, Tan and Gosele proposed [8] in 1988 that diffusion on the group III sublattice in intrinsic and n-type GaAs was controlled by the V_{Ga}^{-3} over a very wide range of temperature. Selected results were taken from the literature to produce an extraordinarily good fit to $D(T, n_i) = 2.9 \times 10^8 \exp(-6 \text{ eV}/kT) \text{ cm}^2 \text{ s}^{-1}$, independent of P_{As} , as shown in Fig. 23. The extremely good fit of the data to the line shown, all within a factor of about 2, appeared to be compelling evidence that this equation described self-diffusion and interdiffusion in intrinsic GaAs.

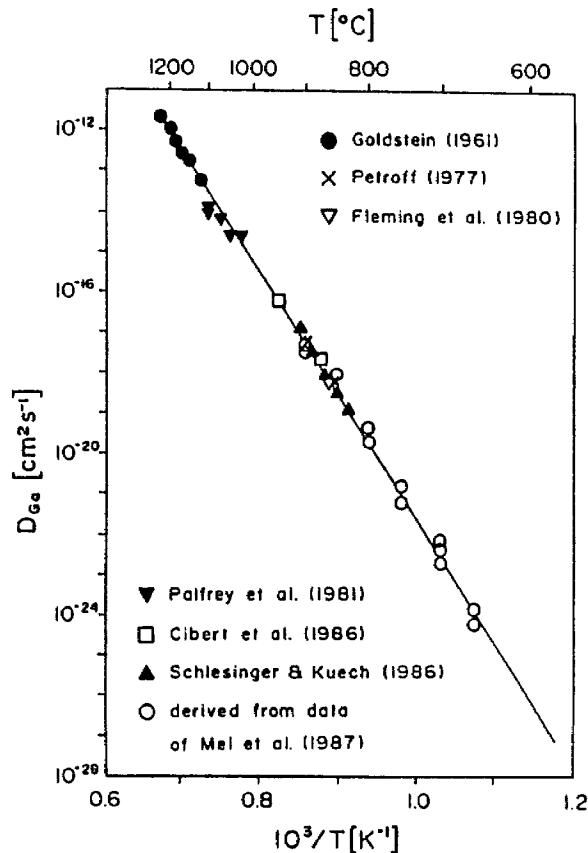


Fig. 23. Tan and Gosele's original D vs. $1/T$ plot. Reprinted from Ref. [8], copyright 1988, with kind permission from Elsevier Science S.A., P.O. Box 564, 1001 Lausanne, Switzerland.

Of course, this equation described diffusion as a function of T alone, i.e. without reference to the correct number of degrees of freedom, and thus it could not be correct. In 1990, Cohen [6] pointed out that P_{As} should be present in any expression describing diffusion in GaAs. Given that (1) the slope of the Tan and Gosele line depended largely upon the Goldstein and the adjusted Mei et al. data, and (2) both of these results had been obtained without any control over the As overpressure, Cohen suggested that the Tan and Gosele equation applied to GaAs with a composition at the edge of the Ga-rich solidus. Since it is customary to define standard conditions at a total pressure of 1 atm, and since the only volatile species present is As, Cohen suggested that standard conditions for diffusion in GaAs should be chosen to be at $P_{\text{As}_4}^{1/4} = 1 \text{ atm}$ in the intrinsic crystal where $n = n_i$. A key implicit assumption here was also that solid-vapor equilibrium, and thus native defect isoconcentration, could be applied. If this were correct, then the Tan and Gosele equation for D should be divided by an expression describing $P_{\text{As}_4}^{1/4}(T)$ over the Ga-rich edge of the solidus to obtain D at $P_{\text{As}} = 1 \text{ atm}$, as discussed in Section 2.5. Using an expression which is valid below approximately $T = 1200 \text{ }^\circ\text{C}$,

$$P_{\text{As}_4}^{1/4} = 1.61 \times 10^4 \exp\left\{\frac{-1.35 \text{ eV}}{kT}\right\} \text{ atm}^{1/4} \quad (141)$$

Cohen proposed that group III interdiffusion controlled by a triply charged vacancy should be described by $D(T, P_{\text{As}} = 1, n_i) = 1.8 \times 10^4 \exp(-4.65 \text{ eV}/kT) \text{ cm}^2 \text{ s}^{-1}$, valid in an intrinsic crystal under $P_{\text{As}} \approx P_{\text{As}_4} = 1 \text{ atm}$.

In principle, the magnitude of the charge associated with the V_{Ga} can be obtained from the dependence of D_{III} upon n . In Section 3.4, it was shown that, if an equilibrium concentration of $\text{V}_{\text{Ga}}^{-3}$ exists and if this is the only defect which affects interdiffusion, then a semi-empirical description of the interdiffusion may be useful. Such a description relates the *measured* diffusivity, after annealing under a chosen value of P_{As_4} , in a doped crystal to an *empirical standard* diffusivity, $D_{\text{V}_{\text{Ga}}^{-3}}(T, P_{\text{As}}^{\text{Std}}, n_i)$, in the intrinsic GaAs, i.e.

$$D_{\text{meas}} = D_{\text{V}_{\text{Ga}}^{-3}}(T, P_{\text{As}_4}^{\text{Std}}, n_i) P_{\text{As}_4}^{1/4} \left(\frac{n}{n_i}\right)^3 \quad (142)$$

In 1991, Tan and Gosele published [5,41] a reanalysis of the same data. They concluded that since enough of the data was taken at high P_{As} , their original equation for a vacancy controlled D_{III} actually applied at $P_{\text{As}} = 1 \text{ atm}$. Since this is clearly erroneous for the data taken at the highest and lowest temperatures, we consider it instructive to reanalyze these early results and reconstruct the D_{III} versus $1/T$ figure, point by point, based upon the results known in 1988. In Fig. 24, we compare selected published data to three lines: the Tan and Gosele line (dashed), the Cohen line (dot-dashed) and the Wang self-diffusion line (solid). The selection criteria for the data is discussed in the following paragraphs. No adjustments to the data have been made, except for the Mei data which have been adjusted for the effect of carrier concentration alone, assuming an equilibrium $\text{V}_{\text{Ga}}^{-3}$ model applies. Although the data superficially appear to follow the Tan and Gosele line, significant insight can be gained by reassessing the data which has been or should have been used in this plot.

5.1.3.1.1. High temperature results

As discussed in Section 5.1.1, Goldstein's samples must have exhibited substantial decomposition in an ampoule containing added Ga when heated close to the melting point of GaAs. This means that a long period of time was required for the native defect concentrations to approach their equilibrium values, if in fact they were ever reached. Thus, it is inappropriate to use Goldstein's data for comparison with any well designed experiments which quickly bring the native defect concentrations to their equilibrium values, and thus they are not included in Fig. 24. In contrast, the data of Palfrey et al. [17]

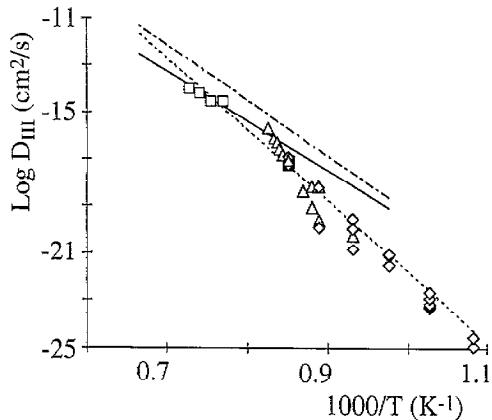


Fig. 24. Reconstructed plot of D vs. $1/T$ using the sources cited by Tan and Gosele in 1988. The text discusses the reasons why this plot differs from that in Fig. 23. For comparison, lines shown are from Tan et al. (dashed), Wang et al. (solid), and Cohen (dot-dashed).

(open squares) were measured after annealing in a quartz ampoule at $P_{As} \approx 1$ atm. These data appear to be of high quality, i.e. the samples were not decomposing and based upon Fig. 22, the native defects presumably approached their equilibrium values in a relatively short time.

5.1.3.1.2. Mid-temperature results

One data point (filled square) comes from Petroff et al., [80] who had calculated D from TEM micrographs of an AlAs–GaAs SL taken before and after annealing at $T = 900$ °C reportedly ‘under arsenic supersaturation condition’. Although the densitometer measurement incorrectly assumed that the film response is linear, it does provide an order of magnitude estimate of D_{III} .

Tan et al. selected a single data point from the work of Fleming et al. [81]. Fleming et al. reported several results (open triangles) for interdiffusion after annealing with a GaAs proximity cap and no added P_{As} . All of their results are shown in Fig. 24. Diffusivities associated with many quantum wells, from nominally 1 ML width to much larger widths, were measured by several different techniques. This work is difficult to evaluate because of all of the experimental variations attempted.

Tan et al. also used data from Cibert et al. [82]. We note that this data did fall upon the desired line, but that it was a result obtained for diffusion *after ion implantation*. Because ion implantation induces a high density of native defects, we consider this data to be inappropriate for comparison with data for which a defect isoconcentration applies, and thus we do not include it in Fig. 24. Similarly, Tan et al. used the 1986 data of Schlesinger and Kuech [83] which also fell on the desired line. However, these data were measured for the interdiffusion of $Al_{0.7}Ga_{0.3}As$ –GaAs quantum wells after being capped with ‘ SiN_x ’. A 1987 report from this same group reported very different results using the same technique but with different Al mole fractions, and those results were not included in the Tan et al. plot. Diffusivities measured after annealing under dielectric caps will be discussed in Section 5.4, and it will become quite clear that they should not be compared here. Thus, these data are not shown in Fig. 24.

5.1.3.1.3. Low temperature results

What had made the original Tan et al. line so interesting was the marvelous fit of the calculated diffusivity in the intrinsic crystal at low temperatures. This data had been obtained by adjusting interdiffusion data, measured for the $T = 700$ °C anneal of an n-type SL, downward by a factor of n/n_i [3]. For comparison with other data, we note that the Mei et al. anneals included a proximity cap but no added P_{As} . Our extrapolations of the Mei data (also adjusting for only carrier concentration,

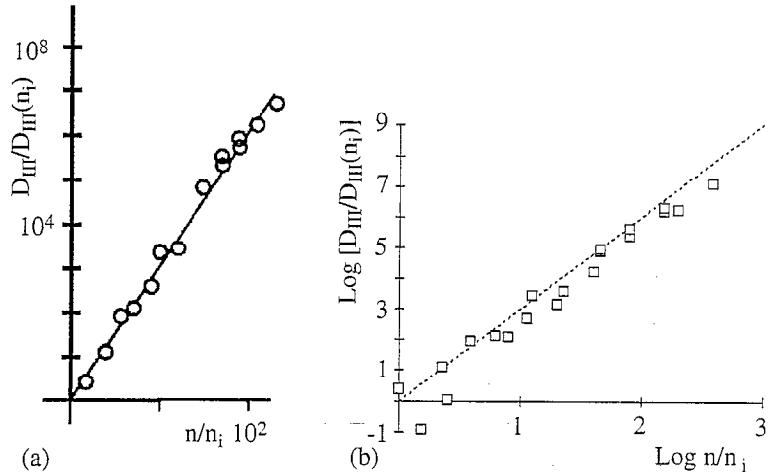


Fig. 25. Plots of the normalized D/D_i vs. n/n_i for the enhanced diffusion of the Si-doped AlAs-GaAs SL of Mei et al. (a) The original plot of Tan and Gosele showed a slope of 3 and strongly suggested that most of the V_{Ga} in the crystal had a charge of -3 . Reprinted from Ref. [8], copyright 1988, with kind permission from Elsevier Science S.A. P.O. Box 564. 1001 Lausanne, Switzerland. (b) The same data have been replotted after using improved adjustment factors. Although the slopes are qualitatively similar to (a), any interpretation of this data would be misleading as discussed in the text.

but not P_{As}) find that several of Mei's 20 data points deviate from the calculated line by more than a factor of 2, as shown in Fig. 4 by the open diamonds. It is unclear whether some of these points were simply not shown in the original Tan and Gosele plots, or whether they overlapped each other because of the values of n_i which were used to make the adjustments.

In our analysis, we have used the carefully determined and generally accepted values for n_i provided by Blakemore [84]. These values account for the relatively low X-minimum in the conduction band which has the effect of changing the effective density of states at elevated temperatures. Tan et al. used a textbook effective mass model for their calculation of n_i , [5] i.e.

$$n_i = \sqrt{N_C N_V} \exp\left\{\frac{E_g}{2kT}\right\} \quad (143)$$

In addition, Tan et al. kept the effective density of states for both the conduction and valence bands independent of temperature. They attempted to correct for this by modifying the temperature dependence of the bandgap, but it is unclear how this correction factor was obtained. Using the values of n_i thus obtained, they plotted $D_{meas}/D(n_i)$ versus n/n_i for several of Mei's points, and demonstrated the strikingly good fit to a cubic dependence shown in Fig. 25(a). Using the more accurate values of n_i obtained from Blakemore, we have replotted the data of Mei et al., as shown in Fig. 25(b). Although the revised plot shows much more noise, the data still appear to exhibit a cubic dependence upon carrier concentration.

It would appear that the reanalysis of the data used by Tan and Gosele provides evidence that a cubic dependence of D_{III} upon n is present. However, the following reanalysis of the key Mei experiment actually demonstrates that the entire model is flawed.

5.1.3.1.4. The key experiment behind it all

Mei et al. [79] reported what initially appeared to be a well controlled group III diffusion study where the interdiffusion of a GaAs/AlAs superlattice (SL) was characterized by secondary ion mass spectroscopy (SIMS). The SL was grown by MBE containing several Si-doped regions of different concentration, each of thickness 2000 Å and separated from each other by undoped spacer layers of

thickness 1000 Å. Pieces of the wafer were annealed at different temperatures, and a family of curves plotted for D_{III} versus $1/T$ as a function of n (assumed to be the same as N_{Si}). Diffusions were performed at various temperatures for $t = 3$ h using a GaAs proximity cap and a H₂:Ar ambient.

Tan and Gosele [8] were the first to recognize that the $T = 700$ °C data of Mei et al. showed a nearly perfect cubic dependence, $D_{III} \propto n^3$, over $D < 10^{-20}$ cm² s⁻¹ to $D \sim 10^{16}$ cm² s⁻¹. However, potentially serious problems with the original data have been generally ignored. These problems include (1) the use of data taken at an unusually low annealing temperature to interpret diffusion data at higher temperatures, (2) data taken at other temperatures disagreed with the cubic dependence on electron concentration, (3) the very low amplitude variation of the SIMS Al signal in the control sample suggested resolution problems for all of the results, (4) the extraordinary SIMS resolution required to obtain several of the reported values of D_{III} , (5) the lack of any reduction in D_{III} in the undoped spacer layers, and (6) the very close proximity, 300–400 Å, of the first AlAs layer to the surface, i.e. a layer which may have oxidized and affected the measurements.

A brief discussion follows. (1) The conclusion that the V_{Ga}⁻³ controlled the diffusion in GaAs was based upon a single plot [79] of data taken at $T = 700$ °C. As already seen in Fig. 22, such low temperatures require very long times for the native defect concentrations to approach their equilibrium values, and this approximation almost certainly does not hold in this case. (2) Data taken in 50 ° increments over $T = 650$ –900 °C suggested that values for an exponent of $x = 1.5$ to 7 could fit an equation of the form $D_{III} \propto n^x$, depending upon the temperature chosen. The substantial scatter in the raw data can be seen in Fig. 26 which plots D versus n for each of the temperatures reported by Mei et al. [79]. This kind of plot is a far more demanding test of D versus n because it shows the raw noise in the slope, rather than the adjusted noise over many decades as shown in Fig. 25. (3) In the control sample, the Al concentration profile showed more than an order of magnitude variation in the 2000 Å wide region next to the surface, but the variation in the Al control signal dropped quickly to a less than three-fold in several of the deeper layers. D_{III} was then estimated by comparing the Al profiles in each of the layers after annealing with their counterparts in the control sample. This is a severe problem which would have reduced the reliability of the calculated diffusivities. (4) For the lowest value of D reported at $T = 700$ °C, we calculate the characteristic diffusion length to be approximately 2 Å, i.e. several of the diffusivities imply characteristic lengths which were simply too small to be determined

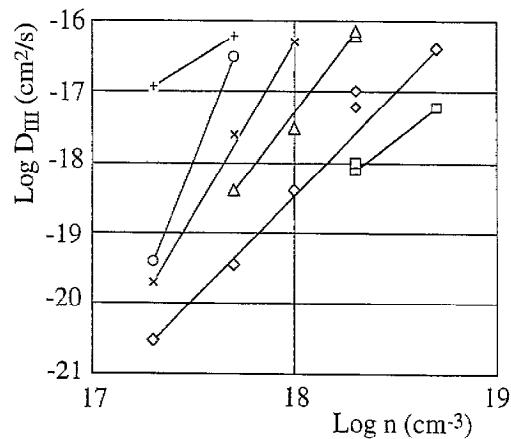


Fig. 26. D_{III} vs. n , measured at different temperatures, summarized from the Si-doped interdiffusion study of Mei et al. [79]. Data has been taken at $T = 650$ °C (squares), 700 °C (diamonds), 750 °C (triangles), 800 °C (×), 850 °C (circles), and 900 °C (+). Lines connect the lowest and highest measurements at each temperature to qualitatively show that the slopes range from about 2 to 7. Little consistency is found between the slopes. This strongly suggests that any quasi-equilibrium interpretation will be unreliable.

by SIMS. (5) In the 1000 Å thick undoped regions between the Si-doped layers, the SIMS profiles showed no indication that D_{III} was reduced by the low carrier concentration. Since the Debye length is only a fraction of the thickness of this layer, D_{III} should have varied substantially. (6) It is known [18,85] that even thick AlAs layers can degrade and convert to an oxide in a few minutes when exposed either directly, or through pinholes, to ambient room air. Evidence discussed in Sections 5.3 and 5.4 suggests that such an oxide is likely to affect diffusion. However, we shall never know the extent to which oxidation may have affected these particular samples. From the first five items above, and particularly Fig. 26, it should be clear that the data obtained from the Si-doped samples of Mei et al. do not provide credible evidence for an interpretation of diffusion controlled by a V_{Ga}^{-3} , or any other mechanism.

It should be noted that Tan et al. [86] recently concluded that the concentration of Ga vacancies must increase, and that the crystal must become more As-rich, as T is reduced. The argument provided appears to be logically consistent assuming that the $T = 700$ °C data of Mei et al. is correct and that the V_{Ga}^{-3} controls diffusion. However, GaAs is generally expected to deviate from stoichiometric perfection as the temperature is increased, rather than decreased, and we consider such a conclusion to be entirely unjustified without strong independent evidence.

5.1.3.2. *The 'No-Fermi level effect' model*

Seshadri et al. [75] reported the results of the interdiffusion of three quantum wells (QWs) grown into a 1 μm wide undoped region lying between n-type and p-type GaAs. The anneals were performed at $T = 888$ °C using either $P_{\text{As}} = 0.01$ or $P_{\text{As}} = 1.0$ atm in an open system. The QWs were grown with slightly different widths, so that they could be easily distinguished by PL spectroscopy. Treating the undoped region as semi-insulating at the anneal temperature, i.e. assuming that the energy band diagram of Fig. 27(a) applies, Seshadri et al. calculated the electron concentration at the position of each of the quantum wells. Using PL, they determined the interdiffusion of each QW, and they found no correlation between D and the calculated carrier concentration. Thus, they concluded that there was no Fermi level effect upon diffusion.

Unfortunately, this conclusion is as valid as the assumptions underlying it. In particular, modeling the undoped layer as a 1 μm wide insulating space charge region may be reasonable at low temperatures, but it does not apply at the annealing temperature. We estimate that a Debye length, $L_D = (\varepsilon kT/q^2n)^{1/2} \approx 100$ Å, was associated with their n- and p-type layers. In such a case, all of the potential drop occurs very close to the doped regions, and the correct energy band diagram corresponding to this experiment is similar to that shown in Fig. 27(b). From this figure, it should be clear that $n \approx \text{constant}$ throughout the layer which contains the QWs, and that this is the reason why D_{III} exhibited negligible variation between the wells.

5.1.3.3. *The singly charged Ga vacancy model*

Not long after reporting the interdiffusion results for Si-doped SLs, Mei et al. [87] reported the results of interdiffusion in Te-doped SLs. Strikingly consistent results were measured for data taken over $T = 800$ –1000 °C, in 50 °C increments. Fig. 28 shows a much better consistency between the slopes obtained at the different annealing temperatures than is shown in Fig. 26, and these slopes appear to vary over the relatively narrow range 1.0–1.5. It is unknown at present how much of the difference may be related to the use of Te rather than Si, which may have moved interstitially and driven diffusion by a kick-out mechanism. However, it is worth noting that these experiments were performed much more carefully than the Si-doped series. In particular: (1) In the control sample used in this work, the Al concentration profile measured by SIMS varied by more than one order of magnitude in each of the doped layers, i.e. it was significantly greater than that of the Si-doped work. (2) The

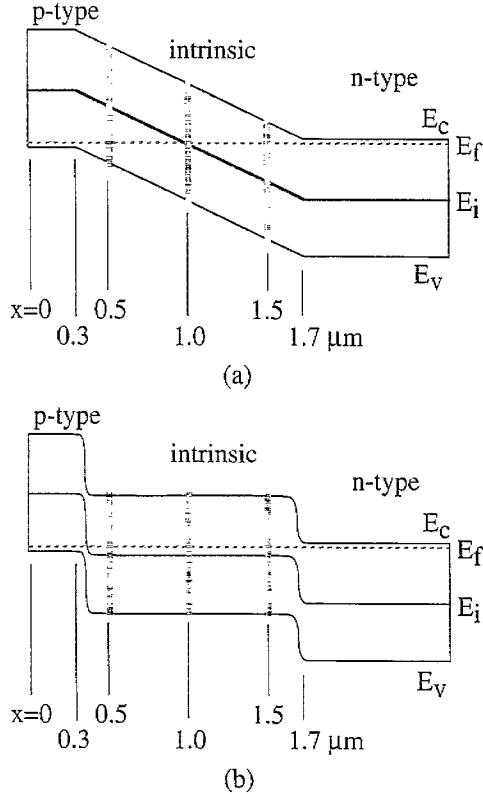


Fig. 27. (a) Schematic diagram of the energy band diagram implicitly assumed by the calculations of Seshadri et al. for the carrier concentration during their anneals. Three quantum wells had been grown at depths of 0.5, 1.0, and 1.5 μm , and they were assumed to be in a highly insulating region at the anneal temperature of $T=888$ °C. This assumption leads one to calculate a very different carrier concentration in each well. (b) Schematic diagram of the correct energy band diagram which applies at an anneal temperature of $T=888$ °C. The extremely small Debye length of $L_D \sim 100$ Å actually leaves the three quantum wells in a region with the same $n = n_i$, and this explains why the diffusivity of each well was found to be the same.

characteristic diffusion lengths associated with the lowest measured values of D were more than an order of magnitude larger than those reported for the Si-doped SLs, and appear to be of reasonable values. (3) A 2000 Å thick GaAs layer covered the SL this time, and made it less likely that oxides of Al may have affected with the diffusion process. Unfortunately, no AlAs spikes were included in the undoped layers for comparison with those in the adjacent doped layers. Given the higher anneal temperatures used, it is much more likely that the native defect concentrations in these samples approached their equilibrium values. Thus, the approximate relation, $D_{\text{III}} \propto n$, obtained from this series of experiments, appears to be a far more reliable description of interdiffusion than the $D_{\text{III}} \propto n^3$ deduced from the prior work of this same group at $T=700$ °C.

Cohen et al. [88] reported results of In interdiffusion at $T=900$ °C in uniformly Te-doped GaAs epitaxial layers, which had been grown by OMVPE at either $T_g = 600$ °C or 700 °C. An essentially linear $D_{\text{III}} \propto n$ dependence was measured. No dependence upon T_g was observed and this was taken to be evidence that the native defect concentrations had approached equilibrium. Li et al. [89] later reported the interdiffusion of In markers in GaAs epitaxial layers grown at $T_g = 650$ °C, but with various growth rates of $r_g = 0.5\text{--}6$ $\mu\text{m h}^{-1}$, in an effort to grow different native defect concentrations into the crystal, similar to those reported in Ref. [42]. No effect of the growth rate was observed after annealing, and the results agreed with the previously reported $D_{\text{III}} \propto n$. Li et al. [89] also grew In markers into npn, pnp, and n-i-n structures, and also obtained $D_{\text{III}} \propto n$ at $T=900$ °C. Except for

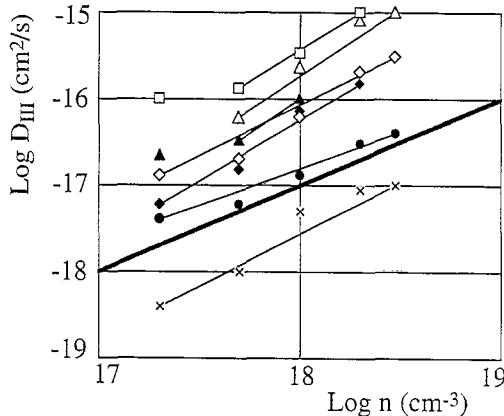


Fig. 28. D_{III} vs. n measured at different temperatures, summarized from the Te-doped interdiffusion study of Mei et al. [87]. Measurement temperatures include $T = 800$ °C (\times), 850 °C (circles), 900 °C (diamonds), 950 °C (triangles), and 1000 °C (squares). Filled symbols represent data from $t = 3$ h anneals. Open symbols indicate data from $t = 0.5$ h anneals. At $T = 800$ °C, $t = 30$ h anneals were used. The highest and lowest measurements at each temperature, except for 950 and 1000 °C, have been connected to qualitatively indicate the slope. At the two highest temperatures, $n_i > 2 \times 10^{17} \text{ cm}^{-3}$, and thus $n = 5 \times 10^{17} \text{ cm}^{-3}$ was used as the lowest point for the lines. The slopes shown vary between about 1.0 and 1.5, and can easily be compared with the heavy shaded line illustrating a slope of 1. The consistency between the data taken across the entire range of temperature is clearly better than that seen in the early Si-doped study reported by this same group, which was shown in Fig. 26.

interdiffusion markers placed exactly at the substrate–epilayer interface, all of the interdiffusion results of Li et al. were found to be consistent with both the Si- and Te-doped results of Mei et al. at $T = 900$ °C, after all were adjusted to the same standard P_{As} . Fig. 29 summarizes the measurements of Li et al. (filled diamonds), the Si-doped (filled circles) and Te-doped (open circles) SL's of Mei et al., and the undoped measurement of Tan et al. (\times). The nominally undoped samples are shown at $n = n_i(900$ °C) $\approx 2 \times 10^{17} \text{ cm}^{-3}$. The Tan result was obtained after annealing at $P_{As} \approx 1$ atm, and is the only result shown in Fig. 29 from a closed ampoule. The Mei data was obtained under $P_{As_4} < 10^{-8}$ atm, i.e. at the edge of the Ga-rich solidus, and the Li data was obtained under $P_{As_3} \approx 10^{-3}$ atm. All of these data, after adjustment to $P_{As} = 1$ atm, are shown by the same symbol (open squares). A line with a slope of unity is shown passing through the open squares, and it provides a reasonably good fit to the adjusted data. In the undoped crystal, $n = n_i \approx 2 \times 10^{17} \text{ cm}^{-3}$, and it appears that a standard diffusivity

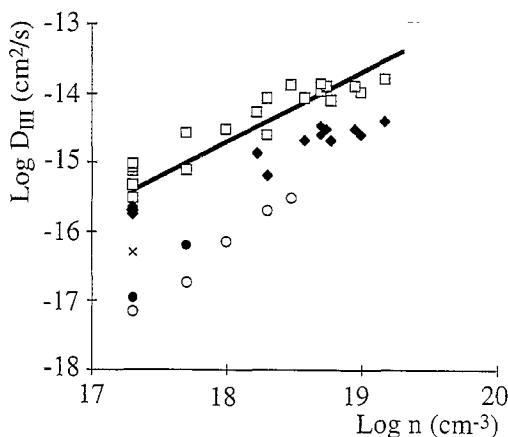


Fig. 29. Measured D_{III} vs. n , at $T = 900$ °C as reported by Li et al. [89] (filled diamonds), Mei et al. (circles), and Tan et al. (\times). After adjustment (open squares) to $P_{As} = 1$ atm, the data fall closely to the line shown, which has a slope of unity. Adapted from Ref. [89].

$D(T, P_{\text{As}} = 1, n_i) \approx 3-4 \times 10^{-16} \text{ cm}^2 \text{ s}^{-1}$ can be used to describe all of the data (except for the single Tan point) to within a factor of approximately 2. This value of D is also in excellent agreement with the $P_{\text{As}} = 1 \text{ atm}$ data of Olmsted et al. associated with the $T = 900 \text{ }^\circ\text{C}$ line of Fig. 22.

The good consistency between samples grown over $T_g = 500-700 \text{ }^\circ\text{C}$ by MBE and OMVPE, which were grown over a wide range of growth rates, and which were annealed over a wide range of P_{As} is important. It shows that the grown-in native defect concentrations, expected to be dissimilar because of the different growth conditions, have approached essentially the same, presumably equilibrium, concentrations at $T = 900 \text{ }^\circ\text{C}$. With this minimum level of reproducibility, it then makes sense to attempt a quasi-equilibrium interpretation. The line shown in Fig. 29 appears to be a reasonably good description of the relationship between D and n , i.e. $D \propto n$. Based upon Section 3.4, Section 5.1.2, and the data above, it appears that the most reliable experimental evidence to date for n-type and intrinsic GaAs, taken under conditions approximating native defect equilibration, are consistent with a model of diffusion controlled by the singly charged $\text{V}_{\text{Ga}}^{-1}$.

It may be worth noting that this conclusion does not contradict the Baraff and Schluter calculations, which predicted that the Ga vacancy with the highest concentration in n-type GaAs would be the $\text{V}_{\text{Ga}}^{-3}$. That prediction could be correct, but the standard diffusivity for the $\text{V}_{\text{Ga}}^{-3}$ may be so low that the defect never plays a measurable role in diffusion for any practical carrier concentration. Indeed, a multiply charged vacancy would presumably be highly polar and this property would presumably affect its diffusivity. We shall leave this topic for others to pursue.

5.1.4. Evidence for the diffusion mechanism in p-type GaAs

It is useful to consider how one might perform an interdiffusion experiment in p-type GaAs, and maintain uniform native defect concentrations close to their equilibrium values. We must choose between two very different kinds of dopant when growing a p-type epilayer: either a group II acceptor which resides primarily on the group III sublattice, or a group IV acceptor which resides primarily on the group V sublattice. In principle, either one will work. However, Section 5.2 will show that a nonequilibrium interstitialcy mechanism often determines the diffusion when group II elements are present. Several of the group IV atoms are amphoteric, i.e. they may switch between sublattices and act as either donors or acceptors. If the fraction of these atoms residing on the group III sublattice were to increase, then the crystal must temporarily accommodate an increased Ga interstitial concentration, and this will complicate any simple interpretation of diffusion. However, the dopant C is believed to remain on the group V sublattice over a very wide range of ambient conditions, and thus it is unlikely to create a nonequilibrium concentration of Ga interstitials which affect group III interdiffusion. Thus, C is the most reasonable dopant to choose when an equilibrium native defect concentration is desired. The extent to which C actually succeeds in satisfying the sufficient condition for native defect equilibration must, of course, be determined in practice.

5.1.4.1. Isoconcentration interdiffusion when Fermi energy is set by C

There have been relatively few analyzable reports of interdiffusion in the presence of C, presumably because the low rates of diffusion encountered are difficult to measure. Two groups have conducted experiments under nominally well-defined equilibrium conditions and the measured values of D_{III} were lower than those measured in the intrinsic crystal under otherwise similar ambient conditions. Using strong nonequilibrium conditions (Ga-rich liquid metal present), another group reported an increased D_{III} .

Li et al. [89] used In markers in GaAs to measure D_{III} at $T = 900 \text{ }^\circ\text{C}$ after $t = 1 \text{ h}$ anneals with As and H_2 in an open system. D_{III} was measured in each layer of a pnp structure in which C was used to set $p \approx N_C \approx 1 \times 10^{19} \text{ cm}^{-3}$, i.e. $p/n_i \approx 50$. Fig. 30 shows the In concentration profile measured by

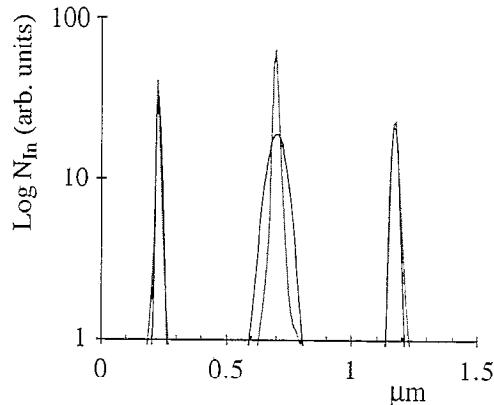


Fig. 30. N_{In} vs. x , measured by SIMS, in the pnp structure of Li et al. [89]. The profile prior to annealing is shown by the lighter curve, and the profile after annealing is shown by the heavier curve. Adapted from Ref. [89].

SIMS in this pnp structure before and after annealing. The diffusion of In is significantly greater in the n-type layer than in the p-type layers. Li et al. calculated $D_{III} < 3.7 \times 10^{-17} \text{ cm}^2 \text{ s}^{-1}$ in the top C-doped layer (or at least $100 \times$ lower than in the n-type region), and determined that D_{III} in the bottom C-doped layer could not be estimated.

The result for D_{III} in this top p-type layer is included in Fig. 31, which shows the measured D_{III} versus n/n_i for all of the $T = 900^\circ\text{C}$ experiments of Li et al., i.e. without any adjustment made for P_{As} . Fig. 31 shows that group III interdiffusion in C-doped GaAs is qualitatively consistent with diffusion via the V_{Ga}^{-1} inferred from the experiments discussed in Section 5.1.3. The lack of a better fit to the slope = 1 line is presumably caused by the much longer time required for the Ga vacancies to approach their very low equilibrium value in the C-doped layer. It is also possible that neutral vacancies control the diffusion in highly p-type GaAs, and no Fermi level dependence would be expected in this case. Further experiments need to be made to determine if this occurs. However, the lack of any increase in D_{III} at high values of p clearly shows that an interstitialcy mechanism does not drive the diffusion under conditions approximating native defect equilibrium.

Fig. 32 shows the measured D_{III} reported in C-doped GaAs. Szafranek et. al [90] measured the PL energy shift of quantum wells after annealing at $T = 825^\circ\text{C}$. For $N_C \approx 8 \times 10^{18} \text{ cm}^{-3}$, they estimated $D_{III} \approx 3.3 \times 10^{-18} \text{ cm}^2 \text{ s}^{-1}$ when using an As overpressure in an ampoule (diamond). A smaller D_{III} was obtained when diffusing at the Ga-rich edge of the GaAs solidus, but the diffusion was considered

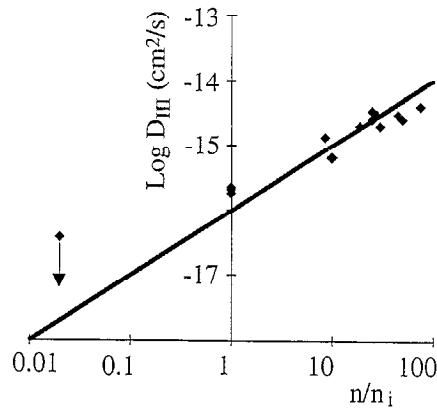


Fig. 31. Plot of measured D_{III} vs. n/n_i , taken from data of Li et al. [89] for $T = 900^\circ\text{C}$ without adjustment for P_{As} . D_{III} in the p-type GaAs appears to be consistent with the V_{Ga}^{-1} defect which also controls D_{III} in the n-type GaAs. Adapted from Ref. [89].

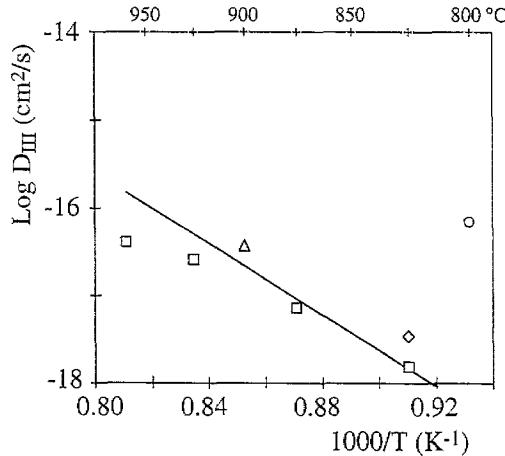


Fig. 32. Measured D_{III} vs. $1/T$ for heavily C-doped GaAs, without any adjustments made for P_{As} or for p . Data taken from Li et al. [89] (triangle), Szafranek et al. [90] (diamond), You et al. [91] (squares), and Chen et al. [66] (circle). The Wang line, for isotopic self-diffusion, is provided for comparison. The relatively good fit of the C-doped data with the interdiffusion line for undoped isotopic interdiffusion suggests that the native defects in the C-doped layers did not equilibrate during the time of the anneals.

too slow to be reliably estimated. Using a quartz ampoule, You et al. [91] reported D_{III} (squares) at several temperatures for heavily doped, $N_C \approx 3 \times 10^{19} \text{ cm}^{-3}$, GaAs at $P_{As} \approx 1 \text{ atm}$. In a nonequilibrium Ga-rich environment, they measured a large increase in D_{III} (discussed in Section 5.3). The result of Li et al. (triangle) at $T = 900 \text{ }^\circ\text{C}$ and $N_C \approx 1 \times 10^{19} \text{ cm}^{-3}$ has already been discussed above. The D_{III} values of Chen et al. [66] were measured after a $T = 800 \text{ }^\circ\text{C}$ anneal in $N_C \approx 6 \times 10^{19} \text{ cm}^{-3}$ C-doped layers, appear to be unusually large (circle). Presumably, this indicates that the vacancy concentration has not equilibrated in the $t = 1 \text{ h}$ anneal time used. For comparison, the line from Wang et al., which describes self-diffusion in intrinsic GaAs at $P_{As} = 1 \text{ atm}$, is similar to most of this data. This probably indicates that the native defects in the C-doped layer have not equilibrated. An additional comparison of the Wang line with different classes of experiments will be summarized in Section 5.3.

Perhaps the most striking feature of interdiffusion in C-doped GaAs is the reduction of D_{III} . Since interstitial based diffusion mechanisms appear to be excluded when C is used as the dopant, one must conclude that the vacancy mechanism also controls diffusion in p-type GaAs and that the total vacancy concentration falls as the Fermi level is lowered. Further measurements need to be made at higher C concentrations, longer anneal times, or higher temperatures, to determine additional reliable values for the interdiffusion in C-doped GaAs.

5.2. Nonequilibrium results

5.2.1. Design of nonequilibrium experiments

There are many experimental designs in which the native defect concentrations can only be described as nonequilibrium. Some of these are worth considering because they can be used to distinguish between diffusion mechanisms. Interdiffusion on the group III sublattice appears to be strongly affected by group IV donors or group II acceptors. It appears likely that group IV donors exhibit amphoteric behavior and may transport interstitially under some conditions. An analysis of such group IV donor behavior would unduly complicate this paper, and will not be presented here. We shall confine ourselves to analyzing the much simpler nonequilibrium results associated with the presence of Zn, and by implication, other group II elements.

5.2.1.1. Strong nonequilibrium

The term strong nonequilibrium will be used to refer to a system in which the number of condensed phases changes during annealing, or a substantial mass transport between existing condensed phases occurs. In general, a sample annealed under strong nonequilibrium conditions will not be subject to a simple, useful, interpretation. However, 'nonequilibrium space' is quite large, and sometimes shifting the system between two very different points in the ternary phase diagram can be quite informative. It is worth discussing two examples of experiments performed under strong nonequilibrium.

5.2.1.1.1. Dopant isoconcentration

The dopant isoconcentration technique was once intended to provide quasi-equilibrium information about Zn self-diffusion in GaAs. In this technique, uniformly Zn-doped GaAs is placed in a closed ampoule with a radioactive isotope of Zn, i.e. ^{65}Zn . Since the diffusion of the radioactive species will occur even without a chemical concentration gradient, some fundamental information about the dopant diffusivity can be obtained. However, one still must be concerned about the necessary and sufficient conditions needed to approximate native defect equilibrium. Unfortunately, the isoconcentration results of Ting et al. [92] and Chang et al. [93] were obtained in the presence of a liquid Zn–Ga melt without any added As in the ampoule, and degradation of the samples was severe. For the same nominal value of N_{Zn} , the measured D_{Zn} was nearly constant over $T = 600\text{--}1000\text{ }^{\circ}\text{C}$, and this result should suffice as evidence that something was drastically wrong. Since the equilibrium activities of Ga and As in the Zn-rich liquid, and the time required to reach those activities, must have differed at each temperature, such nonequilibrium results are not amenable to an interpretation within the equilibrium approximation. Some of the work of Kadhim et al. [94] was conducted under similar conditions, and it is not easily interpreted either. However, some results of Kadhim et al. were obtained with As added to the ampoule. Unfortunately, the characterization of the dopant concentration in the starting wafer was very poor, and it appears that no care was taken to match the activity of the Zn in the solid with that in the source. Given the poor initial characterization, we consider it unwise, if not impossible, to interpret these results within any simple model. Of course, parameterized models may be quite useful for engineering simulations, and we note that Tan et al. [95] published a quasi-equilibrium group II diffusion model which claimed to fit the above results. Evaluation of that model is beyond the scope of this paper.

5.2.1.1.2. Shift of final equilibrium state

A second kind of strong nonequilibrium experiment, that of shifting the system composition between two different regions of the phase diagram, was effectively used by Kendall [3] even before the phase diagram was known for the Ga–As–Zn ternary system. In a 10 min. anneal in a closed ampoule, Zn was diffused into GaAs at $T = 900\text{ }^{\circ}\text{C}$ to form a pn junction a few microns below the surface. When this GaAs was placed in another ampoule without Zn present and annealed at the same temperature for an additional 15 h, no further diffusion of Zn could be measured. There is no known means for the lack of a Zn source to cause either the vacancy concentration or the vacancy diffusivity to drop precipitously. Thus, this simple pair of nonequilibrium anneals eliminated the vacancy mechanism from consideration as the Zn diffusion mechanism and provided the first experimental evidence that Zn must move interstitially before becoming substitutional. In addition, the lack of Zn diffusion during the second anneal demonstrates that substitutional Zn does not become interstitial simply

because of the temperature, i.e. another point defect must be present to make it become interstitial before Zn can diffuse further.

5.2.1.2. Weak nonequilibrium: perturbed equilibrium

Section 3.2 showed how equilibrium for the crystal is defined in terms of the vapor (or other external phase) which acts as a reservoir of atoms. We consider weak nonequilibrium to refer to a modest change in the reservoir, i.e. a sufficiently small change in one or more degrees of freedom such that no large transport of mass between the crystal and reservoir occurs, and no new condensed phases appear. When such a change is made in the system, the final equilibrium state of the system is well defined, and the system will move toward that final state during the experiment (which it may or may not reach). If the concentration of all of the native defects changed at the same rate, then this would be a useless approach. However, different defect concentrations are expected to respond at a different rate to a change in ambient, and a well designed experiment inducing a significant change in one defect concentration during the course of an experiment may provide a probe of the diffusion mechanism.

It is worth discussing two proposed diffusion models, the I-S and the interstitialcy, in terms of one experiment which attempts to simply perturb the equilibrium state. The experiment we shall consider is a common one: measure D_{III} and D_{Zn} as a function of P_{Zn} (at a given T and P_{As}). The commonly observed result is that D_{III} and D_{Zn} are linked, i.e. they increase or decrease together as P_{Zn} is raised or lowered.

5.2.1.2.1. I-S model

The I-S model appeared after Kendall showed that vacancies did not transport Zn in GaAs. The model postulates that Zn atoms diffuse interstitially through the crystal until they encounter a Ga vacancy and become substitutional. Although the concept of Zn interstitial transport appears to have originated with Longini [96], it was Weisberg and Blanc [38] who provided the first logical evidence for it, i.e. a mathematical framework connecting the concept to the non-Fickian concentration profiles associated with Zn. Casey and Pearson [2] used an equilibrium mass-action expression relating the concentrations of I_{Zn} and V_{Ga} to show that the I-S mechanism was qualitatively consistent with the observed Zn concentration profile and effective Zn diffusivity. These qualitative features kept the model alive for many years. Of course, any model based upon an equilibrium mass-action expression relating interstitial and substitutional Zn (for example, the kick-out mechanism discussed in the next paragraph) will show exactly the same consistency with the same measurements. Such attempts to infer kinetic models from equilibrium arguments are fundamentally flawed.

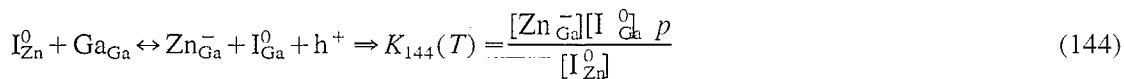
However, when the epilayer uses an isoelectronic marker, i.e. Al, to provide an additional degree of freedom in a slightly nonequilibrium experiment (i.e. the Al concentration profile is not flat), then the dependence of the marker diffusion upon P_{Zn} may help distinguish between atomistic models. For example, if the Zn diffused by an I-S mechanism and the marker moved by a vacancy mechanism as previously expected, then a reduction in D_{III} should be observed with increasing P_{Zn} , because the vacancy concentration should decrease as vacancies become filled by interstitial Zn. However, both D_{III} and D_{Zn} are observed to increase rapidly with increased P_{Zn} , and this suggests that the diffusion of Zn and the marker are intimately linked, i.e. the same mechanism applies to both. Kendall's conclusion that Zn does not diffuse via vacancy mechanism and the conclusion of Section 5.1.4 that the vacancy concentration is very low in p-type GaAs both imply that the observed linkage between D_{III} and D_{Zn} is a result of interstitial atoms. Since the I-S model predicts the wrong behavior, it is

reasonable to suspect that the interstitialcy mechanism controls both the Zn diffusion and the group III interdiffusion.

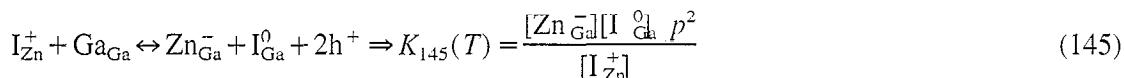
5.2.1.2.2. Interstitialcy model

Experimental results discussed below provide evidence that a nonequilibrium interstitialcy mechanism applies to the in-diffusion of Zn from the vapor, the out-diffusion of Zn to the vapor, and the diffusion of Zn within buried layers in the crystal, i.e. the so-called emitter push effect. Thus, we shall begin by briefly discussing key features of the interstitialcy model.

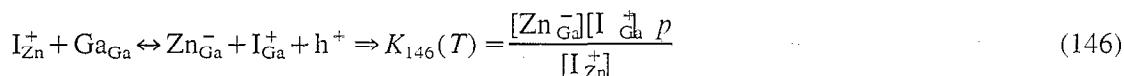
It should be clear from the results discussed in Section 5.1 that all evidence to date indicates that the *interstitialcy plays no role* in a crystal where the native defect concentrations are close to their equilibrium values. However, when appropriate perturbations are made in those equilibrium values, the interstitialcy appears to apply. We shall begin by assuming that Zn is the only dopant present and that $p = N_{\text{Zn}} = [\text{Zn}_{\text{Ga}}]$ is close to its equilibrium value as set by a choice of T , P_{As} , and P_{Zn} . For simplicity, we may consider a reaction involving neutral interstitial atoms,



which predicts a strong dependence for the equilibrium neutral interstitial dopant concentration upon hole concentration, i.e. $[\text{I}_{\text{Zn}}^0] \propto p^2$. It is worth emphasizing that the solubility of the dopant interstitial depends upon its charge state, as can be appreciated by considering a similar reaction for I_{Zn}^+ ,



which predicts an equilibrium relationship, $[\text{I}_{\text{Zn}}^+] \propto p^3$. We note that incorrect models have sometimes been based upon the erroneous assumption that the charge state of the I_{Ga} affects the result. For example, the reaction



superficially appears to predict that $[\text{I}_{\text{Zn}}^+] \propto p^2$. However, it actually predicts that $[\text{I}_{\text{Zn}}^+] \propto p^2[\text{I}_{\text{Ga}}^+]$, and since Eq. (96) showed that $[\text{I}_{\text{Ga}}^+] \propto p[\text{I}_{\text{Ga}}^{0+}]$, it is clear that the correct *equilibrium* relationship remains $[\text{I}_{\text{Zn}}^+] \propto p^3$. In addition, Section 2.2.3 showed that $[\text{I}_{\text{Ga}}^0] \propto P_{\text{Ga}} \propto P_{\text{As}_4}^{-1/4}$, and thus at a given temperature, the defect solubility depends upon the two degrees of freedom, p and P_{As_4} ,

$$[\text{I}_{\text{Zn}}^+] \propto p^3 P_{\text{As}_4}^{-1/4} \quad (147)$$

In practice, however, p is a dependent, rather than an independent, variable. From Section 3.2.1, the equilibrium value of p has been defined in terms of both P_{Zn} and $P_{\text{As}_4}^{1/4}$, and we may thus write an *equivalent* expression describing the equilibrium defect concentration in terms of those two degrees of freedom at the chosen temperature,

$$[\text{I}_{\text{Zn}}^+] \propto P_{\text{Zn}}^{3/2} P_{\text{As}_4}^{1/8} \quad (148)$$

The (apparent) change in the P_{As} dependence occurs because p depends upon the chosen P_{As} as well as P_{Zn} , and there is no conflict between the two descriptions when solid-vapor equilibrium holds. However, it is reasonable to ask what happens when equilibrium is perturbed.

5.2.1.2.3. Interstitial equilibration with vapor versus solid

It is often assumed that the concentration of I_{Zn} equilibrates with the substitutional Zn concentration, independent of the ambient partial pressures chosen for an anneal. However, in Section 5.2.2, it will become clear that Eq. (148) generally is a more appropriate description of $[I_{Zn}]$ than Eq. (147), and that this appears to explain the very different diffusion results associated with (1) Zn in-diffusion from the vapor, (2) Zn out-diffusion from the crystal, and (3) Zn diffusion from a buried layer.

5.2.2. Analysis of selected nonequilibrium experiments

There is a wealth of published data which have been taken under nonequilibrium conditions. However, few of these studies were designed with the intention of perturbing a specific equilibrium state. In this subsection, we shall discuss a relatively small number of experiments which were designed to bring together a few key features of diffusion in a logical and hopefully interpretable manner.

5.2.2.1. Weak nonequilibrium—group III impurity

In Fig. 19, we showed the concentration profiles for an as-grown pnpn structure which contained the dopants C and Te, and the isoelectronic marker, In. Very little diffusion is associated with this structure during growth. Based upon the results presented in Section 5.1, one expects the C-doped top layer to be the first to equilibrate with an external annealing ambient. Fig. 33 shows the concentration profiles [22] measured by SIMS for this same structure after a $t = 1$ h open tube anneal at $T = 800$ °C, using a GaAs proximity cap and a controlled P_{As} . In the top C-doped layer, $D_{III} \sim 10^{-16} \text{ cm}^2 \text{ s}^{-1}$ was measured. If the native defect concentrations in this layer had equilibrated quickly with the ambient vapor, then it is likely that $D_{III} \sim 10^{-18} \text{ cm}^2 \text{ s}^{-1}$ would have been measured (based upon the extrapolation from higher temperature shown in Fig. 32). This appears to indicate that quasi-equilibrium native defect concentrations in the layer were not rapidly approached for the time and temperature used. We also note that the interdiffusion of In markers in the buried C-doped layers was too low to be reliably determined, i.e. $D_{III} < 2 \times 10^{-17} \text{ cm}^2 \text{ s}^{-1}$.

5.2.2.2. Weak nonequilibrium—group II impurity

Another piece of the as-grown sample shown in Fig. 19 was annealed at $T = 800$ °C using the same P_{As} and anneal time as the sample discussed above, but also with a large P_{Zn} added to the ambient. Fig. 34 shows the concentration profiles measured by SIMS after this anneal [66]. D_{III} was found to be unaffected by the presence of the Zn. The unchanged and low value of D_{III} shows that the vacancy concentration has not changed appreciably, and confirms that Zn is not transported by vacancies.

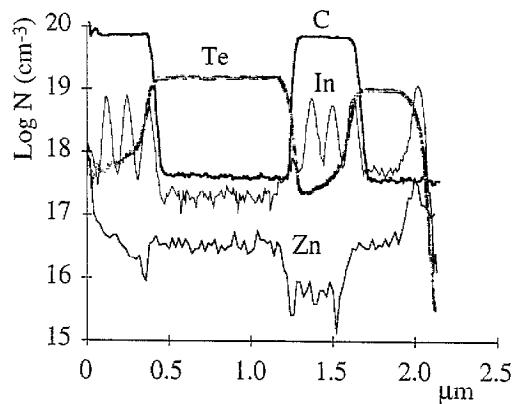


Fig. 33. Pnpn structure of C and Te, measured by SIMS, after annealing the as-grown structure for $t = 1$ h at $T = 800$ °C under P_{As} . The as-grown profiles were shown in Fig. 19, and little change is observed. Adapted from Ref. [22].

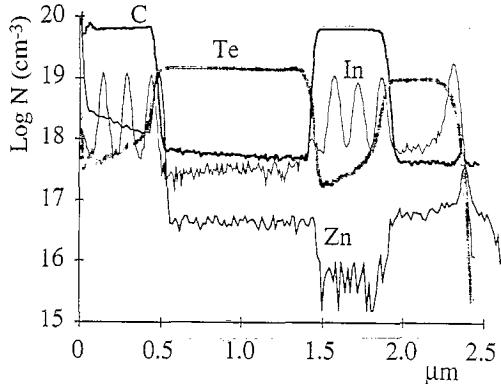


Fig. 34. Another piece of the structure shown in Fig. 19 after annealing for $t = 1$ h at $T = 800$ °C under both P_{As} and P_{Zn} . The Zn has moved into the top p-type layer extremely rapidly, but N_{Zn} remains well below the value expected if C were not present. The lack of any Zn diffusion into the Te-doped layer is direct evidence that the Zn interstitials are positively charged and cannot readily diffuse into the n-type layer. Adapted from Ref. [66].

The apparent diffusivity associated with the Zn profile in the top C-doped layer of Fig. 34, $D_{\text{Zn}} > 10^{-13}$ cm² s⁻¹, is unusually high to be associated with the relatively low Zn substitutional concentration, $N_{\text{Zn}} \sim 2-3 \times 10^{18}$ cm⁻³. Of course, rapid transport of Zn is expected to be associated with a high Zn interstitial concentration set by the four degrees of freedom chosen: T , P_{As} , P_{Zn} , and the large value of p ($= N_C$), as discussed in Section 3.2.1.

5.2.2.3. Solubility of minor impurity

If C had not been present in the crystal, it is known from separate work that the Zn concentration would have been $N_{\text{Zn},1} = p_1 \approx 1.5 \times 10^{19}$ cm⁻³ for the temperature and partial pressures used. However, when the actual hole concentration, p_2 , is larger than this value, then Eq. (92) has shown that the concentration of the minor impurity, $N_{\text{Zn},2}$, will be reduced by ratio of these hole concentrations, i.e.

$$N_{\text{Zn},2}\{T, P_{\text{As}}, P_{\text{Zn}}, p_2\} = N_{\text{Zn},1}\{T, P_{\text{As}}, P_{\text{Zn}}\} \frac{p_1}{p_2} \quad (149)$$

In this case, $p_2 \approx 6.5 \times 10^{19}$ cm⁻³, $p_1/p_2 = 0.23$, and the predicted solubility for the substitutional Zn in the C-doped layer is $N_{\text{Zn},2} \approx 3.5 \times 10^{18}$ cm⁻³, which is very close to the measured value of $N_{\text{Zn}} \approx 3 \times 10^{18}$ cm⁻³ near the surface (which is expected to equilibrate first with the ambient vapor).

At the pn junction between the top C- and Te-doped layers, a large Zn spike is easily seen in Fig. 34. As discussed in Section 3.2.4, the change in potential across the junction causes the equilibrium solubility of any charged defect to change across the junction. The equilibrium substitutional Zn concentration was predicted to increase across a pn junction (formed using two dopants other than Zn) by Eq. (114),

$$[\text{Zn}_G]_n = [\text{Zn}_G]_p \exp \left\{ \frac{q \Delta \phi}{kT} \right\} \quad (150)$$

until the n-type crystal becomes compensated, and a large Zn spike can be observed in the top pn junction in Fig. 34. The Zn concentration spike is also seen to fall rapidly to the background Zn concentration in the Te-doped layer. This is also quite striking, and the lack of a gradual change shows that virtually no interstitial Zn gets beyond the pn junction. It is worth reflecting upon Eq. (113) which describes the equilibrium concentration of positive Zn interstitials, across the pn junction,

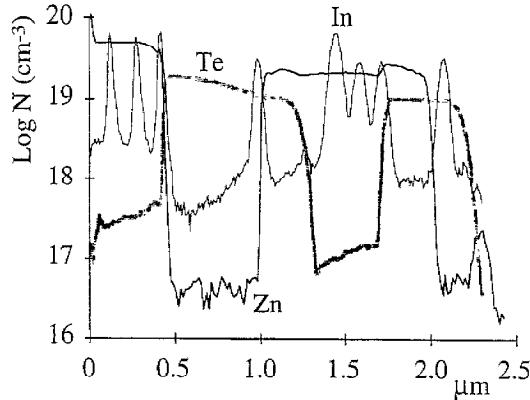


Fig. 35. SIMS profiles of the as-grown pnpn structure previously shown in Fig. 20. The figure is shown again for ease of comparison with the figures which follow. The emitter push effect, i.e. the widening of the buried Zn-doped layer, after growth of the top Te-doped layer is obvious. Adapted from Ref. [66].

$$[I_{Zn}^{\pm}]_n = [I_{Zn}^{\pm}]_p \frac{n_p}{n_n} = [I_{Zn}^{\pm}]_p \exp \left\{ -\frac{q \Delta \phi}{kT} \right\} \quad (151)$$

At the annealing temperature used, this expression predicts that the equilibrium concentration of singly charged I_{Zn}^{\pm} will drop by about 5 orders of magnitude when moving from the p-side to the n-side of the junction. If a significant fraction of the Zn transport took place via neutral I_{Zn} , then the flux of neutral interstitials would have been unaffected by the field of the junction, and a much more gradual drop in the substitutional Zn profile would be expected in the Te-doped layer. These results are direct experimental evidence that the moving Zn interstitial is positively charged, although the magnitude of that charge on the interstitial is not determined.

5.2.2.4. Isolation of emitter push effect

Similar pnpn structures were also grown and annealed [66] using Zn and Te as the dopants. For convenience, Fig. 35 repeats the concentration profiles for the as-grown structure shown in Fig. 20. In this structure, a very large emitter push effect is evident, i.e. an enormous D_{Zn} in the buried Zn-doped layer is associated with growth. It is worth noting again that the Zn source was not turned on until the first Te-doped layer was grown and that the system was purged after the first Zn-doped layer was grown. Thus, the buried Zn-doped layer would have looked like the top Zn-doped layer if diffusion had not taken place. An estimate of $D_{Zn} \sim 1 \times 10^{-13} \text{ cm}^2 \text{ s}^{-1}$ was associated with the substitutional Zn profile in the buried layer. The extremely sharp drop of the Zn concentration profile is consistent with a low solubility for positively charged Zn interstitials in n-type GaAs. As discussed in Section 4.2, the large values for these diffusivities are consistent with a large concentration of positive I_{Ga} in the buried Zn-doped layer, which apparently entered the crystal because of the pinned Fermi energy at the growing surface, and which drive Zn diffusion via an interstitialcy mechanism.

5.2.2.5. Multiple nonequilibrium states in one sample: Low P_{Zn}

Pieces of the as-grown sample were annealed under different conditions to study the variation of diffusion and solubility as the crystal moved toward different equilibrium states. Fig. 36 shows the concentration profiles after annealing at $T = 800 \text{ }^{\circ}\text{C}$ for 1 h under P_{As} , and with $P_{Zn} = 0$. The small rounding in the N_{Zn} concentration profile in the top Zn-doped layer shows that Zn moves from substitutional to interstitial sites relatively slowly even as the concentration of native defects, particularly I_{Ga} , approach their equilibrium values. Using an ambient $P_{Zn} = 0, N_{Zn} = 0$ is the equilibrium Zn

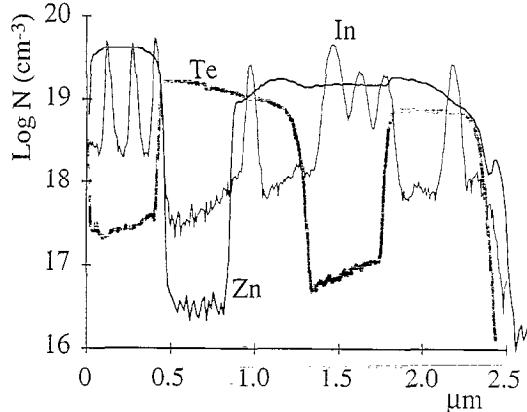
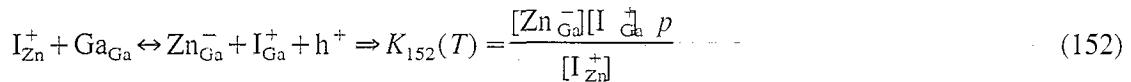


Fig. 36. SIMS profiles for the structure shown in Fig. 35, after annealing for $t = 1$ h at $T = 800$ °C under P_{As} . The widening of the Zn-doped buried layer continues unabated even though the Zn diffusivity out of the top layer is orders of magnitude lower. The native defect concentrations associated with these two p-type layers are clearly different, and those defects remain isolated from each other. Adapted from Ref. [66].

concentration and thus the grown-in Zn present in crystal is highly supersaturated relative to the ambient. In addition, other measurements [42] have shown that the grown-in $N_{Zn} = 5 \times 10^{19} \text{ cm}^{-3}$ exceeds the solubility limit for Zn, by a factor of 2–4× at $T = 800$ °C and the P_{As} used for these experiments. Thus, the thermodynamic driving force tending to reduce N_{Zn} is extremely large, but the rate at which the metastable value of N_{Zn} drops apparently reflects a low concentration of native defects. In the top layer, it is unclear how closely the native defects have approached their quasi-equilibrium concentrations which are defined, as shown in Section 3.2.5, by T , P_{As} , and the electrostatic potential, $\Delta\phi$, between the surface and p-type region. A large contrast exists between the high values of $D_{Zn} \approx 10^{-13} \text{ cm}^2 \text{ s}^{-1}$ and $D_{III} \approx 6 \times 10^{-16} \text{ cm}^2 \text{ s}^{-1}$ in the buried p-type layer, and the relatively small values of $D_{Zn} \approx 6 \times 10^{-16} \text{ cm}^2 \text{ s}^{-1}$ and $D_{III} \approx 3.7 \times 10^{-17} \text{ cm}^2 \text{ s}^{-1}$ in the top p-type layer. This anneal makes it strikingly clear that the highly nonequilibrium concentrations of point defects which were grown into the buried p-type region remain trapped there and continue to drive both the Zn and group III marker diffusion.

To explain the emitter push effect, Chen et al. [55,66] proposed that excess I_{Ga}^+ became incorporated in the buried Zn-doped layer during growth of the following n-type layer, as discussed in Section 4.2. This excess I_{Ga}^+ was proposed to drive the following (or a similar) reaction to the left,



and to produce excess I_{Zn}^+ in the neutral buried p-type region. As the excess I_{Zn}^+ diffused into a nearby pn space charge region, it was suggested that these I_{Zn}^+ would drive reaction Eq. (152) to the right (where p is low) and become substitutional and thus widen the buried p-type region. If the reaction produced positively charged I_{Ga}^+ in the junction space charge region, then the electric field would tend to confine the I_{Ga}^+ to the buried p-type layer. If this occurs, then rapid Zn diffusion out of the buried layer will continue for an extended time. If neutral Ga interstitials had been generated in the space charge region, then they would not be confined by the electric field, and D_{Zn} should have dropped quickly. Thus, it appears that there is a negligible concentration of mobile neutral interstitials present. An attempt to test further the appropriateness of the two-way interstitialcy mechanism for Zn was made by changing the ambient and inducing changes in the point defect concentrations in the top layer.

5.2.2.6. Multiple nonequilibrium states in one sample: High P_{Zn}

Fig. 37 shows the concentration profiles measured by SIMS for another piece of the as-grown sample shown in Fig. 35, after annealing at $T = 800^\circ\text{C}$ for 1 h under the same P_{As} , but also with a large P_{Zn} above the sample. The diffusivities of Zn and In in the buried layer remain essentially unchanged from those measured with $P_{\text{Zn}} = 0$, and thus the point defects in this layer appear to be unaffected by the ambient. However, in the top Zn-doped layer, very significant changes occur. The concentration of Zn has *dropped* by half, but only about 30% of those atoms have diffused into the Te-doped region and the rest have apparently diffused to the surface and left the crystal.

The only difference between the two anneals described above was the applied P_{Zn} . To reiterate, with no Zn present in the vapor, the diffusion in the heavily Zn-doped layer is slow. However, with Zn present in the vapor above a heavily Zn-doped layer, N_{Zn} drops and rapidly approaches its equilibrium value. In addition, there is a large $D_{\text{Zn}} > 10^{-14} \text{ cm}^2 \text{ s}^{-1}$ into the crystal, which appears to be associated with a tenfold increase in $D_{\text{III}} \approx 3.3 \times 10^{-16} \text{ cm}^2 \text{ s}^{-1}$. Since nothing would be changed if the in-diffusing Zn interstitials simply kicked out and replaced other substitutional Zn atoms, another interstitial must be involved. Because of the rapid diffusion of Zn interstitials, Chen et al. proposed that the Zn interstitial concentration tended to equilibrate with the ambient vapor, as described by Eq. (148) rather than with the dopant concentration, as described by Eq. (147). Thus, if I_{Zn}^+ flowed rapidly from the surface to the pn junction and became substitutional, then the concentration of I_{Ga}^+ in the junction would rise as reaction Eq. (152) moved to the right in the space charge region. Because the electric field of the junction would effectively halt diffusion of positively charged interstitials into the crystal, the excess concentration of I_{Ga}^+ would presumably equilibrate by diffusing back toward the surface. It was thus proposed that as this I_{Ga}^+ flowed back toward the surface, some of those interstitials would push out and replace the supersaturated substitutional Zn, i.e. driving reaction Eq. (152) in the neutral p-type region to the left. If this occurred, then N_{Zn} would fall to its equilibrium value at a rate which depended upon the concentration of the native defect, I_{Ga}^+ , and thus upon P_{Zn} . Thus, the single reaction, Eq. (152) appears to explain a wide range of diffusion results often associated with Zn and other group II dopants, i.e. in-diffusion, out-diffusion, and the emitter push effect.

The alternatives to the interstitialcy mechanism appear to involve rather imaginative mechanisms. For example, if the applied P_{Zn} somehow caused the grown-in concentration of substitutional Zn to leave the crystal and leave their substitutional sites vacant, then this would create $[\text{V}_{\text{Ga}}] > 10^{19} \text{ cm}^{-3}$

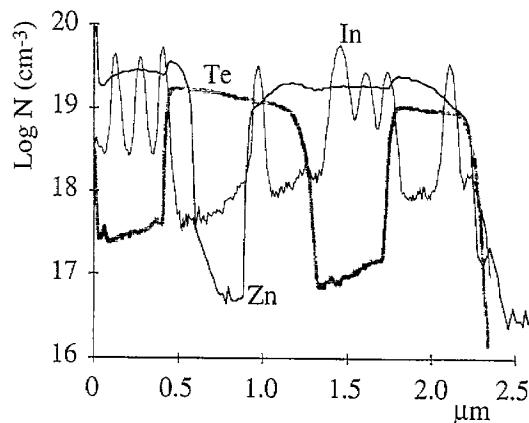


Fig. 37. SIMS profiles for another piece of the structure shown in Fig. 35, after annealing for $t = 1$ h at $T = 800^\circ\text{C}$ under both P_{As} and P_{Zn} . Zn diffusion out of the buried layer is unaffected by the application of P_{Zn} . However, Zn diffusivity in the top layer has increased by orders of magnitude when the high P_{Zn} present causes N_{Zn} to fall rapidly to its equilibrium value with this ambient. Adapted from Ref. [66].

and presumably cause a vast increase in both D_{Zn} and D_{III} . Since some increase in D_{III} was observed, Chen et al. annealed another piece of the as-grown sample under ambient conditions chosen to be close to equilibrium with N_{Zn} in the top layer, i.e. they eliminated the out-diffusion of Zn and thus halted any possible large production of V_{Ga} .

5.2.2.7. Simultaneous equilibrium and nonequilibrium in adjacent layers

Another piece of the as-grown sample was annealed at $T = 650^\circ\text{C}$ for 2 h using values of P_{As} and P_{Zn} which were known to correspond closely to equilibrium with the grown-in $N_{\text{Zn}} = 5 \times 10^{19} \text{ cm}^{-3}$. Thus, this anneal ambient keeps N_{Zn} constant in the top layer and suppresses any increase in the Ga vacancy concentration which might be caused by Zn out-diffusion. Fig. 38 shows the concentration profiles measured by SIMS after this anneal, and it is clear that both $D_{\text{Zn}} \approx 1 \times 10^{-13} \text{ cm}^2 \text{ s}^{-1}$ and $D_{\text{III}} \approx 1.5 \times 10^{-15} \text{ cm}^2 \text{ s}^{-1}$ in the top Zn-doped layer are *larger* than they were at $T = 800^\circ\text{C}$ even as N_{Zn} remains unchanged in the top layer. However, the total number of new Zn atoms entering the top Te-doped layer, $N'_{\text{Zn}} = 2 \times 10^{15} \text{ cm}^{-2}$, is approximately the same as the total number of Zn in the original top Zn-doped layer and these Zn atoms must have arrived as interstitials from the vapor. It is reasonable to expect that this same number of I_{Ga}^+ were displaced from the lattice by the newly substitutional Zn, and that most of this excess I_{Ga}^+ will have to flow back to the surface if the interstitial concentration is to approach its equilibrium value. Thus, the excess I_{Ga}^+ appear to increase D_{III} by driving an increased rate of exchange between group II and III interstitials and substitutionals. This model of a two-way interstitialcy is consistent with the much larger D_{III} and the constant N_{Zn} in the top layer. In contrast, if a large concentration of vacancies had remained from the growth, or had somehow entered the top layer during annealing, then a similar concentration of vacancies should also have been present when annealing with $P_{\text{Zn}} = 0$, i.e. significantly more diffusion should have been observed in the top layer of Fig. 36. Thus, any role for vacancy-driven Zn diffusion in these pnpn structures appears to be eliminated.

In addition to the interstitialcy mechanism, it appears likely that the interstitial concentrations near the surface of GaAs tend rapidly toward equilibrium with the ambient vapor via

$$\text{Zn}_{\text{vap}} \leftrightarrow \text{I}_{\text{Zn}}^+ + e^- \quad K_{153}(T) = \frac{[\text{I}_{\text{Zn}}^+]n}{P_{\text{Zn}}} \quad (153)$$

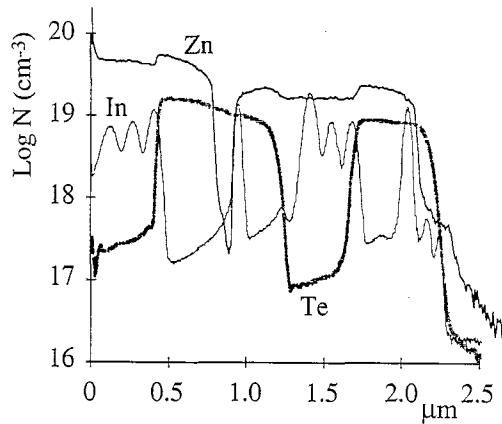


Fig. 38. SIMS profiles for another piece of the as-grown structure shown in Fig. 35, after annealing for $t = 2$ h at $T = 650^\circ\text{C}$ with the same P_{As} and P_{Zn} as those used for the sample shown in Fig. 37. Zn diffusion out of the buried layer is similar to that observed at $T = 800^\circ\text{C}$. However, Zn and In diffusion in the top layer have increased substantially even though the substitutional N_{Zn} in the top p-type layer remains close to equilibrium with the ambient. Adapted from Ref. [66].

and



When P_{Zn} is large, then Eq. (153) implies that Zn entering the crystal as an interstitial will drive reaction Eq. (152) to the right and produce a large Ga interstitial concentration wherever p is small, i.e. in a pn junction. The resulting excess Ga interstitial concentration returns to its equilibrium value as those interstitials diffuse to the surface. If N_{Zn} is above its equilibrium value near the surface, then it will quickly be reduced to its equilibrium value as Eq. (152) shifts to the left and interstitial Zn diffuses away, as seen in Fig. 37. When a large nonequilibrium concentration of Ga interstitials are produced this way, then a large increase in marker interdiffusion is expected, as seen in the top layer of Fig. 38. If P_{Zn} were chosen to be small, then no excess Ga interstitials will be produced in the pn junction and little diffusion is expected. This is why diffusion is very slow in the top layer of Fig. 36, i.e. it is apparently set by the approximately equilibrium Ga interstitial concentration described by Eq. (154).

It is important to appreciate that the different Zn-doped layers discussed above are tending toward entirely *different equilibrium states*, at the same temperature. In Figs. 37 and 38, the substitutional Zn concentration in the top Zn-doped layer is very close to its equilibrium concentration. As discussed in Section 3.2.3, Zn_3As_2 would begin to grow at similar temperatures if P_{Zn} were increased. Thus, the equilibrium solid composition of the GaAs is close to the Zn solid solubility limit defined by the edge of the GaAs solidus near line 1 in the simplified ternary phase diagram shown in Fig. 39. In striking contrast, the equilibrium state for the top layer in Fig. 36 is $N_{\text{Zn}} = 0$ because $P_{\text{Zn}} = 0$. This equilibrium state is near line 2 in Fig. 39 which relates the equilibrium state of the crystal to that of an As-rich ambient. However, the low native defect concentrations in this layer keep diffusion so slow that an impractically long time would be required for the layer to actually reach equilibrium. Thus, the actual crystal stoichiometry of the top layer in Fig. 36 always lies somewhere between its initial and final states. The final equilibrium state for the Zn-doped buried layers in all of the pnpn structures is ultimately determined by the vapor above the annealed sample, but once again a very long time would be required to reach those equilibrium values. For the relatively short times used for the actual anneals, the buried layers give every indication of being rich in both Zn and Ga interstitials. Thus, we approx-

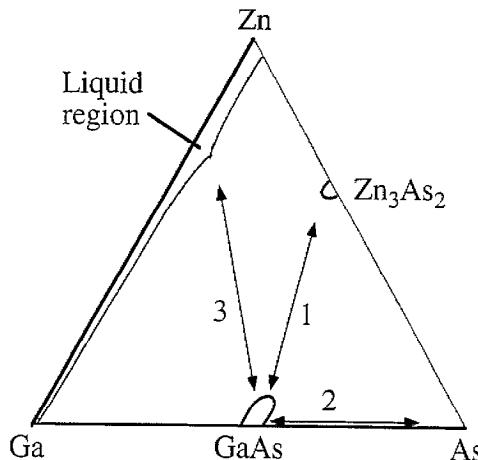


Fig. 39. Isothermal slice of the Ga-As-Zn ternary phase diagram estimated at 800 °C. The As-rich liquid regions are complex and not shown. As discussed in the text, the very different diffusion results observed in the Zn-doped regions in the pnpn structures can be understood by associating each region with one of three different quasi-equilibrium GaAs:Zn stoichiometries.

imate the stoichiometry of the buried Zn-doped layers as that of GaAs approximately in equilibrium with a Ga-Zn melt, i.e. with a composition roughly in the vicinity of line 3 in Fig. 39. At this point, it should be clear that nonequilibrium native defect concentrations may either *enhance or retard* diffusion, and that temperature is only one of the important variables. Other group II dopants in GaAs have not been as well characterized as Zn, but changes in their diffusion and solubility often mimic the Zn results discussed above, and we anticipate that very similar materials issues are involved.

5.2.2.8. Emitter push in C-doped base layer

When Li et al. compared D_{III} in their pnp (C-Te-C) structures at $T = 900 \text{ }^{\circ}\text{C}$, they measured extremely little diffusion in both p-type layers. Li et al. also measured D_{III} in an npn structure which had been annealed simultaneously with the pnp structure. The In concentration profile measured by SIMS is shown in Fig. 40. In the C-doped region of the npn, $D_{\text{III}} \approx 2 \times 10^{-16} \text{ cm}^2 \text{ s}^{-1}$, or almost an order of magnitude higher than obtained for the C-doped layers of the pnp structure discussed in Section 5.1.4. This enhancement in D_{III} appears to indicate that interstitial Ga trapped during growth may also increase D_{III} at a slow rate even in a C-doped base. A more detailed description of this effect awaits clarification in the future.

5.2.2.9. Discussion of nonequilibrium results

Controlled nonequilibrium experiments are useful for inferring diffusion mechanisms. The results discussed above are consistent with the conclusion of Kendall that Zn does not diffuse by a vacancy mechanism. The joint variation of D_{II} and D_{III} during Zn in-diffusion from the vapor argues against the I-S model. An interstitialcy model appears to provide a straightforward explanation of the very large variation of D_{II} and D_{III} observed during in-diffusion, out-diffusion, or diffusion inside of a buried p-type layer. The group II and III interstitial concentrations appear to more or less quickly approach equilibrium with the partial pressures above the surface. A useful model of the interstitialcy mechanism does not require pairing of point defects as some have suggested.

In Section 4.1, we showed that Fermi level pinning at a surface of an n-type layer caused a relatively high concentration of I_{Ga}^{+} to exist near that surface. We concluded that at sufficiently high growth rates, an excess concentration of I_{Ga}^{+} appeared in the growing n-type layer, and that some of these I_{Ga}^{+} diffused into the buried p-type layer where they drove the emitter push effect. An alternative

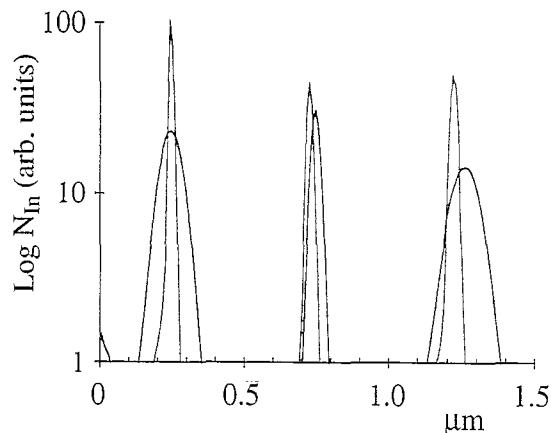


Fig. 40. N_{In} vs. x , measured by SIMS, for the npn structure of Li et al. [89]. The profile prior to annealing is shown by the lighter curve, and the profile after annealing is shown by the heavier curve. D_{III} is large enough to be easily measured in this C-doped layer and shows that a small emitter push effect (on the group III marker) can occur in a C-doped base. This is consistent with the trapping of Ga interstitials during the growth of the second n-type layer, as discussed in the text. Adapted from Ref. [89].

explanation for the interstitial production is a Frenkel reaction in the n-type layer, as suggested by Deppe [65]. However, if Ga interstitial–vacancy pairs could be readily produced by a Frenkel reaction during growth at $T = 650$ °C, then presumably even more pairs would be produced when annealing at $T = 800$ °C. If this occurred, then (1) half of any interstitials produced during annealing should flow into the top Zn-doped layer and increase Zn diffusion there, and (2) D_{III} should increase in the n-type layer because of the increased vacancy concentration. However, the low Zn diffusion in the top layer of Fig. 36, and the low D_{III} in all of the n-type layers of the pnpn structures, indicate that the Frenkel reaction does not rapidly produce interstitials in the n-type layer during annealing, and thus the Frenkel reaction is presumably not significant during growth either.

5.3. Summary of uncapped group III interdiffusion in GaAs

We have attempted above to relate interdiffusion to an equilibrium or a nonequilibrium state. It is also valuable to view effects on interdiffusion which may be related to the experimental technique. We shall begin by comparing values of D_{III} obtained from closed ampoule annealing under either As-rich or Ga-rich conditions which appear to correspond to a rapidly approached solid–vapor equilibrium. These results will be compared with D_{III} obtained from annealing under highly nonequilibrium conditions (liquid Ga intentionally added) in a closed ampoule. The results from these three classes of experiments will then be compared with D_{III} measured after open tube annealing. To minimize complexity of this section, we shall primarily discuss nominally undoped GaAs. Qualitatively, the results are similar to those expected for native defect equilibrium or perturbed equilibrium. However, some unexpected results will also be discussed.

5.3.1. Interdiffusion in a closed ampoule

We shall discuss interdiffusion data taken in a quartz ampoule using conditions which appear to be well-defined. For simplicity, the isotopic interdiffusion data will be left out, but both the Wang and the Tan lines will be provided in the interdiffusion plots for comparison. Anneal conditions will be separated into three categories: (1) As was added to the ampoule, (2) no As or Ga was added to the ampoule, or (3) Ga was added to the ampoule, all prior to annealing. We shall refer to these conditions as (1) ‘As-rich’, (2) ‘Ga-rich’, and (3) ‘strongly nonequilibrium’, respectively. For case (1), many groups have attempted to maintain approximately $P_{\text{As}} \approx 1$ atm. For the few groups who have added As, but gave insufficient information to allow an estimate of P_{As} , we shall assume that $P_{\text{As}} = 1$ atm so their data can be plotted with other data for As-rich GaAs. This will introduce some scatter into the plots, but this noise will be seen to be relatively minor compared with other sources of noise. For case (2), the sample must decompose to a small extent in the ampoule. When $T > 640$ °C, i.e. above the congruent sublimation temperature, the crystal approaches an equilibrium state defined at the Ga-rich edge of the GaAs solidus. If the native defects equilibrate quickly during these experiments, and if no unexpected chemical reactions occur, then these anneals are expected to provide well-defined and reproducible diffusion results for Ga-rich GaAs. For case (3), under the best of conditions, the GaAs sample will dissolve for a period of time into the liquid Ga melt. When the As composition in the melt eventually reaches its equilibrium value, i.e. at the edge of the Ga-rich liquidus, the melting will cease and the native defect concentrations in the GaAs will begin to approach the same well-defined values which were achieved much more quickly in case (2). However, as discussed in Section 2.1.1, this dissolution has been measured and it appears to be less than expected. Although the reasons for this have not been determined with certainty, this behavior appears to be caused by the presence of a contaminant. We consider it likely that contaminant is an oxidant originally present in either the quartz or the added condensed phase, and the result is a chemical reaction with the GaAs which forms a cap

layer of nonvolatile gallia (Ga_2O_3). Formation of a gallia layer may inject native defects, and it would certainly slow the exchange of atoms between solid and vapor and thus slow the rate of solid–liquid–vapor and native defect equilibration. We suspect that the presence of such a contaminant will explain the persistent differences in the interdiffusion observed between closed and open systems as discussed below.

5.3.1.1. High P_{As}

Fig. 41 summarizes measurements of D_{III} versus $1/T$ for annealing in a quartz ampoule at or near $P_{\text{As}} = 1 \text{ atm}$. Four lines describing D_{III} are provided for comparison: (1) the Wang line (solid), (2) the Tan line (dashed), (3) the Olmsted line obtained from the Olmsted high P_{As} data (heavy shaded), and (4) a second Olmsted line (dot-dashed) obtained after adjustment of the Olmsted low P_{As} line (dividing it by the minimum equilibrium value of $P_{\text{As}}^{1/4}$). Lines 3 and 4 are expected to be the same if the V_{Ga} concentration had quickly equilibrated during all of the Olmsted anneals. Fig. 22 showed that this equilibration occurred reasonably quickly at high temperatures, but not at low temperatures. The crossover of the two Olmsted lines at high temperatures is consistent with this.

The data appear to have little consistency with either the Tan and Gosele line or the adjusted Olmsted line. Most of the data appear to follow either the unadjusted Olmsted line or the Wang line. Along the Olmsted line are shown the highest and lowest values of $D_{\text{Al-Ga}}$ (open diamonds) measured at each temperature by Olmsted et al. [73]. Data measured by Hsieh et al. [12] for $D_{\text{In-Ga}}$ (+) are very close to these. An early result of Olmsted et al. [72] at $T = 855 \text{ }^{\circ}\text{C}$ (open triangle) appears to be consistent with their later results. A result reported by Szafranek et al. [90] (open circle) for a C-doped sample, is somewhat below this line.

Near the Wang line are several interdiffusion results for $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$. A fairly wide variation in the results can easily be seen in the data of Baba-Ali et al. [74] who used $x_{\text{Al}} = 1.0$ (smaller filled squares), or from Chang and Koma [97] who used $x_{\text{Al}} = 0.1$ and 1.0 (bars). Other data which are consistent with the noise level of these data (and show no correlation at all to the Al mole fraction

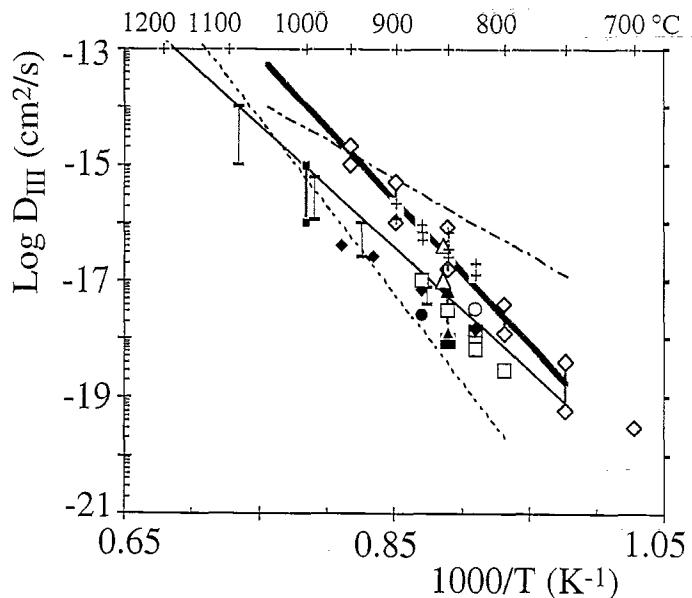


Fig. 41. D_{III} vs. $1/T$ measured after annealing under an As overpressure in a closed ampoule. Proposed fitting lines shown include the Wang line (solid), the Tan line (dashed), the high P_{As} Olmsted line (heavy shaded), and the adjusted low P_{As} Olmsted line (dot-dashed). Specific data are discussed in the text. These values for interdiffusion are among the highest reported for nominally undoped GaAs.

over the range $0.25 \leq x_{\text{Al}} \leq 1.0$) have been presented by Guido et al. [71] (open squares), Hsieh et al. [76] (larger filled square), Camras et al. [98] (filled circle), Furuya et al. [70] (filled triangle), and You et al. [91] (filled diamonds) for heavily C-doped samples.

It is also worth considering possible reasons for the scatter in the data. Two likely causes of experimental error include (1) inaccurate temperature measurement, and (2) inadvertent impurity incorporation in the crystal. Since the data appears to converge at lower temperatures, where radiant heat losses become less important, it is likely that errors in the measurement of temperature account for some of the variation between groups. If this were the only source of error, then the values of D along the Olmsted line should be considered to be closer to the correct values. However, the inexplicable order of magnitude variations (often seen even within individual groups) would require a ~ 50 °C excursion in the same furnace or enormous partial pressure variations, and we consider that neither of these are likely. Since most groups use Al for the interdiffusion marker, and since it is easy to inadvertently grow C into AlGaAs, one must consider the possibility that C is present in some crystals as an acceptor at a high concentration. As seen in Section 5.1, C can significantly reduce D_{III} . Since few groups reported electrical characterization of their GaAs, this may explain some of the apparently random variations in D_{III} . However, if a dopant were the sole cause of the measurement noise, then the data would converge at higher temperatures because $p/n \rightarrow 1$ at high temperatures. The fact that the two groups of data are diverging at higher temperatures suggests that something much more fundamental is occurring. For example, if an oxidant were occasionally present in uncontrolled quantities in the ampoule, then the rock-hard nonvolatile compound gallia, Ga_2O_3 , would occasionally form on the GaAs surface. The presence of any such glass would be expected to affect the rate of native defect equilibration and introduce scatter into D_{III} . After we finish comparing the results from closed and open systems, the subject of gallia formation will be discussed at the end of Section 5.3.

5.3.1.2. Low P_{As}

Measurements of interdiffusion have been made after annealing without any As or Ga present in the ampoule, and these are summarized in Fig. 42. Except for the In marker diffusion of Hsieh et al. [12] (open triangles), the results are for the interdiffusion of Al markers. No consistent dependence upon the Al mole fraction is apparent. Olmsted et al. [73] took data at 50 °C intervals from 750–1050 °C. They did not publish their data for these experiments, but instead reported a line which fit the data and this is shown as the shaded line. However, Drs. Olmsted and Houde-Walter [99] have kindly made those data available for this review, and they are tabulated in Table 7 and shown here for the first time in Fig. 42 (open circles). An early result reported by Olmsted et al. [72] for D_{III} at $T = 855$ °C is shown by the (+). Additional Al marker data shown have been taken from Baba-Ali et al. [74] (filled squares) and Hsieh et al. [76] (filled circle).

Taken as a whole, the results tend to be described reasonably well by the Olmsted low P_{As} line (shaded), but this appears to be similar to the nearby high P_{As} Wang line (heavy solid). Adjustment of the Wang line to the appropriate P_{As} (at the edge of the Ga-rich solidus) gives a line (dot-dashed) which is well below nearly all of the data. Adjustment of the Olmsted high P_{As} line, shown in Fig. 41, to the appropriate P_{As} also gives a line (thin solid) which poorly describes the data. The Tan line (dashed) fits closely to this adjusted Olmsted line. We consider it likely that native defect equilibration has not occurred at the lower temperatures and we shall discuss this later in this section.

5.3.1.3. Strong nonequilibrium

Fig. 43 shows D_{III} versus $1/T$ after annealing in a highly nonequilibrium environment, i.e. with liquid Ga intentionally added to the ampoule before annealing. The results of Hsieh et al. [12] (X) for In marker diffusion and are comparable with the Al marker diffusion summarized by the other

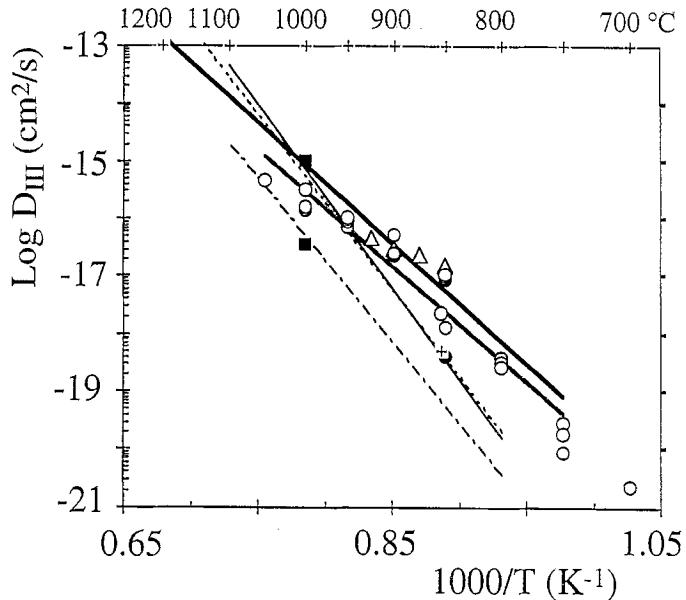


Fig. 42. D_{III} vs. $1/T$ obtained after annealing in a closed ampoule with no As or Ga added to the ampoule, i.e. in a Ga-rich ambient. Possible fitting lines include the original ($P_{As} = 1$ atm) line from Wang (heavy solid), a Wang line adjusted to low P_{As} (dot-dashed), Olmsted low P_{As} (shaded), Olmsted high P_{As} line adjusted to low P_{As} (thin solid), and Tan et al. (dashed). These lines and specific data are discussed in the text.

Table 7

Measured interdiffusion for quantum wells of three different widths which had been annealed in closed ampoules containing no added As or Ga. These previously unpublished data were provided by Drs Olmsted and Houde-Walter. The epilayers were grown on Si-doped substrates

| T (°C) | D (cm ² /s) $W = 521 \text{ \AA}$ | T (°C) | D (cm ² /s) $W = 228 \text{ \AA}$ | T (°C) | D (cm ² /s) $W = 91 \text{ \AA}$ |
|-------------|---|-------------|---|-------------|--|
| 850 | 8.68×10^{-18} | 800 | 3.71×10^{-19} | 700 | 2.22×10^{-21} |
| 850 | 1.05×10^{-17} | 850 | 9.01×10^{-18} | 750 | 2.79×10^{-20} |
| 900 | 2.23×10^{-17} | 900 | 5.19×10^{-17} | 750 | 8.51×10^{-21} |
| 900 | 2.5×10^{-17} | 950 | 8.86×10^{-17} | 750 | 1.77×10^{-20} |
| 950 | 9.97×10^{-17} | | | 800 | 2.98×10^{-19} |
| 950 | 6.93×10^{-17} | | | 800 | 2.53×10^{-19} |
| 1000 | 1.34×10^{-16} | | | 850 | 1.27×10^{-18} |
| 1000 | 3.05×10^{-16} | | | 855 | 2.27×10^{-18} |
| 1000 | 1.6×10^{-16} | | | | |
| 1050 | 4.55×10^{-16} | | | | |
| 1050 | 4.36×10^{-16} | | | | |

data. All of the undoped results of Olmsted et al. [73] (open diamonds), and their first reported measurement at $T = 855$ °C (filled square) are highly consistent with early measurements of Hsieh et al. [76] (open triangle). The results of Szafranek et al. [90] (open square) in heavily C-doped GaAs is somewhat higher, but roughly consistent with these results. However, the data from You et al. [91] (open circles) for heavily C-doped GaAs are clearly inconsistent with the others. For C-doped GaAs in contact with a Ga-rich liquid, it is reasonable to suspect that the interstitialcy mechanism may be acting to increase the measured D_{III} because of the increased tendency for a high concentration of I_{Ga}^+ to be in such a crystal. On the other hand, the only compelling evidence to date for the interstitialcy mechanism has been obtained when group II atoms are present, and it is possible that their measure-

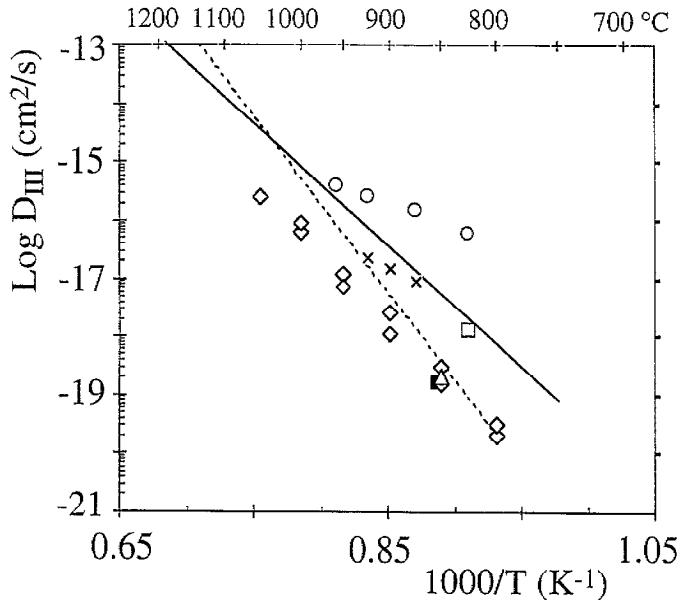


Fig. 43. D_{III} vs. $1/T$ measured after annealing in a closed ampoule in a highly nonequilibrium environment, i.e. with Ga metal intentionally added to the ampoule. Although these results come from strong nonequilibrium conditions, the Wang line (solid) and the Tan line (dashed) are provided to allow easy comparison with the other figures. Specific data are discussed in the text. These results for interdiffusion are among the lowest reported for GaAs.

ments are a result of group II contamination. Thus, until the results of You et al. can be duplicated by others, strong conclusions should not be drawn from them.

All of the data shown in Fig. 43 are the result of nonequilibrium native defect concentrations which change significantly during the anneal period. This is obvious because interdiffusion in the undoped crystals summarized in Fig. 43 is substantially lower than the interdiffusion in their undoped counterparts shown in Fig. 42. Since the final equilibrium state is the same for all of these samples, the additional time required for a significant mass of pure liquid Ga to reach its equilibrium composition at the edge of the liquidus line must cause a nonequilibrium native defect concentration to exist for a significant time.

Since the liquid metal must have a nonequilibrium Ga-rich composition until sufficient GaAs decomposes, it is reasonable to expect that the Ga partial pressure will be higher than its maximum equilibrium value over GaAs. Since we have seen evidence that the Ga interstitial concentration closely follows P_{Ga} , it is reasonable to expect that an increased Ga interstitial concentration plus the Frenkel reaction, i.e. $K = [I_{\text{Ga}}][V_{\text{Ga}}]$, dictates that a reduction will occur in the V_{Ga} concentration for an extended period of time. This fits with several pieces of evidence: D_{III} in undoped GaAs is relatively rapid when the crystal is at the As-rich edge of the solidus, and D_{III} is continually reduced as the solid composition changes to the Ga-rich edge of the solidus, and then D_{III} is reduced further still when the crystal is forced into a nonequilibrium Ga-rich state for an extended period of time! Thus, we conclude that the group III interdiffusion in undoped, n-type, and probably p-type GaAs remains controlled by Ga vacancies at all values of P_{As} .

5.3.2. Interdiffusion using a proximity cap in open tube system

Several groups have measured interdiffusion in GaAs in intimate contact with a proximity cap in an open tube with a flowing gas ambient. If high purity, dry H_2 is present in the flowing gas, then the native oxides are known to be reduced [24] and this may be the cleanest ambient available for the study of diffusion in GaAs. A choice of the appropriate degrees of freedom can easily be made to define the

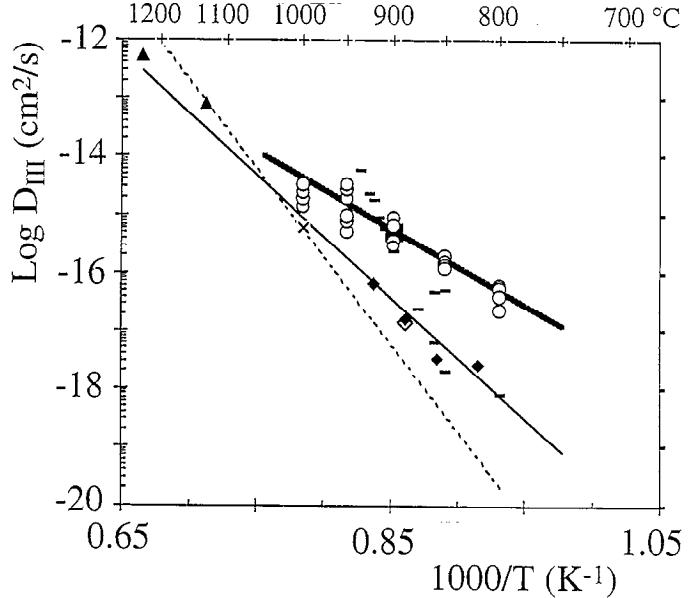


Fig. 44. D_{III} vs. $1/T$ measured after annealing in an open tube system at various values of P_{As} , and n , after adjustment to $P_{\text{As}} = 1 \text{ atm}$, and $n = n_i$, i.e. assuming a singly charged negative Ga vacancy controls the diffusion. Much of the data pass through the Olmsted low P_{As} line, adjusted to $P_{\text{As}} = 1 \text{ atm}$ (shaded). The Wang line (solid) and the Tan line (dashed) are provided for comparison. Specific data are discussed in the text.

equilibrium state of the crystal. As discussed in Section 2.2.2, the use of a proximity cap and convection barriers provides a way of quickly bringing the solid close to this well-defined equilibrium state.

Fig. 44 summarizes the available data after adjustment to standard conditions, i.e. to $P_{\text{As}} = 1 \text{ atm}$ and $n = n_i$. This adjustment requires that a native defect model of diffusion be used because D_{III} was obtained at various values of P_{As} and n . Strong evidence was discussed in Section 5.1 that showed V_{Ga}^{-1} controlled interdiffusion when native defect equilibrium could be approximated. Thus, we have calculated the standard diffusivity for group III interdiffusion as discussed in Section 3.2.2, i.e. using

$$D_{\text{III}, V_{\text{Ga}}}(T, P_{\text{As}} = 1, n_i) = \frac{D_{\text{measured}}(T, P_{\text{As}}, n)}{P_{\text{As}_4}^{1/4} \{n/n_i\}} \quad (155)$$

Using Fig. 1 to obtain P_{As_4} over the Ga-rich solidus (for samples annealed without any As added to the ambient), and using the data of Blakemore [84] for the intrinsic carrier concentration, n_i , the measured diffusivities from several groups have been adjusted to give the diffusivity at standard conditions. In Fig. 44, a single data point (filled square) at $T = 900 \text{ }^{\circ}\text{C}$ summarizes all of the data from Li et al. for intrinsic and extrinsic data already shown in Fig. 29. The Te-doped SL interdiffusion data of Mei et al. [87] (open circles) have been used to calculate a standard diffusivity corresponding to $n = n_i$ at $P_{\text{As}} = 1 \text{ atm}$ for each of their reported results. The shaded line in Fig. 44 shows the low P_{As} line of Olmsted et al., after adjustment to $P_{\text{As}} = 1 \text{ atm}$. This line passes through the 800–1000 °C data of Li et al., Mei et al., and the higher temperature data of Fleming et al. [81] (bars). The fact that this line, rather than the Olmsted high P_{As} line (not shown), fits what we consider to be the more reliable data may reflect a common feature between the open system and the low P_{As} ampoule experiments: when no additional solid As or Ga has been added, then no contaminants associated with those sources have entered the system.

As discussed earlier, the low temperature data of Fleming et al. in Fig. 44 was among the first reported data for interdiffusion of superlattices of nominally monolayer thickness. We consider these

values of D_{III} to be less reliable than their high temperature data for D_{III} which was inferred from the much wider SLs. We have averaged the 1995 data of Seshadri et al., [75] shown in Fig. 22, into a single point (open diamond). These data were taken after annealing the crystals without a proximity cap on the OMVPE growth susceptor at a reported $T = 888$ °C. Given the inherently large heat transfer which occurs by radiation losses at radiant temperatures, such an approach to annealing would almost certainly leave the GaAs colder than the susceptor. Indeed, these data were clearly out of place in Fig. 22, which showed a family of curves corresponding to different temperatures. From this plot, constructed with the data of several groups using furnaces, make it clear that the actual temperature of the Seshadri et al. samples was almost certainly near $T = 830$ °C, or approximately 60 °C lower than reported. We also note that in 1992 Seshadri et al. [100] reported interdiffusion data (filled diamonds) from the same or similar samples using the same experimental technique at several temperatures. If we simply assumed that all of these were annealed at 60 °C below the reported temperature, then these data would lie approximately equidistant between the solid and shaded lines.

A single point from Beernink et al. [101] (×) and the two points from Wang et al. [68] (filled triangles) were taken using N_2 ambients, in contrast to the work described above which used H_2 . Without H_2 present to create a reducing ambient, it is reasonable to wonder if a native oxide hardened into (or if trace oxidants formed) the nonvolatile oxide, Ga_2O_3 . If this occurred, it may have affected the rate of exchange of atoms between the solid and vapor. However, one cannot determine from these data whether the different ambient had any effect on the results for the short anneal periods used.

5.3.3. In situ oxidation and reduction

It is routinely assumed that no oxidants are present in the ampoules used for annealing. We believe that this is a questionable assumption because the common method of sealing an ampoule with an oxyhydrogen flame will introduce a significant quantity of H_2O into the quartz. Indeed, sometimes the inside of the quartz tube is exposed to this flame to remove ‘surface water’. If the H_2O trapped in the quartz diffuses into the quartz ampoule, it is expected to react chemically with the GaAs. However, the extent to which these reactions have actually affected the reported diffusivities is unknown at present.

Conversely, a reducing ambient is expected to exist when H_2 is present above GaAs in an open system. Even if small amounts of H_2O are present in the H_2 carrier gas, the use of a GaAs proximity cap effectively protects the sample surface since the perimeter of the cap and wafer will act as getters for residual oxidants. Thus, we consider an open tube system to be the *cleanest annealing system available* in practice.

Sufficient data collected over many years has shown that there are real differences between the diffusion techniques used, and that there is also substantial experimental noise in the interdiffusion measurements. Even more remarkable is the lack of experimental noise in the high P_{As} isotopic self-diffusion data in Fig. 21 compared with the much larger noise seen within single groups and between groups when Al is used as the marker. The data is too sparse to compare In markers between groups, but within individual groups the noise appears to be less than that associated with Al. This suggests that a contaminant which strongly reacts with Al, such as O, affects the marker diffusion and adds scatter to the results. However, the exact relationship between marker diffusion and oxygen remains to be clarified.

5.3.4. Depth dependence

A dependence in D_{III} upon the distance below the GaAs surface has been reported several times for interdiffusion. In general, the native defect concentrations in a substrate wafer is expected to be different than those in an epilayer which was grown upon that substrate at low temperatures in a short

time. Neither epilayer nor substrate are likely to be close to equilibrium with the ambient at the start of an annealing. Because point defects will enter an epitaxial layer from both the surface and from the substrate, any depth dependence measured for D_{III} needs to be interpreted cautiously. For example, if two interdiffusion markers grown into a GaAs epilayer diffuse at different rates, then did defects associated with the surface affect the top marker, or did defects associated with the substrate affect the bottom marker, or both? This can only be answered with a careful study of marker diffusion as one perturbs the external ambient and/or the initial native defect concentrations in the wafer.

Any depth dependence is further complicated because the driving force for a flux of charged native defects is the gradient in electrochemical potential, as discussed in Section 3.2.4. Thus, *large concentration ratios* of native defects will *exist at equilibrium* across any electrical junction, including any epilayer and substrate with different doping density. It is only when this large concentration ratio, or the potential difference across the junction, is perturbed from its equilibrium value that a net flux of native defects will move between substrate and epilayer.

Wang et al. [68] noted that isotopic interdiffusion was measured for some epilayers grown on n-type, p-type, or semi-insulating substrates and they found no difference in D_{III} , to within their relatively low measurement noise of about $2 \times$. Wang et al. also noted a persistent tendency for the top isotopic heterojunction to have diffusivities slightly higher (about $2 \times$ higher) than that of the bottom heterojunction over a wide range of temperature. Although this result is consistent with a model in which Ga vacancies approach a high equilibrium value set by the external ambient, it is not conclusive evidence.

Seshadri et al. [75] reported that at $T = 888^\circ\text{C}$, $D_{\text{Al-Ga}}$ was slightly higher (about $2 \times$) in their bottom quantum well rather than their top well. They interpreted this result in terms of an outflow of vacancies from the semi-insulating substrate. However, we consider that there are insufficient results from this study to draw any firm conclusions.

Baba-Ali et al. [74] reported values of $D_{\text{Al-Ga}}$ for closed ampoule interdiffusion of a pair of grown-in quantum wells at $T = 1000^\circ\text{C}$ at high and low P_{As} . Faster interdiffusion tended to be associated with the quantum well near the surface [102]. However, the measurement noise was large and only a small number of samples were studied, and no mechanistic conclusions can be drawn from this study.

Li et al. [89] observed an approximately twofold increase in $D_{\text{In-Ga}}$ in the deeper n-type layer of an npn structure grown on a Si-doped substrate. The cause for this is unknown. However, in studies of D_{III} using uniformly doped epilayers, this group reported [89] that very large variations in $D_{\text{In-Ga}}$ were occasionally observed when a marker was coincident with the substrate–epilayer interface. This result suggests that one source of measurement noise may be a point defect source associated with grown-in imperfections at the interface between the substrate and epilayer. However, at present this is a poorly understood and unstudied topic.

5.4. Diffusion in GaAs under dielectric cap layers

For an uncapped GaAs wafer in contact with a vapor containing Ga, As, and an inert gas, $F = 3$ degrees of freedom, i.e. three independent variables such as T , P_{As} , and P_{total} , must be chosen to define the equilibrium state. If the GaAs crystal is covered by an ideal cap which completely isolates the solid from the vapor, then we still have $F = 3$ as discussed in Section 2.4. However, no exchange of atoms takes place between the GaAs and vapor, and the equilibrium native defect concentrations can no longer be defined without contact with the external reservoir of atoms. Thus, another thermodynamic variable must be chosen to define the equilibrium defect concentrations. In practice, there are no variables available for *direct control* by the experimenter during the anneal. However, there are variables which have been *indirectly controlled* in the prior processing step i.e. the precise crystal

composition acts as the third degree of freedom for annealing under an ideal cap, as discussed in Section 2.4. In practice, the defect with the highest concentration is expected to act as a reservoir of atoms and/or native defects, during the capped anneal.

Often, crystals are doped and the associated fourth degree of freedom may be taken to be the carrier concentration. However, the initial native defect concentrations are generally uncontrolled at the beginning of an anneal, and thus one cannot expect a simple power law relationship between D and carrier concentration to be observed. Indeed, in one of the broader series of studies of diffusion in both ion implanted and nonimplanted capped III-V semiconductors, Gillin et al. [103] reported *no dependence* at all of D_{III} upon carrier concentration.

Diffusion in a GaAs epitaxial layer covered by a real (nonideal) cap on one side, and by a substrate on the other side, is far more complicated than suggested by the difficulties associated with the initial native defect stoichiometry. Some fundamental considerations include (1) real substrates are likely to act as a reservoir of native defects, and to cause the native defect concentrations in the epilayer to drift for a significant period of time, (2) real caps can be porous and thus diffusion in the GaAs sometimes depends upon the value of P_{As} above the sample, and (3) real caps may exhibit poorly understood metallurgical reactions with GaAs, and thus act as a source or sink of point defects. Less fundamental, but serious engineering problems likely to affect diffusion include (1) variations in the cap composition, (2) variations in the cap density, (3) stresses induced in the crystal, and (4) systematic errors in the measurement of temperature.

The above discussion is intended to suggest that capped annealing is a fundamentally more complicated technique than uncapped annealing, and thus it will be more difficult to control and interpret the results. Nevertheless, we shall briefly summarize some interdiffusion results in an effort to draw a few preliminary conclusions about diffusion under real dielectric caps.

5.4.1. Group II diffusion

Haddara et al. [104] have reported the effect of SiO_2 and Si_3N_4 caps on the diffusion of Be out of buried layers. Because of the sensitivity of group II element diffusion to the presence of I_{Ga} , a significant increase in group II dopant diffusion can be taken as evidence of an increase in the concentration of I_{Ga} . Fig. 45 shows how the as-grown Be concentration profile, as measured by SIMS, evolves when GaAs is annealed under these two different caps. It is clear that the different caps have affected the native defect concentrations. Although these profiles are not gaussian, we shall assume that they are to obtain a rough estimate of D_{Be} . For Be diffusion under the Si_3N_4 cap,

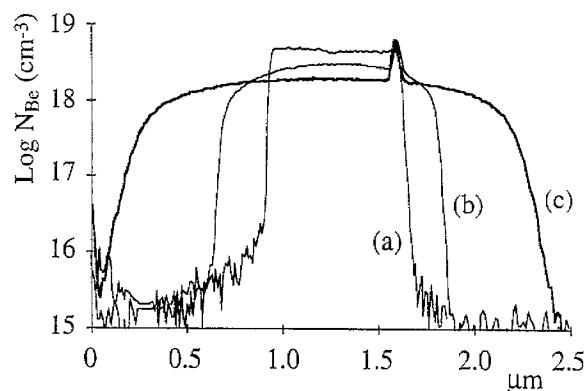


Fig. 45. Evolution of Be concentration profiles measured by SIMS. Reprinted with permission from Ref. [69], copyright 1996, American Institute of Physics. The as-grown Be profile (a) widens, after a $t=2$ h anneal at $T=900$ °C, to (b) when annealing with a SiO_2 cap, but widens significantly more to (c) when annealing with a Si_3N_4 cap. The rapid group II diffusion strongly suggests that I_{Ga} were injected into the GaAs below both caps.

$D_{\text{Be}} \approx 3 \times 10^{-13} \text{ cm}^2 \text{ s}^{-1}$, and under the SiO_2 cap, $D_{\text{Be}} \approx 5 \times 10^{-14} \text{ cm}^2 \text{ s}^{-1}$. Both of these values are large considering that there is no dopant in an external phase driving the diffusion. The discussion of Section 5.2 showed that enhanced group II diffusion appears to always be associated with a large nonequilibrium concentration of Ga interstitials in uncapped GaAs. Indeed, for uncapped GaAs, group II diffusion via a vacancy mechanism is negligibly small. Thus we conclude from this experiment that Ga interstitials have been injected under both caps, but that more interstitials have been injected under the Si_3N_4 cap. This conclusion will be consistent with the results of several other groups discussed below.

Presumably the actual process is sufficiently complicated that such a simple description will be of limited use. Indeed, from the results of their work [105,106], Haddara et al. concluded that they could not simulate the diffusion process solely by a kick-out mechanism induced by injected I_{Ga} . Based upon their simulations, they concluded that the diffusion process involved time dependent concentrations of both I_{Ga} and V_{Ga} . Although some might criticize their approach for simply adopting enough parameters to fit the data, some of the results discussed below indicate that this idea may have a reasonable physical basis.

A key problem in evaluating diffusion under caps is that few groups bother to evaluate the quality of their caps, substrates, stress, etc., and thus there is an enormous range of results which cannot be effectively interpreted. The following discussion of capped experiments is not intended to provide answers to these problems, but to provide a context in which one may begin to appreciate the extent of some of these problems. Because of the various techniques used to deposit dielectric cap layers, the cap composition often varies and we shall generally identify the compositions as ' SiN_x ' or ' SiO_x '. Consistent with the rest of this paper, we shall not discuss crystals which have been ion implanted.

5.4.2. ' SiN_x '

Some of the experimental highlights of the reported work with SiN_x caps are shown in Table 8, and the reported D_{III} versus $1/T$ results are summarized in Fig. 46. The Wang line is shown simply for a convenient comparison.

5.4.2.1. In markers

Several results have been reported for In marker diffusion in GaAs at concentrations up to $x_{\text{In}} = 0.25$. Since high mole fractions are expected to induce significant strain into the crystal, these experiments offer an opportunity to evaluate the effect of strain. Using undoped $\text{In}_x\text{Ga}_{1-x}\text{As}$ wells, with $0.10 \leq x_{\text{In}} \leq 0.25$, placed 50–280 nm below the surface, Gillin et al. [108] found (small filled diamonds) that quantum wells placed near the surface diffused somewhat more rapidly than those placed deeper, but could measure no significant dependence upon x_{In} . Gillin et al. [103,115] have also reported that D_{III} was unaffected by $N_{\text{Si}} \approx 1 \times 10^{18} \text{ cm}^{-3}$ (open circles) or $N_{\text{Be}} \approx 2.5 \times 10^{19} \text{ cm}^{-3}$ (bars). In a separate study, Bradley et al. [109] reported (open diamonds) that essentially the same D_{III} was obtained with or without ion implantation. Presumably, strain caused by the large In mole fraction has largely relaxed through the formation of dislocations, and this appears to have little effect on the results. The lack of dependence upon strain, x_{In} , or doping suggests that the native defect concentrations are consistently far from their equilibrium values. However, 'nonequilibrium space' can be quite large. For example, the results of Taylor et al. [107] (+ signs) show some of the largest values of D_{III} versus $1/T$ ever reported for a SiN_x cap. In contrast to the results of Gillin et al., Taylor et al. reported that wells near the surface diffused at a slower rate than deeper wells.

Table 8
Brief summary of experimental parameters associated with the measured interdiffusion of a marker element in GaAs under an “ SiN_x ” cap layer

| Reference | SiN_x cap | Marker, Mole fract. | Dopant and concentration (cm^{-3}) | Anneal ambient | Structure | Comments |
|---------------------|---------------------------------------|---------------------|--|---|------------------------------------|---|
| [107] (Taylor93) | 80 nm | In, 0.18 | nom. undoped | 10 s RTA, N ₂ | 4 QWs 0.15–0.33 μm deep | D lower near surface |
| [103] (Gillin93a) | 50 nm PECVD at 300 °C | In, 0.2 | $\leq 1 \times 10^{18}$ Si, undoped, $\leq 2.5 \times 10^{19}$ Be | RTA | QW 0.1 μm deep | No Fermi-level effect on D |
| [108] (Gillin93b) | 50 nm PECVD at 300 °C, front and back | In, 0.10–0.25 | undoped | 15–600 s RTA | QWs 50–280 nm deep | D higher near surface; some time dependence noted |
| [109] (Bradley93) | 50 nm PECVD at 300 °C, front and back | In, 0.2 | undoped | RTA, FA, 15 s–6 h | QW 0.1 μm deep | D similar to that after implant |
| [103] (Gillin93a) | 50 nm PECVD at 300 °C | Al, 0.2 | $\leq 1 \times 10^{18}$ Si, undoped | 15–500 s RTA | QW 0.15 μm deep | No Fermi-level effect on D |
| [110] (Choi95) | 300 °C PECVD | Al, 0.2 | nom. undoped, Zn at surface | 35 s RTA, Ar | QW 1.3 μm deep | PECVD affects D before anneal |
| [111] (Beernink95a) | 80 nm | Al, 0.4 | $3 \times 10^{18} - 1 \times 10^{19}$ Si | 1–4 h FA | 1.3 μm SL | |
| [112] (Ralston91) | 250 nm PECVD | Al, 0.32 | 4×10^{18} Si in center of QWs | 15 s RTA, H ₂ | MQW 220 nm | |
| [113] (Major90) | CVD at 720 °C | Al, 0.25 | 2×10^{18} Se or C | RTA, H ₂ /N ₂ , GaAs prox cap | 1.1 μm deep | Very low D |
| [71] (Guido87) | 100 nm CVD at 720 °C | Al, 0.25 | nom. undoped | ampoule + As, 25 h | QW 1.3 μm deep | Very low D |
| [114] (Lee87) | 100 nm PECVD | Al, 0.3, 0.7, 1.0 | undoped | ampoule, 1–8 h | 0.4 μm SL | Low D; 70% Al inconsistent w/ others |
| [90] (Szafranek91) | 100 nm CVD at ~700 °C | Al, 0.3 | 8×10^{18} C | ampoule, Ga or As-rich, 24 h | 0.8 μm SL | Very low D; D changes ~25% w/ ambient |

“Nominally undoped” indicates negligible doping in the QW region, but large dopant concentrations were within 0.2–0.3 μm of the QW (typical laser structure).

“Undoped?” means presumably undoped structure, but carrier concentration either not mentioned or characterization unclear. RTA = rapid thermal anneal. FA = furnace anneal. In all cases, D was determined by PL.

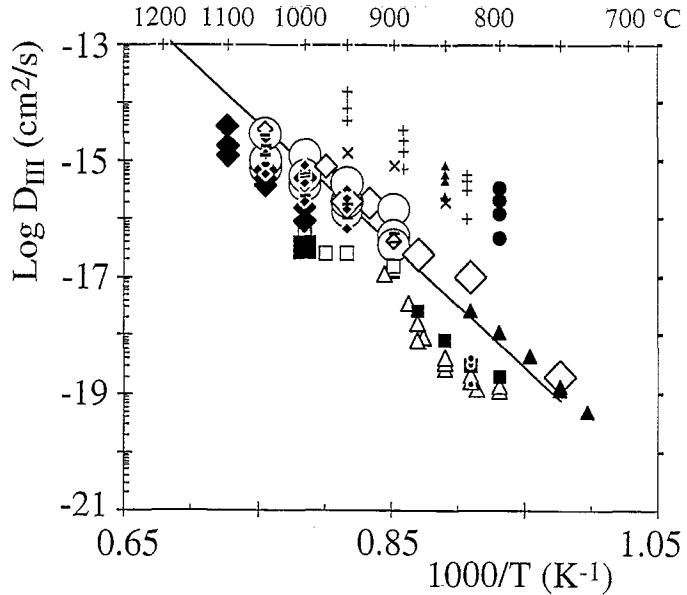


Fig. 46. D_{III} vs. $1/T$ measured after annealing with a ' SiN_x ' cap. The Wang line is shown for comparison. Although there is an enormous scatter in the data, a significant fraction of the results indicate that group III diffusion can be slowed under a nitride cap. The lowest values of D_{III} at any given temperature are roughly comparable to those measured without a cap layer, i.e. uncapped GaAs in a Ga-rich nonequilibrium ambient (liquid Ga present). To the extent that the glass has been characterized, a low D_{III} appears to be associated with a glass of high density. These results are also consistent with the injection of I_{Ga} from the cap. Specific data are discussed in the text.

5.4.2.2. Al markers

There is also a wide range of D_{III} values reported using Al as the interdiffusion marker. Inexplicable variations in D_{III} coming from one group, using nominally identical techniques for growth and characterization, can be found. For example, Lee et al. [114] studied the interdiffusion of $\text{Al}_x\text{Ga}_{1-x}\text{As}$ -GaAs quantum wells for $x_{\text{Al}} = 0.3, 0.7$, and 1.0 . They found nearly identical D_{III} for $x_{\text{Al}} = 0.3$ and 1.0 (open triangles), but an order of magnitude larger D_{III} for $x_{\text{Al}} = 0.7$ (filled triangles). No apparent dependence upon doping has been found. For example, Gillin et al. [103] found that $N_{\text{Si}} \approx 1 \times 10^{18} \text{ cm}^{-3}$ had no significant effect (large filled diamonds) on D_{III} , and Major et al. [113] reported only minor differences have been observed between D_{III} when using $N_{\text{Se}} \approx 2 \times 10^{18} \text{ cm}^{-3}$ (open squares) or $N_{\text{C}} \approx 2 \times 10^{18} \text{ cm}^{-3}$ (large filled square). The relatively high values of Ralston et al. [112] (\times s) and Beernink et al. [111] (larger filled circles) come from n-type GaAs, but their dependence upon the doping appears weak and it is unlikely that doping played a significant role. Some of the lowest values of D_{III} have been reported by Guido et al. [71] (smaller filled squares). Similar values of D_{III} were reported by Szafranek et al. [90] in their heavily C-doped epilayers (smaller filled circles). However, at similar annealing temperatures, Choi et al. [110] (small filled triangles) reported a value of D_{III} which was 3 orders of magnitude larger.

5.4.2.3. Mole fraction

In Fig. 46, there is no evidence to indicate that a significant difference exists between $D_{\text{Al-Ga}}$ and $D_{\text{In-Ga}}$ (for In fractions of up to $\approx 25\%$). Some reports find increasing D_{III} near the GaAs surface but others find decreasing D_{III} near the GaAs surface. There is no trend which can be correlated with the In or Al mole fraction. Indeed, one group found nearly identical results for $x_{\text{Al}} = 0.3$ and $x_{\text{Al}} = 1.0$, but quite different results for $x_{\text{Al}} = 0.7$.

5.4.2.4. Glass quality

A wide range of diffusivities have been measured under a SiN_x cap. The density, or porosity, of the glass appears to play an important role in this variation. The lowest diffusivities at several temperatures appear to be associated with thermal CVD (chemical vapor deposited) nitride [71,113], or with efforts to make PECVD (plasma enhanced CVD) produce a glass as dense as possible [116]. The low rates of interdiffusion associated with these denser SiN_x caps are comparable with those observed after closed ampoule annealing in the presence of intentionally added Ga metal! The discussion of Section 5.1 showed that this extremely Ga-rich annealing environment reduced D_{III} by reducing the vacancy concentration via the injection of I_{Ga} , and we conclude that D_{III} is reduced here because of I_{Ga} injection under a dense SiN_x cap.

5.4.2.5. Processing damage

It is also worth noting that Choi et al. [110] found that the parameters chosen for the PECVD process could induce intermixing during the nitride deposition. This conclusion was reached after PL measurements found an unexpected quantum well energy shift both after PECVD and again after RTA (and both times the nitride had been removed to avoid stress effects in PL!). In the most extreme case, after a 12 min. exposure to a low power plasma, the results showed that 50% of the intermixing (normally attributed entirely to RTA) apparently occurred during PECVD. Similarly, Kupka and Chen [117] have reported that a minor amount of damage close to the surface caused by RIE (reactive ion etch) leads to significant diffusion when processing later under Si_3N_4 or SiO_2 . We suspect that these results reflect metallurgical reactions occurring between the GaAs and the nitride which have been chemically activated by the plasma treatment of the surface. However, the effects of these reactions upon the native defect concentrations are very poorly understood.

5.4.3. ' SiO_x '

Some of the experimental highlights of the reported work with SiO_x caps are shown in Table 9, and the reported D_{III} versus $1/T$ results are summarized in Fig. 47. The Wang line is shown for a convenient comparison.

5.4.3.1. In markers

Burkner et al. [118] appear to have presented the only results for In marker diffusion using $x_{\text{In}}=0.2$. They compared the relative effect on D_{III} from electron-beam (e-beam) deposited SiO_x (larger filled diamonds) a commercial 'spin-on' SiO_x (larger filled squares) and PECVD deposited SiO_x (larger filled triangles). The e-beam deposited glass was found to be associated with high values of D_{III} , and the diffusion associated with the other two glasses was found to be similar.

5.4.3.2. Al markers

Burkner et al. also used e-beam deposited glass to study interdiffusion of Al using $x_{\text{Al}}=0.3$ adjacent to a GaAs quantum well. Using 250 nm of glass, they found D_{III} to be time dependent (open diamonds), and they noted that the glass shrank substantially during annealing. Presumably the glass cracked sufficiently to allow some contact between the crystal and ambient, and an interpretation of this experiment is not simple. They reported some dependence upon well depth and upon the cladding layer composition several hundred nm away, with $x_{\text{Al}}=0.5$ (open squares) associated with slightly faster diffusion than $x_{\text{Al}}=0.8$ (open triangles). Using e-beam deposited glass, Ralston et al. [119] (open circles) have reported fairly high values of D_{III} . Very high values of D_{III} were reported by Gontijo et al. [120] (bars) who annealed under PECVD glass. For comparison, the same group measured relatively low values of D_{III} when using an SrF_2 PECVD glass (filled circles). The lowest

Table 9
Brief summary of experimental parameters associated with the measured interdiffusion of a marker element in GaAs under an “ SiO_x ” or “ SrF_2 ” cap layer

| Reference | SiO_x cap | Marker, Mole frac. | Dopant and concentration (cm^{-3}) | Anneal ambient | Structure | Comments |
|---------------------|--|--------------------|---|---|--|---|
| [118] (Burkner95) | 250 nm e-beam evap | In, 0.2 | undoped? | RTA 15 s, H_2 | QWs 400/800 nm deep | D slightly higher near surface |
| [118] (Burkner95) | 200 nm spin-on | In, 0.2 | undoped? | RTA 15 s, H_2 | QWs 400/800 nm deep | |
| [118] (Burkner95) | 220 nm PECVD | In, 0.2 | undoped? | RTA 60 s, H_2 | QWs 400/800 nm deep | Low D obtained |
| [118] (Burkner95) | 250 nm e-beam evap | Al, 0.3 | undoped? | RTA 15–60 s, H_2 | QWs 400/800 nm deep | D(15 s) > D(60 s). RTA shrinks oxide. |
| [118] (Burkner95) | 35–315 nm e-beam evap | Al, 0.3 | undoped? | RTA 60 s, H_2 | QWs 400/800 nm deep | D up w/ x (ox). Al in distant cladding plays a role. |
| [101] (Beernink95b) | 300 nm CVD | Al, 0.4 | nom. undoped | RTA, FA, 0.1–5 h; N_2 , GaAs prox. cap | QW 1.1 μm deep. AlGaN/P below QW | Very low D |
| [71] (Guido87) | 100 nm 420 °C CVD SiO_2 | Al, 0.25 | nom. undoped | 25 h ampoule w/ As | QW 1.3 μm deep | Very low D |
| [119] (Ralston88) | 300 nm e-beam evap | Al, 0.32 | undoped? | 15 s RTA, GaAs prox cap | 400 nm SL | |
| [113] (Major90) | 420 °C CVD SiO_2 | Al, 0.25 | 2×10^{18} , Se or C | RTA H_2/N_2 ; GaAs prox cap | QW 1.1 μm deep | Very low D |
| [90] (Szafranek91) | 100 nm CVD at ~400 °C | Al, 0.3 | 8×10^{18} C | ampoule, Ga or As-rich, 24 h | 0.8 μm SL | Very low D; D changes ~25% w/ ambient |
| [120] (Gontijo94) | 200 nm PECVD | Al, 0.2 | nom. undoped | RTA 30 s; C/GaAs prox caps | QW 1.1 μm deep | Very high D |
| [120] (Gontijo94) | 200 nm evap SrF_2 covered by PECVD SiO_2 | Al, 0.2 | nom. undoped | RTA 30 s; C/GaAs prox caps | QW 1.1 μm deep | Low D |

“Nominally undoped” indicates negligible doping in the QW region, but large dopant concentrations were within 0.2–0.3 μm of the QW (typical laser structure).

“Undoped?” means presumably undoped structure, but carrier concentration either not mentioned or characterization unclear. RTA = rapid thermal anneal. FA = furnace anneal. In all cases, D was determined by PL.

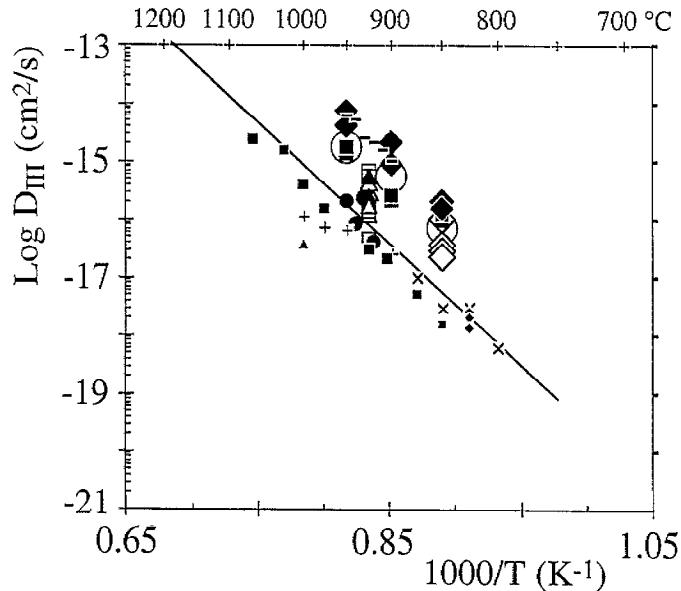


Fig. 47. D_{III} vs. $1/T$ measured after annealing with a ' SiO_x ' cap. The Wang line is shown for comparison. The lowest values of D_{III} at any given temperature are roughly comparable to those measured without a cap layer, i.e. uncapped GaAs in a Ga-rich ambient. To the extent that the glass has been characterized, a low D_{III} appears to be associated with a glass of high density. These results are also consistent with the injection of I_{Ga} from the cap. Specific data are discussed in the text.

values of D_{III} appear to be associated with glass deposited by thermal CVD as reported by Guido et al. [71] (\times), Beernink et al. [101] (smaller filled squares), n-type $N_{\text{Se}} = 2 \times 10^{18} \text{ cm}^{-3}$ (+) and p-type $N_{\text{C}} = 2 \times 10^{18} \text{ cm}^{-3}$ (smaller filled triangle) of Major et al., [113] and p-type $N_{\text{C}} = 8 \times 10^{18} \text{ cm}^{-3}$ of Szafranek et al. [90] (smaller filled diamonds). The results of Burkner et al. indicate that $D_{\text{In-Ga}}$ is slightly larger than $D_{\text{Al-Ga}}$, but we do not conclude that a real difference exists based upon the few results available. Burkner et al. also reported extensive measurements of the variation of D_{III} with the O/N ratio in glass caps, with the Al composition in the cladding layers, and with time.

5.4.3.3. Glass quality

Fig. 47 shows a wide range of results for D_{III} versus $1/T$ after annealing GaAs under a SiO_x cap. SiO_x appears to be associated with interdiffusion which is somewhat higher, on the average, than the interdiffusion associated with a SiN_x cap. Similar to SiN_x , high temperature CVD appears to be associated with reproducible lower values of D_{III} , and presumably this is a result of a denser, less porous, glass. Values of D_{III} associated with the higher quality glasses are comparable with D_{III} measured in ampoule diffusions nominally at equilibrium with the Ga-rich edge of the GaAs solidus. Similar to the case with SiN_x , we conclude that D_{III} scales down as the Ga vacancy concentration is suppressed by Ga interstitial injection from the interface with a high quality SiO_x film.

5.4.4. Cap summary

A high density glass appears to effectively cut off the exchange of atoms between the GaAs and the vapor, and the native defect concentrations are no longer affected by the external ambient. D_{III} , shown throughout Section 5 to be set by a vacancy mechanism, is consistently reduced under a dense glass. Does this mean that (1) the glass–GaAs interface acts as a sink for vacancies, (2) the glass–GaAs interface undergoes metallurgical reactions which inject Ga interstitials (which then reduce the vacancy concentration), or (3) the glass is nonreactive and all of the defects come from the substrate (which is almost never characterized)? The answer cannot be determined simply from the behavior

of D_{III} . However, the results of Haddara et al. indicate that both glasses increase diffusion of grown-in Be markers. Since large values of D_{II} have always been associated with large nonequilibrium concentrations of Ga interstitials, one is led to the tentative conclusion that metallurgical reactions at the glass-GaAs interface cause an injection of I_{Ga} into the crystal. Clearly, further work is needed to improve our understanding of the point defect sources.

6. Summary

This review of diffusion in thin film GaAs began by clarifying several fundamental relationships between the equilibrium native defect concentrations and the external ambient which have not been discussed elsewhere. In particular, we have related impurity diffusion and the native defect concentrations to the independent experimental variables. We have shown that the Gibbs phase rule provides the number of thermodynamically independent variables for a given experimental design. Control over this number of independent variables will satisfy the *necessary condition* to define the equilibrium native defect concentrations. The *sufficient condition*, for actually attaining those concentrations, is satisfied by a well-designed experiment which quickly brings the dependent thermodynamic variables to their equilibrium values.

Experiments have been discussed in which a *native defect equilibrium* approximation appears to be reasonable. When this approximation holds, then it has been shown that the so-called 'Fermi-level model' of diffusion is a natural consequence of the native defect concentrations set by solid-vapor equilibrium. However, this approximation is not automatically valid, and inappropriate application of this model has led to substantial confusion in the literature. We have presented evidence from well-designed experiments using n-type, intrinsic, and p-type GaAs which all strongly argue that diffusion occurs via a singly charged Ga vacancy, V_{Ga}^- , over all practical values of annealing temperature and arsenic partial pressure. The evidence also strongly argues that the interstitialcy mechanism does not occur in any experiment where native equilibrium can be approximated.

There is good evidence that the interstitialcy, or so-called 'kick-out', mechanism occurs under strongly nonequilibrium conditions, i.e. in the presence of an excess Ga interstitial concentration. Many results can be explained by assuming that the group II interstitial concentration equilibrates primarily with the group II dopant partial pressure rather than the dopant concentration in the GaAs. Such an assumption explains how the interstitialcy mechanism can cause the very wide range of values measured for D_{II} and D_{III} during dopant in-diffusion, dopant out-diffusion, and diffusion in buried layers (emitter push effect). Evidence has also been presented that the grown-in dopant concentration often exceeds the solubility limit during annealing. For such cases, there is a substantial driving force for the diffusion of dopant atoms, and that diffusivity may be dramatically accelerated as native defect concentrations are varied via seemingly minor changes in the ambient.

It has been shown that the measurement noise, i.e. unexplained fluctuations in D_{Al-Ga} observed both within and between research groups, often approaches an order of magnitude. Although there are relatively few groups using In diffusion markers, it appears that the measurement noise within each group is relatively low and approaches a more reasonable measurement noise, i.e. a factor of 2. This level of noise is similar to that measured for the isotopic self-diffusion measurements of D versus $1/T$ in GaAs, and also similar to that measured for several metallurgical systems. The noise associated with Al-containing GaAs strongly suggests that an uncontrolled variable exists in much of the published interdiffusion work. We have suggested that small quantities of O_2 or H_2O may have entered ampoules through flame processing of the quartz or via the added As or Ga. We have summarized results from ampoule diffusion, and found values for D_{II} to be consistently less than the corresponding D_{III} values

measured in open tube anneals under the reducing gas, H₂. Until more carefully conducted studies show otherwise, we tentatively conclude that an open tube, proximity capped, anneal with H₂ present is the cleanest practical system available.

Reproducible low rates of marker interdiffusion can be achieved during annealing by capping the GaAs by high density SiO₂ or Si₃N₄. Comparing the diffusion of group II and III elements, we have tentatively concluded that metallurgical reactions occur between the GaAs and both glasses during annealing. The important result of these reactions appears to be the nonequilibrium injection of Ga interstitials into GaAs, and thus the reduction of the (vacancy controlled) group III marker diffusion.

Acknowledgements

I would like to thank the University of Utah for providing support for the sabbatical leave which provided the time to think through and write this review. Thanks are also given to The Australian National University, and its Electronic Materials Engineering Department and its entire staff, for making support facilities available during that sabbatical leave. Particular thanks go to Professors Jagadish and Williams in Australia, and to Professor Virkar in Utah for excellent discussions on the topic of point defects and thermodynamics. Thanks also go to Professor Kendall at New Mexico for inspiring me, through his writing and occasional verbal comments, to think of the right questions. Of course, my students deserve unlimited thanks for doing all of the hard work, and that work could never have reached a logical conclusion without the painstaking efforts of Drs David Simons and Peter Chi at NIST who obtained so many carefully taken SIMS profiles. Grateful thanks are also given to Drs. Olmsted and Houde-Walter for making unpublished data available, and to Drs. Haddara and Gillen for their interaction. The support of both the NSF and the ARO has also made this work possible and is gratefully appreciated. Finally, I must thank Judy and Jessi for putting up with me while this was being written.

References

- [1] D.G. Deppe, N. Holonyak, Jr., *J. Appl. Phys.* 64 (1988) R93.
- [2] H.C. Casey, Jr., G.L. Pearson, in: J.H. Crawford, L.M. Slifkin (eds), *Point Defects in Solids*, vol. 2, Plenum, New York, 1975, p. 163.
- [3] D.L. Kendall, in: R.K. Willardson, A.C. Beer (eds.), *Semiconductors and Semimetals*, vol. 4, Academic Press, New York, 1968, p. 222.
- [4] D. Shaw, in: D. Shaw (ed.), *Atomic Diffusion in Semiconductors*, Plenum, New York, 1973, p. 1.
- [5] T.Y. Tan, U. Gosele, S. Yu, *Crit. Rev. Solid State Mater. Sci.* 17 (1991) 47.
- [6] R.M. Cohen, *J. Appl. Phys.* 67 (1990) 7268.
- [7] F.A. Kroger, *The Chemistry of Imperfect Crystals*, 2nd ed., vol. 2, North-Holland, Amsterdam, 1974, p. 237ff.
- [8] T.Y. Tan, U. Gosele, *Mater. Sci. Eng.* B1 (1988) 47.
- [9] J.R. Arthur, *J. Phys. Chem. Solids* 28 (1967) 2257.
- [10] B.L. Olmsted, S.N. Houde-Walter, *Appl. Phys. Lett.* 63 (1993) 1131.
- [11] M.B. Panish, *J. Cryst. Growth* 27 (1974) 6.
- [12] K.Y. Hsieh, Y.L. Hwang, J.H. Lee, R.M. Kolbas, *J. Electron. Mater.* 19 (1990) 1417.
- [13] C.N. Cochran, L.M. Foster, *J. Electrochem. Soc.* 109 (1962) 149.
- [14] C.D. Thurmond, G.P. Schwartz, G.W. Kammlott, B. Schwartz, *J. Electrochem. Soc.* 127 (1980) 1366.
- [15] O.R. Monteiro, J.W. Evans, *J. Vac. Sci. Technol.* A7 (1989) 49.
- [16] G. Hughes, R. Ludeke, *J. Vac. Sci. Technol.* B4 (1986) 1109.
- [17] H.D. Palfrey, M. Brown, A.F.W. Willoughby, *J. Electrochem. Soc.* 128 (1981) 2224.
- [18] J.M. Dallesasse, N. El-Zein, N. Holonyak, Jr., K.C. Hsieh, R.D. Burnham, R.D. Dupuis, *J. Appl. Phys.* 68 (1990) 2235.
- [19] D.L. Huffacker, D.G. Deppe, K. Kumar, T.J. Rogers, *Appl. Phys. Lett.* 65 (1994) 97.
- [20] M.J. Ries, N. Holonyak, Jr., E.I. Chen, S.A. Maranowski, M.R. Islam, A.L. Holmes, R.D. Dupuis, *Appl. Phys. Lett.* 67 (1995) 1107.

- [21] W.R. Reichert, C.Y. Chen, W.M. Li, J.E. Shield, R.M. Cohen, D.S. Simons, P.H. Chi, *J. Appl. Phys.* 77 (1995) 1902.
- [22] C.Y. Chen, Ph.D. Thesis, University of Utah, 1995.
- [23] C.J. Frosch, L. Derick, *J. Electrochem. Soc.* 105 (1958) 695.
- [24] D.W. Kisker, P.H. Fuoss, K.L. Tokuda, G. Renaud, S. Brennan, J.L. Kahn, *Appl. Phys. Lett.* 56 (1990) 2025.
- [25] F. Rosenberger, *Fundamentals of Crystal Growth I*, Springer, New York, 1979, p. 234.
- [26] F.A. Kroger, *The Chemistry of Imperfect Crystals*, 2nd edn., vol. 2, North-Holland, Amsterdam, 1974, p. 145ff.
- [27] J. Crank, *The Mathematics of Diffusion*, 2nd edn., Oxford University Press, London, 1975.
- [28] Paul Shewmon, *Diffusion in Solids*, 2nd edn., The Minerals, Metals, and Materials Society, Warrendale, PA, 1989, p. 31.
- [29] S.M. Hu, *Phys. Status Solidi B* 60 (1973) 595.
- [30] R.W. Kaliski, D.W. Nam, D.G. Deppe, N. Holonyak, Jr., K.C. Hsieh, R.D. Burnham, *J. Appl. Phys.* 62 (1987) 998.
- [31] C.J. Nuese, G.E. Stillman, M.D. Sirkis, N. Holonyak, Jr., *Solid State Electron.* 9 (1966) 735.
- [32] Y. Matsumoto, *Japan. J. Appl. Phys.* 22 (1983) 1699.
- [33] U. Gosele, F. Morehead, *J. Appl. Phys.* 52 (1981) 4617.
- [34] R.M. Cohen, *J. Electron. Mater.* 20 (1991) 425.
- [35] H.C. Casey, Jr., M.B. Panish, *Trans. Metall. Soc. AIME*, 242 (1968) 406.
- [36] K.K. Shih, *J. Electrochem. Soc.* 123 (1976) 1737.
- [37] F.A. Kroger, *The Chemistry of Imperfect Crystals*, 2nd edn., vol. 1, North-Holland, Amsterdam, 1973, p. 246.
- [38] L.R. Weisberg, J. Blanc, *Phys. Rev.* 131 (1963) 1548.
- [39] H.C. Casey, Jr., M.B. Panish, L.L. Chang, *Phys. Rev.* 162 (1967) 660.
- [40] G.A. Baraff, M. Schluter, *Phys. Rev. Lett.* 55 (1985) 1327.
- [41] T.Y. Tan, S. Yu, U. Gosele, *J. Appl. Phys.* 70 (1991) 4823.
- [42] C.Y. Chen, R.M. Cohen, *J. Cryst. Growth* 167 (1996) 17.
- [43] W. Jager, A. Rucki, K. Urban, H.G. Hettwer, N.A. Stolwijk, H. Mehrer, T.Y. Tan, *J. Appl. Phys.* 74 (1993) 4409.
- [44] W. Shockley, J.T. Last, *Phys. Rev.* 107 (1957) 392.
- [45] D. Hurle, personal communication.
- [46] S. Poykko, M.J. Puska, R.M. Nieminen, *Phys. Rev. B* 53 (1996) 3813.
- [47] H. Seong, L.J. Lewis, *Phys. Rev. B* 52 (1995) 5675.
- [48] J. Dabrowski, J.E. Northrup, *Phys. Rev. B* 49 (1994) 14286.
- [49] J.E. Northrup, S.B. Zhang, *Phys. Rev. B* 50 (1994) 4962.
- [50] J.E. Northrup, S.B. Zhang, *Phys. Rev. B* 47 (1993) 6791.
- [51] S.B. Zhang, J.E. Northrup, *Phys. Rev. Lett.* 67 (1991) 2339.
- [52] C. Wang, Q.M. Zhang, J. Bernholc, *Phys. Rev. Lett.* 69 (1992) 3789.
- [53] J. Bernholc, B. Chen, Q.M. Zhang, C. Wang, S.K. Kajihara, D. Sullivan, in: D.J. Lockwood (ed.), *Proc. 22nd Conf. Phys. Semicond.*, World Scientific, Singapore, 1995, p. 2259.
- [54] Y.M. Haddara, C.C. Lee, J.C. Hu, M.D. Deal, J.C. Bravman, *MRS Bull.* 20 (4) (1995) 41.
- [55] R.M. Cohen, C.Y. Chen, D.S. Simons, P.H. Chi, in: *Defect- and Impurity-Engineered Semiconductors and Devices*, in: S. Ashok, J. Chevallier, I. Akasaki, N.M. Johnson, B.L. Sopori (eds.), *Proc. 1995 MRS Meet.*, vol. 378, Materials Research Society, Pittsburgh, PA, 1995, p. 875.
- [56] K. Okamoto, H. Mawatari, K. Yamaguchi, A. Noguchi, *J. Cryst. Growth* 98 (1989) 630.
- [57] W.M. Li, C.Y. Chen, R.M. Cohen, *J. Cryst. Growth* 156 (1995) 343.
- [58] G. Li, M. Linnarsson, C. Jagadish, *J. Cryst. Growth* 154 (1995) 231.
- [59] H.C. Casey, Jr., M.B. Panish, *J. Crystal Growth*, 13–14, 818 (1972).
- [60] K.H. Zschauer, A. Vogel, *Proc. Symp. GaAs*, 1970 Inst. of Physics, London, 1971, p. 100.
- [61] P.D. Kirchner, T.N. Jackson, G.D. Pettit, J.M. Woodall, *Appl. Phys. Lett.* 47 (1985) 26.
- [62] P. Enquist, J.A. Hutchby, T.J. de Lyons, *J. Appl. Phys.* 63 (1988) 4485.
- [63] P.M. Enquist, *J. Cryst. Growth* 93 (1988) 637.
- [64] W.S. Hobson, S.J. Pearson, A.S. Jordan, *Appl. Phys. Lett.* 56 (1990) 1251.
- [65] D.G. Deppe, *Appl. Phys. Lett.* 56 (1990) 370.
- [66] C.Y. Chen, R.M. Cohen, D.S. Simons, P.H. Chi, *Appl. Phys. Lett.* 67 (1995) 1402.
- [67] B. Goldstein, *Phys. Rev.* 121 (1961) 1305.
- [68] L. Wang, L. Hsu, E.E. Haller, J.W. Erickson, A. Fischer, K. Eberl, M. Cardona, *Phys. Rev. Lett.* 76 (1996) 2342.
- [69] T.Y. Tan, H.M. You, S. Yu, U.M. Gosele, W. Jager, D.W. Boeringer, F. Zyprman, R. Tsu, S.T. Lee, *J. Appl. Phys.* 72 (1992) 5206.
- [70] A. Furuya, M. Makiuchi, O. Wada, T. Fujii, H. Nobuhara, *Jpn. J. Appl. Phys.* 26, L926 (1987).
- [71] L.J. Guido, N. Holonyak, Jr., K.C. Hsieh, R.W. Kaliski, W.E. Plano, R.D. Burnham, R.L. Thornton, J.E. Epler, T.L. Paoli, *J. Appl. Phys.* 61 (1987) 1372.
- [72] B.L. Olmsted, S.N. Houde-Walter, *Appl. Phys. Lett.* 60 (1992) 368.
- [73] B.L. Olmsted, S.N. Houde-Walter, *Appl. Phys. Lett.* 63 (1993) 530.
- [74] N. Baba-Ali, I. Harrison, B. Tuck, *J. Mater. Sci. Mater. Electron.* 6 (1995) 127.
- [75] S. Seshadri, L.J. Guido, P. Mitev, *Appl. Phys. Lett.* 67 (1995) 497.
- [76] K.Y. Hsieh, Y.C. Lo, J.H. Lee, R.M. Kolbas, in: J.S. Harris (ed.), *Gallium Arsenide and Related Compounds 1988*, vol. 15, IOP, Atlanta, GA, 1988, p. 393, as taken from B.L. Olmsted, S.N. Houde-Walter, *Appl. Phys. Lett.* 60 (1992) 368.

- [77] L.J. Guido, N. Holonyak, Jr., K.C. Hsieh, in: J.S. Harris (ed.), Gallium Arsenide and Related Compounds 1988, vol. 15, IOP, Atlanta, GA, 1988, p. 353, as taken from B.L. Olmsted, S.N. Houde-Walter, *Appl. Phys. Lett.* 60 (1992) 368.
- [78] B.L. Olmsted, S. Houde-Walter, personal communication 1996.
- [79] P. Mei, H.W. Yoon, T. Venkatesan, S.A. Schwarz, J.P. Harbison, *Appl. Phys. Lett.* 50 (1987) 1823.
- [80] P.M. Petroff, *J. Vac. Sci. Technol.* 14 (1977) 973.
- [81] R.M. Fleming, D.B. McWhan, A.C. Gossard, W. Wiegmann, R.A. Logan, *J. Appl. Phys.* 51 (1980) 357.
- [82] J. Cibert, P.M. Petroff, D.J. Werder, S.J. Pearton, A.C. Gossard, J.H. English, *Appl. Phys. Lett.* 49 (1986) 223.
- [83] T.E. Schlesinger, T. Kuech, *Appl. Phys. Lett.* 49 (1986) 518.
- [84] W. Blakemore, *J. Appl. Phys.* 53 (1982) 520.
- [85] D.L. Huffaker, D.G. Deppe, K. Kumar, T.J. Rogers, *Appl. Phys. Lett.* 65 (1994) 97.
- [86] T.Y. Tan, H.M. You, U.M. Gosele, *Appl. Phys. A56* (1993) 249.
- [87] P. Mei, S.A. Schwarz, T. Venkatesan, C.L. Schwartz, E. Colas, *J. Appl. Phys.* 65 (1989) 2165.
- [88] R.M. Cohen, C.Y. Chen, D.S. Simons, P.H. Chi, in: S. Ashok, J. Chevallier, I. Akasaki, N.M. Johnson, B.L. Sopori (eds.), Defect- and Impurity-Engineered Semiconductors and Devices, Proc. 1995 MRS Meet., vol. 378, Materials Research Society, Pittsburgh, PA, 1995, p. 959.
- [89] W.M. Li, R.M. Cohen, D.S. Simons, P.H. Chi, *Appl. Phys. Lett.* 70 (1997) 3392.
- [90] I. Szafranek, M. Szafranek, J.S. Major, Jr., B.T. Cunningham, L.J. Guido, N. Holonyak, Jr., G.E. Stillman, *J. Electronic Mater.* 20 (1991) 409.
- [91] H.M. You, T.Y. Tan, U.M. Gosele, S.T. Lee, G.E. Hofler, K.C. Hsieh, N. Holonyak, Jr., *J. Appl. Phys.* 74 (1993) 2450.
- [92] C.H. Ting, G.L. Pearson, *J. Appl. Phys.* 42 (1971) 2247.
- [93] L.L. Chang, G.L. Pearson, *J. Appl. Phys.* 35 (1964) 1960.
- [94] M.A.H. Kadhim, B. Tuck, *J. Mater. Sci.* 7 (1972) 68.
- [95] S. Yu, T.Y. Tan, U. Gosele, *J. Appl. Phys.* 69 (1991) 3547.
- [96] R.L. Longini, *Solid State Electron.* 5 (1962) 127.
- [97] L.L. Chang, A. Koma, *Appl. Phys. Lett.* 29 (1976) 138.
- [98] M.D. Camras, N. Holonyak, Jr., R.D. Burnham, W. Streifer, D.R. Scifres, T.L. Paoli, C. Lindstrom, *J. Appl. Phys.* 54 (1983) 5637.
- [99] B. Olmsted, S. Houde-Walter, personal communication 1996.
- [100] S. Seshadri, L.J. Guido, T.S. Moise, J.C. Beggy, T.J. Cunningham, R.C. Barker, *J. Electronic Mater.* 21 (1992) 33.
- [101] K.J. Beernink, D. Sun, D.W. Treat, B.P. Bour, *Appl. Phys. Lett.* 66 (1995) 3597.
- [102] N. Baba-Ali, I. Harrison, B. Tuck, H.P. Ho, M. Henini, O.H. Hughes, *J. Mater. Sci. Mater. Electron.* 1 (1990) 133.
- [103] W.P. Gillin, I.V. Bradley, L.K. Howard, R. Gwilliam, K.P. Homewood, *J. Appl. Phys.* 73 (1993) 7715.
- [104] Y.M. Haddara, M.D. Deal, J.C. Bravman, *Appl. Phys. Lett.* 68 (1996) 1939.
- [105] Y.M. Haddara, M.D. Deal, H.G. Robinson, J.C. Bravman, Proc. SOTAPOCXXI, Electrochemical Society, Pennington, NJ, 1995, p. 96.
- [106] Y.M. Haddara, M.D. Deal, J.C. Bravman, in Proc. 4th Int. Symp. Process Physics and Modeling, Electrochemical Society, Pennington, NJ, 1996, p. 142.
- [107] W.J. Taylor, N. Kuwata, I. Yoshida, T. Katsuyama, H. Hayashi, *J. Appl. Phys.* 73 (1993) 8653.
- [108] W.P. Gillin, D.J. Dunstan, K.P. Homewood, L.K. Howard, B.J. Sealy, *J. Appl. Phys.* 73 (1993) 3782.
- [109] I.V. Bradley, W.P. Gillin, K.P. Homewood, R.P. Webb, *J. Appl. Phys.* 73 (1993) 1686.
- [110] W.J. Choi, S. Lee, J. Zhang, Y. Kim, S.K. Kim, J.I. Lee, K.N. Kang, K. Cho, *Jpn. J. Appl. Phys.* 34, L418 (1995).
- [111] K.J. Beernink, R.L. Thornton, G.B. Anderson, M.A. Emanuel, *Appl. Phys. Lett.* 66 (1995) 2522.
- [112] J.D. Ralston, M. Ramsteiner, B. Dischner, M. Maier, G. Brandt, P. Koidl, D.J. As, *J. Appl. Phys.* 70 (1991) 2195.
- [113] J.S. Major, Jr., F.A. Kish, T.A. Richard, A.R. Sugg, J.E. Baker, N. Holonyak, Jr., *J. Appl. Phys.* 68 (1990) 6199.
- [114] J.C. Lee, T.E. Schlesinger, T.F. Kuech, *J. Vac. Sci. Technol.* B5 (1987) 1187.
- [115] W.P. Gillin, I.V. Bradley, S.S. Rao, K.P. Homewood, B.J. Sealy, L.K. Howard, A.D. Smith, A.T.R. Briggs, *Mater. Sci. Eng.* B21 (1993) 281.
- [116] W. Gillin, personal communication 1996.
- [117] R.K. Kupka, Y. Chen, *Appl. Phys. Lett.* 67 (1995) 1612.
- [118] S. Burkner, M. Maier, E.C. Larkins, W. Rothmund, E.P. O'Reilly, J.D. Ralston, *J. Electronic Mater.* 24 (1995) 805.
- [119] J.D. Ralston, S. O'Brien, G.W. Wicks, L.F. Eastman, *Appl. Phys. Lett.* 52 (1988) 1511.
- [120] I. Gontijo, T. Krauss, J.H. Marsh, R.M. De La Rue, *IEEE J. Quantum Electron.* 30 (1994) 1189.