Dopant diffusion in Si and SiGe

Doctoral Thesis by

Jens S. Christensen

Material and Semiconductor Physics
Stockholm 2004
Dopant diffusion in Si and SiGe
Jens S. Christensen

A dissertation submitted to the Royal Institute of Technology, Stockholm, Sweden, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ISRN KTH/FTE/FR-2004/2-SE
ISSN 0284-0545
TRITA – FTE
Forskningsrapport 2004:2

Printed by Universitetsservice US AB, Stockholm 2004
Abstract

Dopant diffusion in semiconductors is an interesting phenomenon from both technological and scientific points of view. Firstly, dopant diffusion is taking place during most of the steps in electronic device fabrication and, secondly, diffusion is related to fundamental properties of the semiconductor, often controlled by intrinsic point defects: self-interstitials and vacancies. This thesis investigates the diffusion of P, B and Sb in Si as well as in strained and relaxed SiGe. Most of the measurements have been performed using secondary ion mass spectrometry on high purity epitaxially grown samples, having in-situ incorporated dopant profiles, fabricated by reduced pressure chemical vapor deposition or molecular beam epitaxy. The samples have been heat treated both under close-to-equilibrium conditions (i.e., long time annealings in an inert ambient) and conditions which resulted in non-equilibrium diffusion (i.e., vacuum annealing, oxidation, short annealing duration, and proton irradiation).

Equilibrium P and B diffusion coefficients in Si as determined in this thesis differ from a substantial part of previously reported values. This deviation may be attributed to slow transients before equilibrium concentrations of point defects are established, which have normally not been taken into account previously. Also an influence of extrinsic doping conditions may account for the scattering of the diffusivity values reported in literature. B and Sb diffusion in Si under proton irradiation at elevated temperatures was found to obey the so-called intermittent model. Parameters describing the microscopic diffusion process were derived in terms of the intermittent diffusion mechanism, and it was found also that the presence of Sb strongly affected the B diffusion and vice versa.

In relaxed Si$_{1-x}$Ge$_x$-alloys, which has the same lattice structure as Si but a larger lattice constant, P diffusion is found to increase with increasing Ge content ($x \leq 0.2$). In Si/SiGe/Si heterostructures, where the SiGe layer is biaxially strained in order to comply with the smaller lattice parameter of Si, P diffusion in the strained layer is retarded as compared with relaxed material having the same Ge content. In addition, P is found to segregate into the Si layer via the Si/SiGe interface and the segregation coefficient increases with increasing Ge content in the SiGe layer.
Contents

PUBLICATIONS .............................................................................................................. III
APPENDED PUBLICATIONS ........................................................................................... III
PUBLICATIONS NOT INCLUDED ...................................................................................... III

ACKNOWLEDGEMENTS .................................................................................................. V

CHAPTER 1 INTRODUCTION ............................................................................................ 1

CHAPTER 2 DOPANT DIFFUSION ..................................................................................... 3

2.1 THERMODYNAMICS AND PHENOMENOLOGY OF DIFFUSION .................................. 3
2.2 MECHANISMS OF DIFFUSION .................................................................................... 5
  2.2.1 Intermittent diffusion ......................................................................................... 7
  2.2.2 Fickian diffusion ............................................................................................... 8
2.3 THE DIFFUSION COEFFICIENT .............................................................................. 10
2.4 POINT DEFECTS ....................................................................................................... 13

2.5 PERTURBATIONS IN POINT DEFECT CONCENTRATIONS ....................................... 15
2.6 EFFECT OF HYDROSTATIC PRESSURE ..................................................................... 17
2.7 DIFFUSION IN HETEROSTRUCTURES ...................................................................... 18
  2.7.1 Composition effect ............................................................................................ 18
  2.7.2 Segregation ....................................................................................................... 20
  2.7.3 Strain ................................................................................................................ 20

CHAPTER 3 SAMPLE PREPARATION AND CHARACTERIZATION ........................................ 23

3.1 SAMPLE PREPARATION ........................................................................................... 23
3.2 HIGH RESOLUTION X-RAY DIFFRACTION ............................................................... 25
  3.2.1 $\omega/2\theta$-scan ................................................................................................. 27
  3.2.2 Reciprocal lattice map ..................................................................................... 30
3.3 SECONDARY ION MASS SPECTROMETRY ................................................................. 31

CHAPTER 4 DOPANT DIFFUSION IN SI ........................................................................... 35

4.1 EQUILIBRIUM DIFFUSION ....................................................................................... 35
  4.1.1 Phosphorus diffusion ...................................................................................... 35
  4.1.2 Boron diffusion ............................................................................................... 42
4.2 NON-EQUILIBRIUM DIFFUSION .......................................................................... 44
  4.2.1 Surface reactions .............................................................................................. 44
  4.2.2 Ion implantation .............................................................................................. 46

CHAPTER 5 DOPANT DIFFUSION IN SIGE ...................................................................... 52

5.1 COMPOSITION EFFECT ON DOPANT DIFFUSION ................................................. 52
5.2 EFFECT OF STRAIN .................................................................................................. 56
5.3 DIFFUSION MECHANISM ....................................................................................... 57
5.4 P SEGREGATION IN HETEROSTRUCTURE INTERFACES .......................................... 60
5.5 P DIFFUSION IN SIGE WITH IMPERFECT CRYSTALLINITY .................................... 61

CHAPTER 6 SUMMARY OF PAPERS .............................................................................. 64
Publications

Appended publications

   Phosphorus diffusion in Si; influence of annealing conditions Mat. Res.

II. J. S. Christensen, A. Yu. Kuznetsov, H. H. Radamson, B. G. Svensson
    Phosphorus and Boron Diffusion in Silicon Under Equilibrium Conditions

III. J. S. Christensen, A. Yu. Kuznetsov, H. H. Radamson, B. G. Svensson
     Phosphorus Diffusion in Strained and Relaxed Si$_{1-x}$Ge$_x$ J. Appl. Phys. 94,
     6533 (2003)

IV. P. Lévêque, J. S. Christensen, A. Yu. Kuznetsov, B. G. Svensson, A.
    Nylandsted-Larsen Influence of boron concentration on the enhanced
diffusion observed after irradiation of boron delta-doped silicon at 570°C

V. P. Lévêque, A. Yu. Kuznetsov, J. S. Christensen, B. G. Svensson, A.
   Nylandsted-Larsen Irradiation enhanced diffusion of boron in delta-doped

VI. P. Lévêque, J. S. Christensen, A. Yu. Kuznetsov, B. G. Svensson, A.
    Nylandsted-Larsen Influence of boron on radiation enhanced diffusion of

Publications not included

J. S. Christensen, A. Yu. Kuznetsov, H. H. Radamson, B. G. Svensson
Phosphorus diffusion in Si$_{1-x}$Ge$_x$ Defect and Diffusion Forum vols. 194-199 pp. 709-716 (2001)

Radamson, J. Grahn, G. Landgren Diffusion of phosphorus in strained Si/SiGe/Si

M. S. Janson, M. K. Linnarson, J. S. Christensen, P. Lévêque, A. Yu. Kuznetsov, H.
H. Radamson, A. Hallén, A. Nylandsted-Larsen, B. G. Svensson Diffusion of dopants
and impurities in device structures of SiC, SiGe and Si Defect and Diffusion Forum vols. 194-199 pp. 597-610 (2001)


Acknowledgements
Chapter 1
Introduction

Diffusion can, in short, be described as a process resulting from random motion of particles, a motion which results in a redistribution of particles from regions with high concentration to regions with low concentration. Similar to a drop of milk in a cup of coffee, the milk will slowly dissolve in the coffee until it is practically invisible.

Diffusion in solid-state materials has been studied systematically for a little more than a hundred years \[1, 2\]. The early studies were made in connection with fabrication of coins, where both material and instruments were available. Naturally, they dealt with diffusion of metals in metals. With the birth of the semiconductor technology, materials such as germanium and silicon attracted much attention. Especially, with the development of semiconductor devices such as diodes and transistors, which rely on semiconductor materials doped with impurities in confined regions in the semiconductor, dopant diffusion has been important \[3\].

The current development, towards smaller and faster devices, forces the developers to look into other materials than silicon. A natural choice is a Si\(_{1-x}\)Ge\(_x\) alloy (x denotes the relative Ge concentration in the alloy), which is relatively easy and cheap to incorporate into existing standard Si processes \[4\]. Si\(_{1-x}\)Ge\(_x\) has the same lattice structure as Si, but its lattice constant increase with increasing x, and the band gap decreases with increasing x. The fact that the band gap varies with x opens for the possibility of band gap engineering of devices, for example, an electric field built into a device with a graded SiGe layer. Most devices using SiGe are based on silicon on which a layer of SiGe is grown. As Si and SiGe have different lattice parameters, the SiGe layer will be biaxially compressed or strained in order to accommodate the lattice of the substrate. This strain may completely change the material and electronic properties compared with the unstrained or relaxed SiGe. An example of such a device is the heterojunction bipolar nnp-transistor with a SiGe base. Figure 1.1 shows a secondary ion mass spectrometry (SIMS) measurement of the dopant profiles in such a transistor. It consists of an emitter with a high phosphorus concentration, a B-doped SiGe base, and a P and As doped collector. The advantage of using SiGe in the base instead of Si, is that the potential barrier for electron injection from the emitter to the base of the transistor is lowered due to the smaller SiGe bandgap. The lower bandgap cause an exponentially enhanced current gain for a constant emitter base voltage drop.
Diffusion of P into the base region will degrade the performance of the device. Similarly, for B diffusion out of the base. During the fabrication of the device, some diffusion is inevitable, so device manufacturers need to know how the dopants will diffuse to account for their redistribution during the fabrication processes.

Consequently, much research has been done on all relevant dopants in silicon. However, only a few elements in SiGe have been studied, and the results from these studies have not been fully understood.

Dopant diffusion in Si is controlled by intrinsic point defects; self-interstitials and vacancies, through the microscopic diffusion mechanism. Thus, a study of diffusion will provide information of fundamental thermodynamic properties of the semiconductor.

The present work is mainly focused on an experimental determination of phosphorus and boron diffusion in Si and SiGe. The results of these investigations will be related to other results from similar systems. This thesis gives a general introduction to the phenomenology of dopant diffusion in Si and SiGe in Chapter 2, Chapter 3 deals with experimental aspects of determining diffusion. Chapter 4 and 5 present the results of P and B diffusion in Si and in SiGe, respectively.
Chapter 2
Dopant diffusion

2.1 Thermodynamics and phenomenology of diffusion

The Gibb’s free energy of a crystalline material will change, when impurities are included in the system. The energy change $\Delta G_A$ for a material with a density of lattice sites $N$, which is $\sim 5 \times 10^{22}$ cm$^{-3}$ for Si, doped with an impurity $A$, e.g., phosphorus, with a concentration $C_A$ is given by

$$\Delta G_A = C_A(x)g^f - k_BT \ln \left( \frac{N!}{(N-C_A(x))!C_A(x)!} \right),$$

(2.1)

where $g^f_A$ is the Gibb’s free energy of formation for the single impurity, $x$ is the distance in a 1-dimensional model, $k_B$ and $T$ are Boltzmann’s constant and absolute temperature, respectively. The second term in equation (2.1) is a configurational entropy term given by the number of ways the impurity atoms can be distributed in given number of lattice sites. A calculation of $g^f_A$ is a complicated matter, as it contains changes in internal energy and vibrational entropy associated with the exchange of a host atom with an impurity atom, as well as energy related to electrical effects of the dopant [6, 7]. In equation (2.1), $C_A$ is given a spatial variation in one dimension. This can be extended to three dimensions, but usually only the direction perpendicular to the materials surface is of interest, as in the example of the transistor structure (Figure 1.1).

The chemical potential $\mu_A$ can be calculated as the derivative of the Gibb’s free energy with respect to the concentration of the element:

$$\mu_A = \frac{\partial G_A}{\partial C_A} = g^f + k_BT \ln \left( \frac{C_A}{N-C_A} \right) = g^f + k_BT \ln \left( \frac{C_A}{N} \right),$$

(2.2)

where we assume that $N >> C_A$ and use Stirling’s formula on the second term in equation (2.1). The explicit spatial variation is omitted for simplicity and will be so henceforward. The chemical potential of an element can be viewed as a generalized force, in analogy with an electrostatic potential. At a given concentration of element $A$
the Gibb’s free energy exhibit a minimum, and the chemical potential is zero. This concentration is called the solid solubility, $C^\text{sol}_A$, of the element and is determined by

$$C^\text{sol}_A = N \exp\left(- \frac{g^f}{k_B T}\right). \quad (2.3)$$

Thus, equation (2.2) can be rewritten as

$$\mu_A = k_B T \ln\left(\frac{C_A}{C^\text{sol}_A}\right). \quad (2.4)$$

Consequently, if $C_A$ has a spatial variation, then so will the chemical potential and this will constitute a non-equilibrium situation from a thermodynamical viewpoint. In order to restore equilibrium, i.e. reach a constant chemical potential, dopants will redistribute until the chemical potential is constant. The impurities $A$ will experience an exerted force $f$ given by

$$f = -\left.\frac{\partial \mu_A}{\partial x}\right|_{\text{const}} = -k_B T \left(\frac{1}{C_A} \frac{\partial C_A}{\partial x}\right). \quad (2.5)$$

working to restore a constant chemical potential. It is assumed that the solid solubility is $x$-independent. In section 2.7.2, the case of a varying solubility is considered. The force will cause a flux $J_A$ of impurities:

$$J_A = C_A v_A = C_A M_A f = -k_B T M_A \frac{\partial C_A}{\partial x}, \quad (2.6)$$

where $v_A$ is the velocity of the impurities given by the product of the exerted force and $A$’s mobility $M_A$. Usually, diffusion is described by the diffusion coefficient $D$ instead of the mobility. $M$ and $D$ are related through the Einstein relation: $M = D/k_B T$. Then equation (2.6) may be written as

$$J_A = -D_A \frac{\partial C_A}{\partial x}. \quad (2.7)$$

The diffusion coefficient $D_A$ expresses the amount of particles that diffuse across a unit area in one time unit, when the gradient is 1 unit. Equation (2.7) is known as Fick’s first law, a generalization of Fick’s first law, where the flux depends on the gradient in the chemical potential rather than the gradient in impurity concentration, can be derived from equation (2.5) and (2.6).

Since particles are conserved in the diffusion process, a continuity equation can be set up, stating that at a given point the rate of change of the concentration is given by the flux:

$$\frac{\partial C_A}{\partial t} = \frac{\partial}{\partial x}(-J_A) = D_A \frac{\partial^2 C_A}{\partial x^2}, \quad (2.8)$$
where it is assumed that the diffusivity is independent on position and concentration. This is true for many cases, but there are exceptions, for example when diffusion takes place in a heavily doped region as will be discussed in Chapter 4.

The diffusion model given by equation (2.8) is known as Fick’s second law, named after the German scientist Adolf Fick who was the first to come up with the quantitative model of diffusion about 150 years ago [1]. With this model it is possible to predict how a given profile will evolve. There are only analytical solutions of the equation (2.8) in a few special cases, but the equation can be solved numerically for all kinds of profiles (see appendix A).

### 2.2 Mechanisms of diffusion

Knowledge of the mechanisms that govern dopant diffusion on a microscopic level are crucial for the understanding and modeling of diffusion, and it is well established that the important dopants in Si will diffuse by interacting with native point defects, such as vacancies (V) or Si self-interstitials (I), which are always present in the crystal. The impurity traps a point defect and forms a highly mobile complex that is able to move through the crystal, until the complex breaks up and the impurity again occupies a substitutional site. The so-called concerted exchange mechanism, where no point defect are required, because the dopant simply changes place with a neighboring Si-atom, is usually regarded to be negligible in silicon [2].

Typical reactions responsible for forming the mobile dopant-point defect complexes are:

\[
A + V \leftrightarrow AV, \quad (2.9a)
\]

\[
A + I \leftrightarrow AI \text{ or } A_i, \quad (2.9b)
\]

\[
A \leftrightarrow (A_i \text{ or } AI) + V, \quad (2.9c)
\]

where \(AV, AI\) are the impurity-point defect complexes and \(A_i\) is the impurity in an interstitial position. The first reaction (2.9a) is responsible for the vacancy mechanism of diffusion. Diffusion via interstitials can take place in two ways as represented by reaction (2.9b): the interstitialcy mechanism, where a substitutional impurity and a self-interstitial diffuse as an \(AI\) pair, or the kick-out mechanism where the impurity atom is kicked out from a lattice site and diffuse as an isolated interstitial \(A_i\) [8]. The distinction between \(AI\) and \(A_i\) is only relevant for theoretical considerations, experimentally it is not possible to distinguish between the kick-out and the interstitialcy mechanism. The last reaction (2.9c) is the so-called dissociative or Frank-Turnbull mechanism.

Generally, the impurity \(A\) may be either positively or negatively charged, for \(n\)- or \(p\)-type dopants, respectively, and the point defects may be charged as well or neutral.
Furthermore, the defects in the reactions \((2.9a) - (2.9c)\) may exchange electrons and/or holes with the conduction and/or valence bands, so the charge of the defect complexes is not necessarily given by the charge of its two components. This means that each of the above reactions may represent several reactions, for example the vacancy and the \(AV\) complex in \((2.9a)\) can in principle have charges varying independently between \(+2\) and \(−2\). So far, no systematic work has been done to determine the charge states of diffusing complexes, and in most cases below reference to charge states will be omitted. However, it should be kept in mind that different charge states might play a role in the diffusion processes.

The individual dopant-point defect complexes diffuse with a diffusivity \(D_{AI}\) or \(D_{AV}\), which is related to the total diffusivity \(D_A\) in \((2.8)\) in steady state by

\[
D_A = D_{AI} \frac{C_{AI}}{C_A} + D_{AV} \frac{C_{AV}}{C_A}.
\]

\(C_{AV}\) and \(C_{AI}\) are determined by the reaction coefficients for reactions \((2.9a) - (2.9c)\) and are usually much smaller than \(C_A\). The diffusion mechanism is quantified by a single parameter called the interstitialcy fraction, \(f_I\). The interstitialcy fraction is the ratio of the part of the mass transport taking place via interstitials to the total diffusion coefficient in steady state conditions:

\[
f_I = \frac{D_{AI} \frac{C_{AI}}{C_A}}{D_A}.
\]

For an impurity diffusing only by the interstitial mechanism \(f_I\) equals 1 while for an impurity diffusing only via the vacancies \(f_I\) equals 0.

A host of experimental works has shown that B and P diffuse in silicon predominantly through the interstitial mechanism, whereas Sb diffuses via the vacancy mechanism. Arsenic diffusion and Si self-diffusion are known to be combinations of the two mechanisms [9]. This thesis is mainly concerned with phosphorus and boron diffusion, so in the following we will develop the diffusion formalism only for the interstitial mechanism. All equations below can easily be extended to account for a dual mechanism diffusion, but on the expense of clarity.

The model given by equation \((2.8)\) is not always able to describe the diffusion. For instance, it may be necessary to consider the specific diffusion reaction \((2.9b)\) complexes are formed with a generation rate constant \(g\), break up with a recombination rate constant \(r\), and they diffuse with a diffusion coefficient \(D_{AI}\). Practically all diffusion happens through the complex, and the substitutional impurity is considered to be immobile. With this assumption a continuity equation for the complex concentration \(C_{AI}\) can be written as
$$\frac{\partial C_{Al}}{\partial t} = D_{Al} \frac{\partial^2 C_{Al}}{\partial x^2} - rC_{Al} + gC_A. \quad (2.12)$$

The concentration $C_{Al}$ depends on the concentration of Si self-interstitials $C_I$ through $g$, and the concentration of self-interstitials is connected to the concentration of vacancies, because vacancies and interstitials can recombine. Also $C_{Al}$ may depend directly on the vacancy concentration through the Frank-Turnbull mechanism (2.9c). A complete and general set of reaction-diffusion equations should account for all possible mechanisms, recombination between point defects and indirect recombination of point defects [8], however, under normal diffusion conditions equation (2.12) describe the diffusion fairly well.

The generation rate $g$ can be estimated by the probability for a substitutional $A$ to capture a self-interstitial, and the recombination $r$ is estimated by an attempt frequency $\nu$ times the probability for the $A$-complex to dissociate [10, 11]:

$$g = 4\pi a D_I C_I \exp \left( -\frac{E_{form}}{k_B T} \right), \quad r = \nu \exp \left( -\frac{E_{diss}}{k_B T} \right), \quad (2.13)$$

where $E_{form}$ and $E_{diss}$ are the barriers for the complex formation and dissociation, respectively. $D_I$ is the diffusion coefficient for Si self-interstitials, and $a$ is the capture radius, which is on the order of a few Å. From the formation and dissociation barriers a binding energy between the self-interstitial and the dopant can be defined: $E_{AI}^b = E_{diss} - E_{form}$. Contributions to the binding energy can be of electrostatic nature and/or related to a relaxation of the surrounding lattice.

For the total concentration of $A$, which is the sum of the complex concentration and the substitutional $A$ concentration the continuity equation will be

$$\frac{\partial (C_{Al} + C_A)}{\partial t} = D_{Al} \frac{\partial^2 C_{Al}}{\partial x^2} \quad (2.14)$$

The two last terms in equation (2.12) will govern the process if the diffusion time is short enough or if $g$ is small enough (e. g. if the concentration of Si self-interstitials are very low), so the complexes do not reach a steady state concentration. This diffusion mode is called intermittent diffusion. Another case is when a balance between the generation and recombination of the $AI$ complexes is reached, so the two last terms in equation (2.12) vanish. This mode is called Fickian diffusion.

### 2.2.1 Intermittent diffusion

For short diffusion times, i. e. $t < 1/g$, there is no equilibrium between the generation and recombination of the $AI$ complex. In that case each dopant will be in its fast diffusing state only a few times, and the diffusion will be given by the mean migration
distance of the individual complexes, \( \lambda = \sqrt{D_{\text{eff}}/r} \). This mode of diffusion is called intermittent diffusion. It is known to occur for boron diffusion in Si \([111]\). In this thesis more examples of the intermittent diffusion mode is demonstrated, e. g. P in relaxed SiGe (Section 5.5). Intermittent diffusion profiles are characterized by an exponential decay with a characteristic length given by \( \lambda \).

### 2.2.2 Fickian diffusion

When the diffusion times are so long that all dopants will participate in many diffusion steps, \( t \gg 1/g \), a stationary state between generation and recombination will prevail, and the two last terms in equation (2.12) vanish. Approximating the total amount of impurities \( C_A \) by the amount of substitutional impurities, equation (2.14) can be simplified to

\[
\frac{\partial C_A}{\partial t} = D_A \frac{\partial^2 C_A}{\partial x^2} = D_A \frac{\partial^2 C_A}{\partial x^2},
\]

(2.15)

where \( D_A \) is an effective diffusion coefficient. It is the value of \( D_A \) that is derived from diffusion experiments, rather than the diffusion coefficient of the fast diffusing species, since the measured profiles always show the depth dependence of the total amount of dopants \( C_A \). The effective diffusivity coefficient is coupled to the parameters characterizing the intermittent diffusion; \( D_{\text{eff}}, g, \) and \( \lambda \) by

\[
D_A = D_{\text{eff}} \frac{g}{r} = \lambda^2 g.
\]

(2.16)

---

1 For \( gt > 5 \) there is practically no difference in diffusion profiles simulated after the Fick model and the intermittent model.

---

Figure 2.1 Comparison between intermittent and Fickian diffusion. The as-grown profile (full line) and profile after 750 °C, 10 hour anneal (o), shown together with simulations using the Fick model with \( D_p = 4.3 \times 10^{-15} \text{ cm}^2/\text{s} \) (dotted line) and the intermittent model with \( g = 3 \times 10^{-5} \text{ s}^{-1} \) (\( gt = 1.08 \)), \( \lambda = 120 \text{ nm} \) (dashed line).
Figure 2.1 shows simulation results given by the Fick and the intermittent diffusion models. Parameters $g$, $\lambda$ and $D_A$ used in the intermittent and the Fick model respectively obey equation (2.16).

2.3 The diffusion coefficient

As shown above from a thermodynamic point of view, dopants will redistribute until a uniform concentration throughout the sample is reached. However, the time scale on which this occurs is given by the diffusion coefficient $D_A$, and it cannot be determined from thermodynamical, macroscopic considerations. Instead, it is necessary to consider what happens on the microscopic scale.

In case of diffusion through an interstitialcy mechanism, the diffusion process includes formation of an I, which diffuses until it meets a substitutional atom $A$, and forms a diffusing complex $AI$. This complex then moves through the lattice until it breaks up and the I is released. The diffusion coefficient $D_A$ depends on the Gibb’s free energy of this process $\Delta G_A$, and $D_A$ can be written as

$$D_A = K_1 \exp \left( -\frac{\Delta G_A}{k_B T} \right), \quad (2.17)$$

where $K_1$ is the product of several geometrical factors characterizing the lattice symmetry and an attempt frequency (see e.g. [12] and [3]). The Gibb’s free energy is usually split into a temperature dependent term and a temperature independent term, using enthalpy ($H$) and entropy ($S$):

$$\Delta G = \Delta H - T\Delta S. \quad (2.18)$$

The part of the exponential containing the entropy will be temperature independent and is usually included in the proportionality or pre-factor, so the only physical parameter in the temperature dependent part is the enthalpy. The temperature dependence of $D_A$ can be expressed with the entropies and enthalpies characterizing the diffusion process as

$$D_A = K_1 \exp \left( \frac{\Delta S_f + \Delta S_m}{k_B} \right) \exp \left( -\frac{\Delta H_f^i - E_a^b + \Delta H_m}{k_B T} \right) = D_A^0 \exp \left( -\frac{E_a}{k_B T} \right), \quad (2.19)$$

where superscripts "f" and "m" indicate formation and migration parameters. The sum in the numerator of the fraction in the exponential term is called the activation energy $E_a$ and it includes all the enthalpies required for making the dopant-self-interstitial complex and moving it. The temperature independent factors are usually lumped together into a single factor called the pre-exponential factor $D_A^0$. The two parameters $D_A^0$ and $E_a$ can be derived by measuring the temperature dependence of $D_A$ under conditions where the point defect concentrations are in equilibrium.
If charged defects are involved in the diffusion process, the diffusivity will be dependent on the Fermi-level \( E_F \), which can be given by the diffusing dopant itself, or by other dopants. The enthalpy of formation for a defect \( X^q \) with a charge \( q \) is given by

\[
\Delta H_{X^q}^f = \Delta H_{X^q}^{f0} + qE_F, \tag{2.20}
\]

where \( \Delta H_{X^q}^{f0} \) is the enthalpy of formation if the Fermi-level is coinciding with the valence band. Typically the charge will be between \(+2\) and \(-2\). When the enthalpy of formation depends on the Fermi-level, then so will the diffusivity according to equation (2.19). This dependence is usually given in terms of the concentration of free charge carriers (electron or holes), which is related to the Fermi-level by

\[
\frac{n}{n_i} = \exp\left(\frac{E_F - E_i}{k_B T}\right), \tag{2.21}
\]

where \( E_i \) and \( n_i \) are the Fermi-level position and carrier concentration in intrinsic material, respectively, while \( n \) is given by the concentration \( C \) of dopants and can be calculated from the mass action law assuming charge neutrality:

\[
n = \frac{1}{2}\left(C + \sqrt{C^2 + 4n_i^2}\right). \tag{2.22}
\]

In the case of phosphorus, it is usually assumed that the diffusing species may exist in three charge states, giving rise to a diffusivity characterized by three terms:

\[
D_p = D^0 + D^+\left(\frac{n}{n_i}\right) + D^-\left(\frac{n}{n_i}\right)^2. \tag{2.23}
\]

This relation is derived in [3] under the assumption that it is only the diffusing dopant which determines \( n \). Hence, the above \( n \)-dependence is an effect of both a Fermi-level dependent formation energy and an electric field, caused by the gradient in the dopant profile, which exerts an additional driving force on charged defects. Usually, experimental determination of the Fermi-level effect is done by measuring the diffusion of one dopant with a background of some other dopant which provides a uniform Fermi-level throughout the sample (see Chapter 4).

The absolute value and temperature dependence of the intrinsic carrier concentration \( n_i \) are obviously of importance when evaluating the Fermi-level dependence of the diffusion coefficient. \( n_i \) is given by [13]

\[
n_i = 2\left(\frac{2\pi k_B}{\hbar^2}\right)^{3/2}\langle m\rangle^{3/2}T^{3/2}\exp\left(-\frac{\Delta E_g}{2k_B T}\right), \tag{2.24}
\]
where \( \langle m \rangle \) is the geometric mean of the electron and hole effective masses and \( \Delta E_g \) is the bandgap, both are temperature dependent. Based on experimental data of \( \Delta E_g \), effective masses of the charge carriers and of the intrinsic charge carrier concentration, Thurmond [13] has calculated the intrinsic charge carrier concentration in Si from 150 K to the melting temperature, and found that the intrinsic carrier concentration can be described by the Arrhenius expression

\[
n_i = 1.8 \cdot 10^{21} \exp\left(-\frac{0.66 \text{ eV}}{kT}\right) \text{cm}^{-3}.
\]  

(2.25)

This expression includes temperature dependence of the band gap and the densities of states for holes and electrons, and it represents the value of \( n_i \) used in the present work. The expression is in accordance with values determined recently by Sproul and Green [14] and Green [15] in the 77 – 300 K and 300 – 500 K temperature range, respectively.

The Arrhenius type of temperature dependence in [2,19] implies that dopant diffusion is negligible at low temperatures. For example, with typical parameters for silicon a considerable diffusion, that would be destructive for devices, can be observed after an anneal for 10 hours at 800 °C. The same diffusion at room temperature would require more than \( 10^{31} \) years.

### 2.4 Point defects

Native point defects, the self-interstitial and the vacancy, play a crucial role in dopant diffusion through the reactions (2.9a)-(2.9c), and the diffusion coefficient directly depends on their formation and migration energies (equation (2.19)). Therefore, an introduction to point defects and their energetics will be given in the following section. The point defects discussed are those in silicon, unless otherwise stated.

The self-interstitial can exist in several configurations [8]. It can be situated as a tetrahedral (T) or hexagonal (H) interstitial, where it exerts a considerable strain on the surrounding substitutional Si atoms, or it can be in an interstitial configuration, also called the \( <110> \)-dumbbell configuration, with two Si atoms sharing one lattice site. In the case of a vacancy, an empty lattice site, the surrounding atoms relax inwards toward the empty site. Point defects can be generated by the Schottky or the Frenkel process [8]. In the Schottky process a self-interstitial is made by removing an atom from the surface and moving it into the crystal, and similarly with a vacancy. In the Frenkel process an atom moves out of its lattice site to an interstitial site, creating both a self-interstitial and a vacancy.
Theoretically calculated estimations of the properties of point defect have been based on molecular dynamics studies [16, 17] or density functional theory (DFT) [18, 19] (see Appendix B). DFT has proven itself to be a valuable tool for solving many-body problems, and it is currently the most commonly used method for calculating defect properties in semiconductors. Unfortunately specific approximations have to be introduced, often resulting in disagreements between the reported calculations. In Figure 2.2 calculated values of the formation energies for the Si self-interstitial are shown for different defect configurations. General trends can be identified from the figure: In $p$-type Si it is a doubly positive charged tetrahedral interstitial which has the lowest formation energy, while in $n$-type it is the negatively charged interstitialcy configuration. For the neutral interstitials the energy for the tetrahedral configuration is higher than for the hexagonal and interstitialcy configurations. On the other hand, there is a disagreement as to which configuration is the most stable in intrinsic Si, and there is a rather large difference in the absolute energy values. The calculations for the neutral interstitialcy differ by as much as 1 eV. This is a general tendency for DFT calculations that, while relative energies agree reasonably well, the absolute energies may differ substantially. This is reflected in Table 2.1, which lists ranges for the calculated values for formation and migration enthalpies of both Si self-interstitials and vacancies.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta H_{f}^{I}$</td>
<td>$3.2 - 4.2$ eV</td>
<td></td>
</tr>
<tr>
<td>$\Delta H_{m}^{I}$</td>
<td>$0.2 - 1.1$ eV</td>
<td></td>
</tr>
<tr>
<td>$\Delta H_{f}^{V}$</td>
<td>$2.4 - 4$ eV</td>
<td></td>
</tr>
<tr>
<td>$\Delta H_{m}^{V}$</td>
<td>$0.1 - 0.5$ eV</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1 Theoretical values for formation and migration enthalpies for the Si self-interstitial and vacancy (Refs. [16-19])
Total energy calculations on the self-interstitial in Ge showed that the interstitial configuration is the one with the lowest formation energy, 2.3 eV. In addition it was shown that this was much lower than other possible configurations for the Ge self-interstitial. This means that the migration energy for the Ge self-interstitial is large, since the self-interstitial diffuse by switching between the different configurations. Thus, the Ge self-interstitial is expected to have a small diffusion coefficient [21].

The values given in Ref. [16] can be used to get an idea of the magnitude of the concentrations of self-interstitials and vacancies in silicon. It is found that at 800 °C the concentrations are \(1 \times 10^{10} \text{ cm}^{-3}\) and \(1 \times 10^{8} \text{ cm}^{-3}\) for self-interstitials and vacancies, respectively. Since dopant concentrations are up to 10 orders of magnitude higher, the assumption, that the amount of dopants that enter into the fast diffusing complexes is negligible compared with the total dopant concentration, is reasonable.

The entropy of formation for a single defect can also be estimated theoretically. It is given by a sum of a configurational entropy, related to the number of ways in which the defect can be incorporated into a given lattice site, and a vibrational entropy, which can be viewed as a measure of the disorder introduced into the crystal by changing the vibrational properties of the neighboring atoms, or how tightly the defect is constrained in its equilibrium structure [22]. The configurational entropy of formation has been calculated in Ref. [22] and is \(-1k_B\) and \(-2k_B\) for vacancies and Si self-interstitials, respectively. For both types of point defects the vibrational entropy is around \(4k_B\). Additional entropy changes may arise when a point defect forms a complex with a substitutional dopant, as the vibrational entropy may decrease or increase depending on the spatial extension of the complex relative to the isolated point defect and dopant. Usually the entropy is regarded as temperature independent, although some calculations suggest that entropies of formation for point defects increase with increasing temperature [17]. Since the entropy is lumped together with other parameters in the pre-exponential factor, \(D_0\), the entropy of a diffusion process can only be estimated by making some assumptions about the constant \(K_1\) in equation [2.17].

A relatively low concentration of point defects in silicon makes a quantitative determination of their concentrations quite difficult. It is, however, possible to measure their diffusivity or the Si self-diffusivity which is a combination of interstitial and vacancy diffusion. Metals like Au, Pt, and Zn can be used as probes for measuring point defect properties, because the diffusion of these metals depends on the properties of the Si self-interstitial [23–25]. The self-interstitial diffusivity and concentration has for example been deduced from measuring zinc diffusion in silicon [24].
Direct measurements of the self-diffusion coefficient $D_{Si}$ in Si using isotopically enriched structures has also been reported [26, 27]. Typical structures consist of a layer where the concentration of the isotope used to monitor self diffusion, e. g. $^{30}$Si, was reduced from its natural abundance 3.10% to around 0.002%, so that the self-diffusion can be measured with conventional chemical profiling. Bracht et al. [26] compared their extracted value of $D_{Si}$ with results from earlier zinc diffusion experiments [24] to calculate the diffusion parameters for vacancy and interstitialcy mechanisms, respectively. Ural et al. [27], on the other hand, measured the self-diffusion during vacancy and interstitial injection and were able to directly split $D_{Si}$ into its vacancy and interstitial components. They found that the vacancy and the interstitialcy mechanism are equally important for Si self-diffusion. The activation energy values, i. e., the sum of the formation enthalpy and the migration enthalpy, for interstitial diffusion are 4.68 eV [27] and 4.95 eV [24]. So the experimental values lie within the range given in Table 2.1. For vacancy diffusion activation energies of 4.86 eV [27] and 4.14 eV [26] are found. According to the theory this value should be between 2.5 – 4.5 eV. The value for vacancy migration in Table 2.1 may, however, be too low, according to a recent publication by Bracht et al. [28]. They have reported a migration enthalpy of $1.8\pm0.5$ eV.

The relative importance of the bulk or the silicon surface as sources for point defects was measured by Fang et al. [29], who used diffusion of the interstitialcy mediated boron as a marker for self-interstitials. B diffusion in two boron spikes, which were separated by a relaxed SiGe layer, that Si self interstitials could not penetrate, showed that the surface was the principal source of I’s. Moreover, they discovered that the influence of the bulk was dependent on the type of Si substrate that was used. With Czochralski (Cz) silicon as substrate the influence of the bulk was more pronounced than if Float zone (FZ) silicon was used as a substrate. This is explained as an effect of impurities in the substrate, in particular oxygen. This may be related to the discovery by Hu [30], who found that over a time scale of hours, bulk oxygen, in an interstitial position, causes a temporary supersaturation of I due to formation of SiO$_2$, which has a larger atomic volume than substitutional silicon atoms. Si self-interstitials are emitted to relax the strain on the surrounding lattice.

### 2.5 Perturbations in point defect concentrations

In general, a change in point defect concentrations will cause non-equilibrium diffusion. For example, annealing in some reactive gases, such as O$_2$ or NH$_3$, is known to inject interstitials and vacancies, respectively, and thereby enhance or retard diffusion. The diffusivity under non-equilibrium conditions relative to the equilibrium value can be given in terms of the super/under-saturation of I and V respectively:
\[
\frac{\langle D_i \rangle}{D_i^0} = f_i \frac{\langle C_i \rangle}{C_i^*} + (1 - f_i) \frac{\langle C_V \rangle}{C_V^*}.
\] (2.26)

Asterisks indicate equilibrium values, and the brackets means values averaged over a time \(T\), i.e.

\[
\langle D_i \rangle = \frac{1}{T} \int_0^T D_i(t) dt.
\] (2.27)

Equation (2.26) can be used to determine the interstitialcy fraction if the non-equilibrium point defect concentrations are known, or it can be used to determine point defect concentrations by measuring diffusion of dopants with known interstitialcy fraction.

The vacancy concentration is also affected during injection of self-interstitials. Since the self-interstitial can recombine with a vacancy, there is generally an undersaturation of \(V\)’s when there is a supersaturation of \(I\)’s and vice versa. This may be expressed in terms of a mass-action law, which relates the equilibrium concentrations to the non-equilibrium concentrations:

\[
C_i C_V = C_i^* C_V^*.
\] (2.28)

This equation is only valid when thermal excitation is responsible for excess defects. In general \(C_i C_V\) may be larger under non-equilibrium conditions [31].

In practice, the supersaturation of Si self-interstitials or vacancies can be caused by surface reactions (oxidation or nitridation) during processing of a Si wafer. Other process steps may also influence the point defect concentrations. Ion implantation, where energetic ions kick out host atoms of their lattice sites, thus generating both interstitials and vacancies in equal amounts. Also, the presence of other elements may affect the defect concentrations. It was mentioned earlier that oxygen in silicon might enhance the interstitial concentration. Furthermore, it has been observed that diffusion of interstitialcy-mediated diffusers, such as B and P, is retarded in carbon rich samples \((C > 10^{19} \text{ cm}^{-3})\) [32]. This was explained by a trapping of the self-interstitials by carbon, leading to an undersaturation of \(I\)’s and a supersaturation of \(V\)’s which was supported by the fact that antimony diffusion, which is mediated by a vacancy mechanism, is enhanced in C rich samples.

### 2.6 Effect of hydrostatic pressure

A change in enthalpy caused by some process in the presence of an applied pressure is described by

\[
\Delta H = \Delta U + p \Delta V,
\] (2.29)
where $\Delta U$ is a change in internal energy, and $p\Delta V$ is an energy associated with the change in volume $\Delta V$ caused by the process. At a constant pressure, the latter term can be lumped together with the internal energy. For a diffusion process the change in volume is called the activation volume $\Delta V_a$, and the specific pressure dependence of the diffusivity is

$$D_A(p) = D_A^0 \exp \left( - \frac{E_a + p\Delta V_a}{k_B T} \right).$$

(2.30)

Thus, a negative $\Delta V_a$ increases the diffusivity, while a positive $\Delta V_a$ decreases the diffusivity. $\Delta V_a$ consists of several terms related to the processes of defect formation, complex formation and migration. For diffusion via the interstitial mechanism it becomes

$$\Delta V_a = \Delta V_{I}^f + \Delta V_{I}^m + \Delta V_{I}^m .$$

(2.31)

For an interstitial generated through the Schottky process, the formation volume consists of two parts; a decrease in the crystal volume by one atomic volume unit $\Omega$ ($\Omega$ is the volume occupied by one atom, in Si $\Omega \sim 2 \times 10^{-23}$ cm$^3$), because a Si atom is removed from the surface, and a change in volume caused by relaxation of the lattice surrounding the interstitial. The magnitude of this so-called relaxation volume is found, by calculation [16], to be of the same order of magnitude as the atomic volume unit, giving a formation volume for interstitials, $\Delta V_{I}^f = -0.10 \Omega$ [16]. Generating a vacancy by the Schottky process involves moving an atom from the bulk to the surface, which increases the lattice volume by $\Omega$. However, the lattice relaxes inward around the vacancy resulting in a total formation volume for vacancies smaller than $\Omega$. $\Delta V_{V}^f$ has been calculated to be between $-0.08 \Omega$ and $0.2 \Omega$ [33]. Antonelli and Bernholc [34] calculated the formation enthalpy, equation (2.29), as a function of pressure. Their results are consistent with formation volumes for self-interstitials of $-0.3 \Omega$ and for vacancies $0.3 \Omega$.

Diffusion of point defects involves both the formation and the migration of the defects. Sugino et al. [35] calculated the pressure effect of both the processes for vacancy diffusion. They found that the total activation volume for the vacancy mechanism of Si self-diffusion is $-0.46 \Omega$, while for the dopants Sb and As, they find activation volumes of $0.06 \Omega$ and $0.2 \Omega$, respectively, assuming only the vacancy mechanism to be active.

The activation volume term is usually omitted for diffusion at atmospheric pressure as $p\Delta V_a$ is approximately $1 \times 10^{-5}$ eV, using the atomic volume unit for $\Delta V_a$, so even with uncertainties of the order of 0.05 eV in the activation energy for diffusion, this term is indeed negligible. In order to be able to measure the pressure dependence on diffusion,
pressures in the GPa range should be applied. Such experiments have been performed, and data exist for the activation volume of boron diffusion in Si, which is \(-0.17\Omega\) \(\text{[36]}\), arsenic in Si \((\Delta V_a = -0.47\Omega \text{ [37]\])\) and in Ge \((\Delta V_a = -0.12\Omega \text{ [38]\})\). Arsenic is known to diffuse via a combination of both interstitial and vacancy mechanism in Si, and this may explain the disagreement between the measured value in Si \((-0.47\Omega\)\) and the theoretical value \((0.2\Omega\)\), which was calculated under the assumption of a vacancy mechanism. The activation volume for Sb is measured to be \(0.07\Omega \text{ [39]\) in agreement with the calculated value of \(0.06\Omega \text{ [35]\).}

2.7 Diffusion in heterostructures

So far, only diffusion in a homogeneous material has been considered. However, diffusion in multi-component structures is of great interest, especially since epitaxial growth techniques like chemical vapor deposition (CVD) and molecular beam epitaxy (MBE) have made high quality heterostructures accessible. This thesis deals with Si/Si$_{1-x}$Ge$_x$/Si heterostructures, where a layer of SiGe, with an atomic percentage of Ge given by \(x\), is grown pseudomorphically on a Si-substrate followed by another Si-layer. Diffusion in such a structure is affected by several effects: the diffusion coefficient may vary as the Ge content in the layer is changed, differences in the solid solubility of the dopants will cause segregation of the dopants near the interfaces in the heterostructures, the strain in the pseudomorphically grown layers may affect the diffusivity in a similar manner as an externally applied pressure does, and structural defects affect the diffusion by perturbing the point defect concentrations. These effects will be discussed in the following sections.

2.7.1 Composition effect

In a first approximation, since SiGe is a nearly ideal solution, the average atomic interactions might be almost the same regardless of the atom considered to be a Si atom or a Ge atom. This means that the formation enthalpies of point defects should be the same irrespective of the surroundings. There may, however, be some difference in the case of self-interstitials, which can be either a Si atom or a Ge atom. The microscopic strain, which the bigger Ge self-interstitial imposes on the surrounding lattice, may be higher than that imposed by the Si self-interstitial, thus leading to different formation enthalpies.

In order to form a self-interstitial it is necessary to break a number of chemical bonds. The bonding energy is related to the melting temperature \(T_m\) of the material; the higher the bonding energy the higher the melting temperature. So the formation enthalpy of point defects in materials of the same lattice structure can, in a first approximation, be assumed to be proportional to the melting temperature. The phase diagram of a SiGe
alloy (Figure 2.3) show that $T_m$ decreases with increasing Ge content, which means that the formation enthalpy of point defects is likely to decrease with increasing $x$.

The band gap also changes with Ge content (Figure 2.4), and this affects the concentration of intrinsic charge carriers, which we approximate by

$$n_i(x) = n_i(0) \exp \left( \frac{\Delta E_g(x)}{2k_B T} \right), \quad (2.32)$$

where $n_i(0)$ is the intrinsic carrier concentration in Si given by equation (2.25), and $\Delta E_g(x)$ is the change in band gap compared to Si. According to Figure 2.4, the band gap decreases with increasing Ge content, so the intrinsic carrier concentration in SiGe increases with Ge content.

The entropy change associated with the diffusion process may as well be composition dependent. For an ideal alloy where the individual atoms are randomly distributed, there will be an entropy related to the disorder of the mixing, called the entropy of mixing, which is given by \[ H_{mix} \]

$$\Delta S_{mix} = k_B \ln \left( \frac{N!}{N_{Si}!N_{Ge}!} \right) = -Nk_B \left[ x \ln x + (1-x) \ln (1-x) \right]. \quad (2.33)$$

For $x = 0.5$, the entropy of mixing is $0.7k_B$ per atom. It is hypothesized in Chapter S4 that the entropy change associated with the diffusion process in SiGe has a similar $x$-dependence.

Finally, a change in the diffusion mechanism should be considered. Dopant diffusion in Ge is believed to be via a vacancy mechanism \[ H_{43} \], which implies that the interstitialcy fraction for P diffusion decreases from 1 in pure Si to 0 in pure Ge.
Especially, in the case of P diffusion the composition effect is large, since the diffusivity in Ge is approximately 5 orders of magnitude larger than that in Si, for the temperature range of 700°C – 900°C [43].

2.7.2 Segregation

Returning to equation \(\{2.5\}\), it is seen that the solubility, \(C_{eq}^A\), is assumed to be constant in the derivation of the force exerted on the dopant. This is not necessarily true in a heterostructure, where \(C_{eq}^A\) may vary from layer to layer. Considering a spatial variation of the solubility at the interfaces in the structure, the force exerted on the dopants will have an additional term:

\[
f = -k_B T \left( \frac{1}{C_A} \frac{\partial C_A}{\partial x} - \frac{1}{C_{eq}^A} \frac{\partial C_{eq}^A}{\partial x} \right). \tag{2.34}
\]

The flux of dopants then becomes

\[
J = -D_A \left( \frac{\partial C_A}{\partial x} - \frac{C_A}{C_{eq}^A} \frac{\partial C_{eq}^A}{\partial x} \right), \tag{2.35}
\]

where the first term accounts for the ordinary Fick diffusion due to concentration gradients, while the second term is non-zero only in the vicinity of the interface, where it causes an extra flux of dopants into the material with the highest solubility. When the ratio of concentrations at either side of the interface equals the ratio of the corresponding dopant solubilities, the segregation flux will be of the same magnitude, but opposite to the flux due to the concentration gradient, and an equilibrium situation will occur. The ratio of the dopant solubilities is called the segregation coefficient \(k\).

The model of segregation given above is rather simple, but turns out to be applicable to the experimentally observed segregation of P in Si/SiGe/Si heterostructures, as it will be shown in Chapter 5. A more refined treatment of the segregation should include the influence of the electric field at the interfaces, caused by the difference in band gap between Si and SiGe, and on charged dopant-point defects \(\{44\}, \{45\}\). Also, both the macroscopic and microscopic strain dependence should be included \(\{42\}, \{47\}\).

Boron tends to segregate into the SiGe part of the heterostructure (\(k=0.74\) at 950°C, 8% Ge \(\{46\}\) (\(k=0.33\) at 850°C, 20% Ge \(\{47\}\)). Phosphorus segregates into the Si part of the heterostructure (\(k=1.35\) at 950°C, 8% Ge \(\{48\}\) and \(k\) increases with increasing Ge content \(\{45\}\).
2.7.3 Strain

The silicon lattice has a diamond structure, which is a face-centered cubic lattice with a two point basis (see e.g. Ref. [48]). The lattice parameter of Si is $a_{\text{Si}} = 0.5431$ nm. Germanium has the same lattice structure but a different lattice parameter; $a_{\text{Ge}} = 0.5657$ nm. The lattice parameter for Si$_{1-x}$Ge$_x$ can be approximated by Vegard’s rule [49]:

$$a_{\text{SiGe}}(x) = a_{\text{Si}}(1-x) + a_{\text{Ge}}x.$$ (2.36)

When a SiGe layer is grown on a Si-substrate the SiGe layer will be tetragonally distorted to fit the substrate lattice. The lattice parameter in the direction perpendicular to the substrate surface will be elongated, this is known as the Poisson effect. The amount of distortion of the lattice is characterized by the strain parameter $s$, defined as the difference between the lattice constant of the strained crystal and that of the relaxed crystal normalized to the relaxed lattice constant:

$$s = \frac{\Delta a}{a}.$$ (2.37)

A SiGe layer on a silicon substrate will have a negative in-plane strain $s_{\parallel}$, and this parameter is usually referred to as the strain $s$ of the layer, but there will also be a strain $s_{\perp}$ of the perpendicular lattice parameter due to the Poisson effect. The ratio between the in-plane strain and the perpendicular strain can be determined by simple stress considerations [50], and it is given by

$$\frac{s_{\perp}}{s_{\parallel}} = \frac{2c_{12}}{c_{11}},$$ (2.38)

where $c_{12}$ and $c_{11}$ are the stiffness constants of the material. They are known for Si and Ge [49], but for the Si$_{1-x}$Ge$_x$ alloy they are usually calculated by using a Vegard law like equation (2.36). When the Si$_{1-x}$Ge$_x$ is grown on a Si substrate, $a_{\text{SiGe}} = a_{\text{Si}}$ so the strain becomes

$$s_{\parallel} = \frac{a_{\text{SiGe}} - a_{\text{Si}}}{a_{\text{Si}}}.$$ (2.39)

From the two equations above, the lattice parameter perpendicular to the surface can be calculated as

$$a_{\text{SiGe}\perp} = a_{\text{SiGe}} - \frac{2c_{12}}{c_{11}} (a_{\text{SiGe}} - a_{\text{Si}}).$$ (2.40)

This relation between the parallel and perpendicular lattice parameter is important for X-ray measurements, as it will be explained in Section 3.2. For a positive $s_\parallel$, the lattice parameter of the substrate is larger than that of the layer (e.g. Si on relaxed SiGe) so in this case the strain is tensile. If SiGe is grown on Si the strain is compressive ($s_\parallel < 0$).
Cowern et al. [51] found that the diffusivity of boron at 975 °C in strained SiGe samples with five different Ge concentrations decreased exponentially with increasing Ge content. This lead them to define the apparent activation energy $Q'$ with units of energy per unit strain, which is a quantitative measure of how the diffusion coefficient responds to strain, in the same way that the activation volume is a measure of the response to pressure (equation (2.30)):

$$D_A(s) = D_A^0 \exp \left( -\frac{E_a + Q's}{k_B T} \right).$$

(2.41)

The connection between pressure, activation volume, strain and apparent activation energy was used by Aziz [52] to find a relation between $\Delta V_a$ and $Q'$:

$$\frac{\Delta V_a}{\Omega} + \frac{3}{2} \frac{Q'}{Y\Omega} = \pm 1 + \left( V_{m\parallel}^\infty - V_{m\perp}^\infty \right)/\Omega,$$

(2.42)

where $Y$ is a material constant containing Young’s modulus and Poisson’s ratio, and $Y = 180.5$ Gpa for Si. $V_{m\parallel}^\infty$ and $V_{m\perp}^\infty$ are migration volumes in directions parallel with and perpendicular to the surface, respectively. The “+” in the “±1” term in equation (2.42) is for a vacancy mediated diffusion and the “−” is for an interstitial mechanism. The term is related to the point defect formation at the surface, i.e., the Schottky mechanism as discussed previously.

There are only a few calculations on the effect of strain on diffusion. Antonelli and Bernholc [53] calculated the strain dependence of the formation energy for a neutral tetrahedral interstitial, the energy is found to decrease with increasing tensile strain, and is suggested to increase with compressive strain. The formation energy of the vacancy did not show any strain dependence. The data are, however, too limited to draw any definite conclusions about the actual diffusion modes, but one may anticipate a retardation of interstitial mediated diffusion in compressive strained material, which has actually been observed experimentally (see Chapter 5).

### 2.7.4 Relaxation

There is a limit for the amount of elastic energy that can be stored in a strained layer. Above a certain thickness it will be energetically favorable for the layer to relax by formation of misfit dislocations. Matthews and Blakeslee (MB) [54] found an expression for the theoretical critical thickness $h_c$ for a strained SiGe layer on a Si substrate based on equilibrium considerations; the MB result is illustrated in Figure 2.5. In reality, metastable strained layers with thickness above $h_c$ can be accomplished since the relaxation process is controlled by nucleation of the misfit dislocations, which relax the strain [55, 56]. The nucleation is thermally activated, which means that...
the onset of the relaxation will take place during an annealing of the strained structure. The extent of the metastable region in Figure 2.5 is not well defined, because the formation of misfit dislocations depends on the amount of nucleation sites. The nucleation is a complex process, which hinges on several parameters, where different materials exhibit different stability of strained layers above $h_c$.

The most common misfit dislocation is the 60° dislocation, which is a dislocation lying in the heterostructure interface with a Burger’s vector in the [011] direction $[\overline{55}2 \overline{5}]$. The accommodation of compressive strain by the 60°-type misfit dislocations can be estimated by

$$\Delta s = \frac{d a_{Si}}{\sqrt{2}},$$  \hspace{1cm} (2.43)

where $d$ is the number of dislocations per unit length, and $\Delta s$ is the change in compressive strain. Equation (2.43) gives that ~1 dislocation per $\mu$m causes a 1% strain relaxation of a strained $Si_{0.72}Ge_{0.18}$ layer grown on Si.

The formation and presence of misfit dislocations may have an effect on dopant diffusion due to an interaction between the dislocations and the point defects. For low temperature and short time annealing, the strained SiGe layer in a heterostructure has been shown to be transparent to Si self-interstitials $[47, 58]$. On the other hand, an unstable strained SiGe layer act as a sink for Si self-interstitials during I-injection by oxidation $[29]$. Kuznetsov et al. $[59]$ observed that oxidation anomaly degraded a strained SiGe layer, and this was interpreted as interstitials facilitating the formation of misfit dislocations, which relax the strain. Because of this interaction between misfit dislocations and Si self-interstitials, it is essential to avoid relaxation when diffusion of interstitial mediated diffusers is studied in strained heterostructures.

![Figure 2.5 Critical thickness MB-model. Above this thickness metastable structures can exist.](image-url)
Chapter 3
Sample preparation and characterization

3.1 Sample preparation

Samples with simple dopant profiles, like spikes or boxes, are advantageous for diffuson studies, and a conventional method for producing such samples is epitaxy. Specifically, the samples used in the present work have been made by reduced pressure chemical vapor deposition (RPCVD) or molecular beam epitaxy (MBE). Epitaxy denotes the process of growth, on a crystalline substrate, of crystalline layers that have the same crystal structure as the substrate.

In CVD, layers are grown using gases like silane (SiH₄) or dichlorosilane (SiH₂Cl₂) for silicon layers and germane (GeH₄) for growing layers containing Ge. Doping is done by adding diluted B₂H₆ or PH₃ in H₂ for B or P doping, respectively. Relative concentrations of Ge and dopants are controlled by the partial pressure of the precursor gases. Deposition of atoms on the substrate occurs through the dissociation of the reactant gases, for these reactions to proceed the sample is heated to 600 – 700 °C.

MBE uses beams of atoms or molecules of the relevant elements evaporated from a source, the beam hit the substrate, which is situated in an ultra high vacuum chamber. As the beam hits the sample some of the atoms will be adsorbed on the surface and will eventually bind to the surface. MBE grown material contains carbon and oxygen $C_C \approx 10^{18}$ cm$^{-3}$, $C_O \approx 10^{16}$ cm$^{-3}$ as the main residual impurities, whereas the corresponding values for layers grown by CVD are of an order of magnitude lower.

The epitaxy process ensures that Si and SiGe layers will have the same crystallinity as the underlying Si substrate, which for the SiGe layers means that they will be strained. It is another matter to grow relaxed SiGe layers with good crystallinity on a Si substrate. In practice they are made by first growing a linearly graded layer with a grading of approximately 10% Ge per µm. The graded layer will have a high density of misfit dislocations [57]. These do not cause problems, if they are confined in the graded layer, and if they are thermally stable during subsequent annealing, so they do
not affect the point defect concentration in the upper layer. The surface of the graded layer will have a crystallinity with lattice parameters reflecting the Ge concentration, so it can be used as a substrate for a relaxed SiGe layer.

Sample profiles of the structures with relaxed SiGe layers are shown in Figure 3.1 and Figure 3.2. The as-deposited P profile has a spike with a concentration of $3 \times 10^{19}$ cm$^{-3}$, which is large enough to cause a concentration dependent diffusion. This is undesirable, and in order to lower the maximum concentration of the P profiles, the samples were pre-annealed in vacuum at 950°C for 1 hour prior to the diffusion treatment. Any transient effects, which may be caused by structural imperfections in the graded SiGe layer, are expected to be reduced by this pre-anneal.

Annealing was done in a temperature range of 700 – 1100 °C using either a tube furnace or a vacuum furnace. The duration of the annealing was determined for every temperature to give a significant diffusion and, specifically, annealing durations were in the interval of 30 minutes to 60 hours.

The tube furnace provides the possibility of annealing in a non-reactive N$_2$ ambient or a reactive O$_2$ ambient. Temperature and ambient flow are recorded and stored as a function of time, and the temperature is given with an accuracy of ±2 °C. The tube furnace has the disadvantage that the samples are introduced into the furnace at ~620°C and then heated to the desired anneal temperature. An example of the temperature evolution during the annealing is given in Figure 3.3. The heating rate is 5 - 6 °C/min, while the rate of cooling down is 1.5 °C/min irrespective of the chosen annealing time. This means that for high temperatures, where a short annealing duration is required, the heat up and cool down ramps are important, since a
considerable diffusion occurs during ramping, which has to be accounted for when extracting the diffusivities (see Appendix A).

The vacuum furnace operates at a pressure of \( \sim 10^{-6} \) Torr. The vacuum annealing has been found to cause a non-equilibrium diffusion for dopants, which diffuse through an interstitial mechanism (Section 4.2.1). The process is time dependent and may as well be depth dependent. For the furnace, which has been used in the present work, it was found that some chemical reaction took place on the sample surface during the vacuum anneal. According to optical ellipsometry a surface layer of unknown identity is formed. Rutherford backscattering spectrometry (RBS) analysis showed no oxygen, but revealed the presence of carbon and a small content of some transition metal. Indeed, mass spectroscopy revealed the presence of both chromium and zinc, metals that probably originate from a contamination of the furnace by previously annealed samples.

3.2 High resolution X-ray diffraction

The strained heterostructures used in the present work are metastable, so some relaxation may occur during annealing. Relaxation takes place by the formation of a dislocation network at the interfaces, and hence point defect concentrations and diffusion will be affected. Furthermore, diffusion may be strain dependent, so it is crucial that the strain status is preserved during annealing, and therefore X-ray diffraction measurements have been employed. Two modes of X-ray diffraction will be discussed in the following: (i) \( \omega 2\theta \) scan, which is a fast method from which strain, Ge content and layer thickness of the heterostructure can be determined and (ii) reciprocal lattice mapping, a method which is more time consuming than the \( \omega 2\theta \) scan, but it detects a relaxation of a few percent in the structure, also it can reveal the presence of structural defects.

Figure 3.3 Temperature evolution during a tube furnace anneal at 950°C for \( \sim 2 \) hours.
A schematic of the high-resolution X-ray diffraction setup, used in the present work, is shown in Figure 3.4 [60]. The X rays, with a wavelength of $\lambda = 1.540562$ Å (Cu Kα transition) pass through a monochromator, consisting of 4 Ge (220) crystals, which assure a monochrome beam with a divergence below 12 arcsec for the incident X-ray beam. The incoming X-ray beam angle is $\omega$, and the diffracted X rays are counted by a detector which is at an angle of $2\theta$ relative to the incoming beam line. A two-crystal collimator in front of the detector decreases the diffracted beam divergence to 12 arcsec, which is especially desirable when measuring reciprocal lattice maps [60].

According to Bragg’s law, X rays scattered from parallel planes, separated by a distance $d$, will interfere constructively when the difference in distance traveled is equal to an integer $n$ times the wave length of the X rays, which is the same as satisfying the equation

$$n\lambda = 2d \sin(\omega'),$$

(3.1)

where $\omega'$ is the angle of incidence relative to the diffraction plane (see Figure 3.5). Each family of planes defines a vector $\mathbf{q}$, which is perpendicular to the planes, and has a length, which is the reciprocal of the distance between the planes. If the vectors $\mathbf{k}$ and $\mathbf{k}'$ are defined in the direction of the incoming and diffracted beam, respectively, and with a length of $1/\lambda$, it can be shown that constructive interference is obtained when

$$\mathbf{k}' - \mathbf{k} = \mathbf{q}.$$

(3.2)

In fact, this equation is equivalent to the Bragg condition. Symmetric diffraction, i.e. when $\mathbf{q}$ is perpendicular to the surface, only gives information on the lattice parameter perpendicular to the surface. However, the planes do not necessarily have to be parallel to the surface. For example, in Figure 3.5, diffracting planes with a $\mathbf{q}$ vector in the
direction are shown. Diffraction on those planes will also contain information on the lattice parameters in the $x$ and $y$-directions.

For any family of lattice planes a vector $\mathbf{q}$ can be defined. The space spanned by the $\mathbf{q}$’s from a given lattice is called the reciprocal lattice or reciprocal space. Since X rays can be diffracted on any family of lattice planes, the points in the reciprocal lattice can be understood as representing reflections.

In Figure 3.6, points in a plane in the reciprocal lattice of a Si/SiGe heterostructure are shown. The plane is spanned by the [001] and [110] directions, the directions which in real space are the $z$-direction (perpendicular to the sample surface) and a diagonal in the $x$-$y$ plane (parallel to the sample surface). The black points represent the reciprocal lattice of the Si substrate, the gray points represent reflections from the strained SiGe layer, which has a lattice constant in the direction perpendicular to the sample surface $a_{\text{SiGe}}$ which is larger than $a_{\text{Si}}$. In the $x$- and $y$-directions the strained layer has the same lattice parameter as the substrate, so a reciprocal lattice point of a strained layer will appear directly below the lattice point of the substrate. If the SiGe layer is relaxed its lattice parameter in the $z$-direction will be smaller compared to the strained layer, but larger than the one for the substrate. In the $x$-, $y$-directions it will also be larger than the parameter of the substrate. So reciprocal lattice points of a relaxed layer will be located as the white points in Figure 3.6.

Using equation (3.2), the conditions, in which constructive interference will occur, have been visualized geometrically with the so-called Ewald sphere in Figure 3.6; The sample is placed in the center of a sphere with radius $1/\lambda$, the incoming wave vector, $\mathbf{k}$
points from the sample to the origin of the reciprocal lattice and a vector representing the diffracted wave, \( \mathbf{k}' \) points from the sample to the circumference of the sphere. If the difference between the two vectors (represented by the dotted line in the figure) hits a point in the reciprocal lattice, constructive interference will occur. Thus, by varying \( \omega \) and \( 2\theta \) a line (\( \omega/2\theta \) scan) or an area (reciprocal lattice mapping) can be probed. In Figure 3.6 part of the Ewald sphere is shown for 3 different values of \( (\omega, 2\theta) \), and a reflection will only be observed for the sample represented by black in Figure 3.6.

The silicon lattice is not a simple cubic lattice as the one shown in Figure 3.5 but a face-centered cubic (fcc) lattice with a two atom basis (the so-called diamond structure). The reciprocal lattice of an fcc lattice is a body-centered cubic (bcc) lattice, of which the cubic lattice is a sub-set, and due to the two-point basis a geometrical structure factor will cause the intensity of the reflections to vary (for example the point (002) will have no reflection). For a more rigorous treatment of reciprocal lattices see e.g. Ref. [48]. The purpose of the X-ray diffraction measurements performed in this work is not to determine the lattice structure, but rather changes in lattice constants, and for this specific reflections are selected.
3.2.1 $\omega/2\theta$-scan

The $\omega/2\theta$ scan is a one-dimensional measurement which scans $\omega$ and $2\theta$ angles with a ratio 1:2. In this case different planar spacings in a certain crystal orientation are probed. Usually, it is the (004) reflection that is measured, because it provides a high intensity. For a heterostructure, the reciprocal lattice points, corresponding to the silicon substrate and the SiGe-layer, are observed as two peaks, as exemplified in Figure 3.7. The distance between the two peaks is given by $d_{\text{SiGe}}$ and $d_{\text{Si}}$, which means that the Ge content in the SiGe layer can be determined by applying equation (2.40), together with Vegard’s law. This method is based on some assumptions that must be fulfilled; there must be a perfect fit between substrate and layer (i. e. no mismatch or partly relaxation), and that the layer is elastically distorted.

Due to the finite size of the sample, the peaks are not perfect $\delta$-functions. In fact, the width of the SiGe peak is qualitatively related to the thickness of the SiGe layer; the broader the peak, the thinner the layer. The oscillations around the peaks in Figure 3.7 are also related to the thickness of the layer. The oscillations are known as finite thickness fringes, and they are caused by X ray reflections at the layer interfaces. They only occur in a perfect structure where no interfacial defects disrupt the X-ray wavefield, so the diffraction is coherent across the interface [61]. In other words, the fringes are only observed in samples with no interfacial defects.

An example of an $\omega/2\theta$ scan of a partly relaxed sample is given in Figure 3.8, where both the scan of the as-grown sample and the annealed sample are shown. In the scan of the as-grown sample the finite thickness fringes are not very clear, which indicates an onset relaxation. This becomes more pronounced after annealing, where the maximum of the layer peak has moved toward the substrate peak, showing a decrease in the average value of $d_{\text{SiGe}}$ due to the annealing.

![Figure 3.7 Scan of a strained Si/Si$_{0.82}$Ge$_{0.18}$/Si structure.](image)

![Figure 3.8 Scan of a partly relaxed Si/Si$_{0.79}$Ge$_{0.21}$/Si structure before (solid line) and after anneal (dashed line).](image)
3.2.2 Reciprocal lattice map

Though \(\omega/2\theta\) scans give valuable results on thickness and composition of the layers in a Si/SiGe heterostructure, they are generally not sensitive enough for detailed studies of relaxation in the heterostructures [62]. Instead, a more precise method for determining strain, the reciprocal lattice map (RLM), is usually preferred. The RLM is basically a series of \(\omega/2\theta\) scans, with an offset of \(\Delta\omega\) between individual scans. Typically the area around the [113] reflection is measured, and this area is indicated in Figure 3.6 by the dotted rectangle. Two examples of reciprocal lattice maps recorded on two types of SiGe samples are shown in Figure 3.9.

Figure 3.9(a) is the RLM of a relaxed SiGe sample grown on a Si substrate. The reflection from this substrate appears as a well resolved point above the reflection from the SiGe layer. The layer peak is the collective reflection from a graded SiGe layer and a relaxed SiGe layer. Since the samples with a relaxed SiGe layer also contain a graded layer with a high density of dislocations, the RLM of such samples will be rather complex and not suitable for any exact interpretations. It is a different matter with the RLM of a strained heterostructure, of which an example is shown in Figure 3.9(b). This RLM only contains reflections from the substrate and the SiGe layer and has the same features as the \(\omega/2\theta\) scan in Figure 3.7. The substrate and the layer peaks appear as two islands, and the thickness fringes are clearly resolved between the peaks. The broadening of the peaks and the relaxation due to different annealing conditions are illustrated in Fig. 5, Paper III.

![Figure 3.9](image)

Figure 3.9 Reciprocal lattice maps of (113)-reflection. (a) relaxed SiGe on graded layer. (b) strained Si/Si\(_{0.82}\)Ge\(_{0.18}\)/Ge heterostructure.
The broadening around the peaks in the RLM is a consequence of the finite sample size (similar to the width of the peak in a \( \omega/2 \theta \) scan), but also of diffuse scattering of the X rays, caused by long-range strain fields associated with defects and structural imperfections [61]. The structural imperfections can be locally relaxed regions close to dislocations [62].

Strain is relieved through generation of misfit dislocations as described in Section 2.7.4. Dislocations cause an increase in \( a_{\text{SiGe}} \) and a decrease in \( a_{\text{SiGeL}} \). This will manifest itself as an asymmetric broadening in the \( \omega \)-direction on the left-hand side of the layer peak. An example of this kind of broadening is clearly seen in Fig. 5(c) in Paper III, where the peak from the strained layer has moved toward the substrate peak, corresponding to a strain relaxation of approximately 4%.

### 3.3 Secondary ion mass spectrometry

Dopant profiles, like the ones in Figure 3.1 and Figure 3.2, are measured with secondary ion mass spectrometry (SIMS). SIMS is based on a physical phenomenon known as sputtering, which is the ejection of particles from a sample bombarded by energetic ions. To obtain depth profiles, part of the sputtered atoms is analysed by mass spectrometry as described below [63, 64].

A schematic overview of a magnetic sector SIMS instrument, like the Cameca ims 4f used in the present work, is shown in Figure 3.10. The beam of primary ions is extracted from an ion source and accelerated towards the sample, where they are rastered over an area of, typically, \( 200 \times 200 \mu \text{m}^2 \). The ions impinge on the sample with an energy, which can be varied between 3-15 keV depending on the measurement conditions. The primary ions undergo a series of collisions with the atoms in the target,

![Figure 3.10 Schematics of the secondary ion mass spectrometer.](image-url)
and thereby transfer energy to the target atoms, some of which will receive enough energy to escape the target. The volume, in which the primary ions are stopped, is called the cascade collision volume, and the ejected atoms originate from this volume. The extent of the collision volume is related to the projected range for the primary ions. For a 13.5 keV Cs\(^+\) ion at 45° incidence relative to the surface normal, the projected range in silicon is \(~100\ \text{Å}\) according to simulations.

The sputtered particles are in the form of charged or neutral atoms or molecules, but only the secondary ions are used for further analysis, as they can be manipulated by electric and magnetic fields. The secondary ions are accelerated towards an electrostatic sector analyser (ESA) by a target potential of 4.5 kV. The ESA is an energy filter, which consists of two curved plates, between which there is an electric field \(E_0\) that deflects the ions. Only those ions that follow the path with a radius of curvature \(r_e\) will pass the analyser. After the ESA the beam passes through a magnetic sector analyser (MSA), where a magnetic field \(B\) perpendicular to the beam deflects the beam according to the ion momentum. The effect of the two filters is that only ions with a mass to charge ratio, given by the parameters set by the instrument, will pass through to the detector. The relation between the mass to charge ratio and the parameters are

\[
\frac{M}{q} = \frac{\left(\frac{r_e B^2}{q E_0}\right)}{r_e E_0}.
\] (3.3)

By varying \(B\) during the measurement, several elements can be analysed simultaneously. For example, a matrix element is usual included, in addition to the elements of interest, in the measurement. The matrix element should give a constant signal, unless irregularities, such as variation in the primary beam current, are encountered during the measurement. Also, in the case of the Si/SiGe heterostructures the Ge profile can be studied as well.

Sputtering leads to a successive erosion of the target material and a crater is formed. The crater depth is typically measured using a surface stylus profilometer. The secondary ion signal is recorded as counts per second versus sputtering time. This is then converted to concentration versus depth profile using reference samples, as discussed below. When measuring a depth profile, the secondary signal should come from the bottom of the crater and not the crater walls. This is ensured by an electronic gating, which only accept the signal from the center part of the crater bottom.

The time-to-depth conversion is usually done by assuming a constant erosion rate throughout the measurement. This approximation is valid for homogenous samples, but may fail in the case of heterostructures. Specifically, the erosion rate in SiGe as a function of Ge has been studied in the present work. The results are shown in Figure 3.11, where it is evident that the erosion rate indeed depends on the Ge content.
The conversion of counts/sec to concentration depends on the total sputtering yield and the secondary ion yield of the elements of interest. The total yield is the number of sputtered target atoms per primary ion, and the ionization yield is the number of ionized atoms relative to the total amount. The ionization yield is strongly enhanced by using reactive primary ions, e.g. using a primary beam of $\text{O}_2^+$ enhances the yield of positive ions, whereas $\text{Cs}^+$ enhances the yield of negative ions. The secondary yields can not be accurately determined theoretically, so the conversion to concentration is done by measuring a “standard”-sample with known concentration, assuming that secondary ion intensity of a given dopant is proportional to the concentration of the dopant.

The secondary ion yield of a particular element depends strongly on its chemical environment, a dependence known as the matrix effect. This effect for P in SiGe has been studied with $\text{Si}_{1-x}\text{Ge}_x$ samples, where $x$ covers the whole range from 0 to 1. The samples were implanted with P, and the relative yield was deduced from the measured depth profiles. The results are shown in Figure 3.12, and for a Ge content of 30% and below, the secondary yield is constant within less than 10%, which means that standard Si samples can be used for the calibration. In Ref. [46] similar results on boron were presented, and it was found that the secondary ion yield and erosion rate differ by less than 10% for a 2-3 keV $\text{O}_2^+$ primary beam and $0 < x < 0.1$.

The sputtered material will consist of atoms and ions from the matrix (e.g. Si, Ge), dopants, primary beam element (O or Cs), and residual elements from the atmosphere (H, C, O, and N), even though the SIMS instrument is operating under ultra high vacuum conditions (~10^{-10} torr). Material is also sputtered in the form of molecules consisting of these elements. Thus, there may be a mass interference with the signal from the dopant of interest. For example, $^{30}\text{Si}^{3+}$ will interfere with $^{10}\text{B}^+$, since the signal is filtered according to the mass to charge ratio (equation (3.3)), and $^{30}\text{SiH}^-$ interferes with $^{31}\text{P}^-$. In some cases, as with B, a different isotope can be chosen, i.e. $^{11}\text{B}$, but this can not be done for phosphorus, where $^{31}\text{P}$ is the only stable isotope.

![Figure 3.11 Erosion rate as a function of Ge content](image1)

![Figure 3.12 Normalized yield for P⁻ as a function of Ge content](image2)
Instead a high mass resolution can be used to resolve $^{30}\text{SiH}^-$ and $^{31}\text{P}^-$, a resolution of $M/\Delta M > 3000$ is required.

The bombardment of primary ions causes a redistribution of the particles within the collision cascade volume. This mixing causes a perfectly abrupt profile to have an exponential trailing edge with a characteristic decay length $\lambda_{\text{mix}}$. Also the leading edge show an exponential behavior, although with a smaller characteristic length than the trailing edge \cite{64}.

In diffusion studies where the intrinsic shape of the profile is important, the relative effect of atomic mixing is minimized by making sure that the diffusion length is much longer than the characteristic decay length. The measured decay length $\lambda_{\text{trail}}$, is given by the intrinsic $\lambda_{\text{intr}}$ and the mixing decay lengths using

$$\lambda_{\text{trail}}^n = \lambda_{\text{intr}}^n + \lambda_{\text{mix}}^n,$$

where $n$ is a value between 2 \cite{64} and 4 \cite{65}. The trailing edge for all the as-grown P and B profiles used in the of the present work is characterized by $\lambda_{\text{trail}} \approx 10 \text{ nm}$, which then can be regarded as an upper limit of the mixing decay length for the SIMS conditions used.
Chapter 4
Dopant diffusion in Si

4.1 Equilibrium diffusion

There is a considerable spread in diffusion coefficients reported for both phosphorus and boron in silicon. This indicates that some of the data have not been obtained under equilibrium conditions. In order to find the “true” dopant diffusion coefficient, determined only by the microscopic properties of the diffusion process, as described in Chapter 2, contributions from non-equilibrium diffusion have to be identified and avoided when possible.

4.1.1 Phosphorus diffusion

The diffusion of phosphorus has been measured using a wide range of experimental conditions applying different analysis techniques, and the extracted activation energies deviate within a range of 1 eV. In this section the results obtained in the present work will be compared with literature data.

A value for the intrinsic diffusivity of phosphorus in silicon ($D^\text{Si}_P$), often used as a reference, is given in the review paper by Fahey et al. [3]. Based on experimental results from several papers (with activation energies in the range of 3.3 eV to 3.7 eV), it was suggested that P diffuses with an activation energy of 3.66 eV. This diffusivity is shown in Figure 4.1; and will in the following be referred to as the “Fahey-diffusivity”. The Fahey-diffusivity does represent a large number of results on P diffusion, e. g. Makris and Masters [66], Ishikawa et al. [67], Lin et al. [68], and Fair and Tsai [69].

The diffusivities determined in the present work are shown as open circles in Figure 4.1; and they obey the Arrhenius dependence

$$D^\text{Si}_P = (8 \pm 5) \cdot 10^{-4} \exp \left( -\frac{2.74 \pm 0.07 \text{ eV}}{kT} \right) \text{ cm}^2 \text{s}^{-1}. \quad (4.1)$$
Clearly, our results do not agree with the Fahey-diffusivity, and neither do an increasing number of publications. For example, the early results of Ghoshtagore [70], who used a P doped epi-layer with a concentration less than $2 \times 10^{19} \text{ cm}^{-3}$ as a diffusion source and high annealing temperatures (1130-1405°C) to assure intrinsic conditions. The results of Ghoshtagore’s measurements are shown in Figure 4.1 together with an extrapolation to lower temperatures. Hill [71], who used an ion implanted P profile as diffusion source, found values which agree with Ghoshtagore’s low temperature extrapolation. Further, Pelleg and Ditchek [72], Wittel and Dunham [73] and Haddara et al. [74] have reported data different from the Fahey-diffusivity. The results of Pelleg and Ditchek [72] show an overall poor Arrhenius dependence when the whole temperature range of 765 – 1230 °C is considered. For the high temperature results Pelleg and Ditchek derive an activation energy of 3.64 eV. The low temperature diffusivities are enhanced as compared to those expected when extrapolating the high temperature diffusivities. The authors have no explanation for this apparent enhancement, but comparing with our results, there is, in fact, no enhancement. Wittel and Dunham [73] and Haddara et al. [74] have reported $D_p$ values coinciding with the results of the present work. Both Ref. [73] and Ref. [74] use P profiles with peak concentrations below $3 \times 10^{18} \text{ cm}^{-3}$ to assure intrinsic diffusion conditions. In addition, Haddara et al. [74] found no time dependence of $D_p$ at 850 °C for annealing durations greater than 10-16.

![Figure 4.1](image)  

**Figure 4.1** Arrhenius plot of the P diffusivity in silicon obtained in this work compared with other published results.
up to 5 days. The recent results of Zangenberg et al. \cite{75}, who used in-situ doped P profiles in MBE grown samples, also agree with the P diffusivities found in the present work.

A review of the literature does not reveal any systematic relation between the experimentally obtained P diffusivities and the specific experimental conditions used in each study. For example, Haddara et al. \cite{74} and Wittel and Dunham \cite{73} obtained similar diffusivities using different purity silicon, Cz and FZ, respectively. In the present work we compared P diffusion in CVD grown Si and in Cz Si and found no difference in the diffusivity [paper I]. Hill \cite{71} and Ishikawa et al. \cite{67} measured the diffusivity in different lattice orientations, and no difference between the <111>, <110>, and <100> orientations. The experimental techniques used in dopant diffusion studies in Si are (i) a marker technique \cite{71}, where the dopant profile is determined using a copper stain on a beveled sample, (ii) profiling using radioactive tracers \cite{70,66}, (iii) profiling by electrical measurements \cite{67,68,70}, and (iv) SIMS profiling \cite{72–75}.

The diffusivity may especially be sensitive to the methods by which the samples have been doped with phosphorus. The best way to incorporate P into the lattice is by in-situ growth of a doped epi-silicon layer \cite{70,75, present work}. Ion implantation is a widely used technique \cite{68,71,73,74}, but in order to avoid non-equilibrium effects, the implantation has to be followed by an anneal to reduce lattice damage, and the time dependence of the diffusivity should be investigated for any transients (see Section 4.2.2). Finally, a method commonly used in the early experiments were in-diffusion of P from a vapor source or a source deposited on the wafer surface \cite{66,67,69}. Doping by in-diffusion is often an unfavorable method, because it generally results in profiles with very high phosphorus concentrations at the surface, possibly containing immobile phosphorus precipitates, a fast diffusing tail due to non-equilibrium conditions, and a kink which divides the two diffusion regimes. Examples of some profiles obtained by in-diffusion are shown in Figure 4.2.

![Figure 4.2 Example of P diffusion profiles using a surface source. (from Ref. [77])](image)

In an early attempt to describe these complicated profiles, Fair and Tsai \cite{69} proposed a model based on a vacancy mechanism for P diffusion, where the P-atoms diffuse through a PV complex, which, according to the
model, could exist in different charge states, with different diffusivities. This model
could describe many of the features observed. For instance, in accordance with the
model in Ref. [69], the P concentration in the kink corresponds to a Fermi-level
position which coincides with a change of charge state of the diffusing complex.
However, the model can not explain the so-called emitter-push effect, an effect of
enhanced boron diffusion in buried boron doped layers. Hu et al. [76] suggested that
an interstitial mechanism should be included in the P diffusion model, and a similar
proposal has been made by Uematsu [77]. Specifically, in the Uematsu model P
diffusion takes place via V2- and the pairing between P and I0 in highly doped regions
and in intrinsic regions, respectively. This model has proven to be successful in
describing the complex diffusion profiles with kinks, fast diffusing tails, and emitter
push. Simulations using this model indicate that in the tail region an equilibrium
between I, PI and substitutional P is not reached, and a supersaturation of interstitials
occurs, thus explaining the emitter-push effect.

Despite of the success of the phenomenological models, the type of profiles shown in
Figure 4.2 is not suitable for determining the intrinsic diffusivity, mainly because
many parameters are required in the models describing the profiles. For example,
Uematsu uses 5 parameters in the simulation, to describe to diffusion profiles in Figure
4.2. This underlines the importance of designing samples with profiles that are simple,
like samples with P doped epi-layers or pre-annealed ion-implanted phosphorus, where
the effects on the diffusion associated with high P concentration are avoided.

As described in Section 2.3 the Fermi-level dependence on the diffusivity may be
given as a function of the free carrier concentration:

\[ D'_{\text{eff}} = D^0 + D^+ \left( \frac{n}{n_i} \right) + D^- \left( \frac{n}{n_i} \right)^2, \]  (4.2)

where \( D^0, D^+ \) and \( D^- \) are the diffusivities for P+ paired with Si self-interstitials in
different charge states, and generally, they have different prefactors and activation
energies. Fair and Tsai [69] found that all the three terms contributed to the effective
diffusion coefficient, but the activation energies for \( D^- \) and \( D^- \) were so high that \( D^0 \) can
be regarded as the effective diffusivity under intrinsic conditions.

Fair and Tsai [69] generally based their model on P profiles (as the ones shown in
Figure 4.2), where the point defects participating in the diffusion process are unlikely
to reach equilibrium, because the defect concentrations are influenced by the diffusion
process, and not only by the Fermi-level position. To avoid those complications, the
diffusivity can be determined using a constant background of some other dopant,
which controls the extrinsic carrier concentration. This was done by Wittel and
Dunham [73] who found that a boron background retards the P diffusion. Wittel and
Dunham assumed that this is due to a pairing between the P and B atoms, rather than a Fermi-level effect, but they presented no good arguments for this assumption. In an As background (i.e. n-type material) they found that the P diffusion is enhanced. John and Law [78] came to the same conclusion in a similar experiment. Wittel and Dunham concluded that the effective diffusivity is given by the constant term and the term linear in $n$ in equation (4.2), whereas John and Law found that is it the constant term and the term quadratic in $n$ which determine the effective diffusivity. In either case, the diffusion coefficient is definitively dependent on the concentration of free charge carriers.

Using the unstable $^{32}\text{P}$ in a constant background of $^{31}\text{P}$ to determine the diffusion in extrinsic material, Makris and Masters [66] found that the effective diffusion coefficient could be explained by only the term linear in $n/n_i$, i.e.

$$D_p^{\text{eff}} = D_p \left( \frac{n}{n_i} \right). \tag{4.3}$$

Uematsu [77] used the same $n$ dependence for the diffusivity of P via the I mechanism, which in his model dominates at low P concentration. The very simple model given by equation (4.3) can be tested on the experimental results of Refs. [73] and [78], without the use of any free parameters. Using the P diffusion coefficient obtained in our work, we have found that the model fits well with the results of Wittel et al. and reasonably well for the low temperature measurements of John et al., as illustrated by Figure 4.3 and Figure 4.4: If the P diffusion coefficients reported by us are reviewed with this model in mind, the values should be divided by $n/n_i$ in order to find the Fermi-level independent value of $D_p$. Using $C_p = 7 \times 10^{17}$ cm$^{-3}$, which is the concentration of the plateau of the phosphorus profile (Paper I, Fig. 1), the values will decrease by ~20% for the low temperatures and ~5% for the high temperatures. This will, however, have only a minor effect on the extracted activation energy (from 2.74 eV to 2.81 eV for the corrected values).
It is not possible to deduce the individual terms in the activation energy (equation (2.19)) from the experiments described above, and thus, to study and understand diffusion at an atomic level, theoretical calculations are needed. Nichols et al. [19] used DFT with the local density approximation (LDA) to calculate activation energies for B, P, As, and Sb. For phosphorus they found that the vacancy mechanism has the lowest activation energy, 3.4 eV, while for the kick-out mechanism they calculated an energy of 3.8 eV. According to their calculations the interstitialcy mechanism (termed by the authors as “coordinated push”) generally can be ignored because the activation energy for this mechanism is larger than the one for the kick-out mechanism. However, they only consider neutral defects and thus do not account for any coulomb interaction which may increase or decrease the binding energy of dopant-point defect complexes.

In a recent report, Liu et al. [79] published results of DFT calculations on substitutional P in silicon interacting with both interstitials and vacancies. They identified the four lowest energy structures for the PI complex (Figure 4.5). The complex with the lowest formation energy in intrinsic material is labeled P_X, and is a <110> dumbbell interstitial, where a phosphorus atom and a silicon atom share a lattice site, a structure similar to the Si self-interstitial with the lowest formation energy (see section 2.4). P_X has a similar structure as P_X, but with a local structural distortion, and has the same binding energy at midgap level. P_H, where the P atom is in the hexagonal interstitial site, and P_S, a dumbbell structure oriented in the <100> direction have binding energies of 0.5 eV and 0.6 eV, respectively, and they are the most stable P-I complexes with a positive charge.

The energetics of the diffusion process, according to Ref. [79], is shown in Figure 4.6 for the neutral and the positively charged cases. The neutral P_X migrates by changing into a P_H and back to a P_X in another lattice site, and so the barrier for this process is

![Figure 4.5 Formation energy for PI complex relative to substitutional P+ (from Ref. [79]). The vertical lines indicate the position of the Fermi-level at 900 °C for n/n_I = 0.1 and 10, respectively.](image)

![Figure 4.6 Energy diagrams for the P diffusion process based on the results in Ref.[79].](image)
the difference between the formation energy for the two complexes which is 0.6 eV. If the complex is allowed to change charge state during the process, i.e., into the $P_{1}^{H+}$ complex, the barrier is lowered to 0.2 eV as indicated by the broken line in Figure 4.6. Using $\Delta H_f = 3.7$ eV the overall activation energy for diffusion through the neutral PI complex is 3.5 eV (or 3.1 with charge variations).

The $P_{1}^{H+}$ complex has a binding energy of 0.6 eV, and it has the lowest formation energy of the charged complexes. It migrates by changing to the $P_{1}^{X+}$ that has an energy of formation which is 0.3 eV larger. This lead to an overall activation energy of 3.4 eV for the positively charged complex (see Figure 4.6).

The DFT calculations suggested a migration energy for the negatively charged PI complex of 1.4 eV, and this is so large, relative to the neutral and positively charged complexes, that its contribution to the diffusion can be neglected. Liu et al. [79] also did calculations on diffusion via the vacancy mechanism, which suggested that it does not play a role in the intrinsic region, but the $P^{+}V^{-}$ complex may contribute in heavily $n$ doped regions. These calculations support the model for the P diffusivity given by equation (4.3). In Figure 4.5, it is seen that at the Fermi-level positions which correspond to $n/n_i$ of 0.1 – 10 the formation energy of $P_{1}^{X}$ is the lowest, since the $P_{1}^{X_2}$-complex does not contribute to the diffusion. This implies that the Fermi-level dependence of the P diffusivity will be given by the Fermi-level dependence of the formation energy of $P_{1}^{X}$.

### 4.1.2 Boron diffusion

Boron diffusion in silicon has, like phosphorus, been studied extensively, due to its importance as a $p$-type dopant in device technology. Similar to that for phosphorus a diffusion model which includes all observed phenomena is not available. In this thesis, equilibrium and intrinsic diffusivities are treated, as well as enhanced diffusion during proton irradiation.

Some of the papers, concerned with phosphorus diffusion, described in the previous section, also include results on boron diffusion, and they are summarized in the Arrhenius plot in Figure 4.7. Fahey et al. [3] reported activation energies for boron diffusion between 3.25 – 3.87 eV under intrinsic conditions. The diffusivities from the present work are represented in Figure 4.7 by the open circles, and the results are fitted by the Arrhenius expression

$$D_B = (0.06 \pm 0.02) \exp \left( \frac{-3.12 \pm 0.04 \text{ eV}}{k_B T} \right) \text{cm}^2/\text{s}.$$  (4.4)
Even though P and B diffuse through the same mechanism, similar experimental conditions do not give similar results, as it is evident from the Arrhenius plot in Figure 4.7. For example, both Haddara [74] and we [Paper II] have determined both P and B diffusivity, but whereas the P diffusivity results in the two works agree (Figure 4.1), there is a large difference in the B diffusivities.

As for phosphorus, boron has been suggested to diffuse via point defects in different charge states, giving a Fermi-level dependence similar to the one for P diffusion

\[ D_{\text{eff}} = D_0^B + D_n^B \left( \frac{P}{n_i} \right). \] (4.5)

The Fahey values for \( D_0^B \) and \( D_0^+ \) are of equal magnitude, and for \( p < 4n_i \) the effective diffusivity will vary by less than a factor of 2. However, Willoughby et al. [80] found that in n-type material the boron diffusivity was retarded by up to a factor of 10, which suggest that \( D_0^B \) is small. Furthermore, the results of Willoughby et al. [80] and Miyake [81] shown in Figure 4.8 indicate that \( D_0^B \) is negligible for \( p > n_i \). The lines in Figure 4.8 represent the effective B diffusivity using only the \( D_0^+ \) term in equation (4.5). Although, in this case \( D_0^+ \) has been used as a fitting parameter, and it is
generally about 2 times smaller than the values reported in the present work. However, as will be seen in a following section, the boron diffusion is very sensitive to the annealing conditions.

In the next section it will be mentioned that presence of carbon retards both the B and P diffusion due to an undersaturation of Si self-interstitials caused by the carbon atoms. Other impurities may cause similar perturbations of the point defect concentrations, and for example, the retardation of B diffusion in a background of As could be explained by an undersaturation of interstitials. However, this would have a similar effect on P diffusion in contrast to the results of Wittel and Dunham and John and Law. This indicates strongly that the effect can be attributed to the Fermi-level effect.

So boron diffusion can seemingly, as for phosphorus, be described with a single diffusivity which for the intrinsic case equals \( D_B^{0} \). With this model the results of the present work can be reviewed, by calculating \( D_B (p/n_i) \), using \( C_s = 2 \times 10^{18} \text{ cm}^{-3} \) (which is the boron concentration in the box profile) and we find the following prefactor \( D_B^{0} = 0.34 \text{ cm}^2/\text{s} \) and \( E_a^{B} = 3.35 \text{ eV} \). This is quite different from the Arrhenius dependence given in Paper II and by equation (4.4), which illustrates the importance of knowing the correct Fermi-level dependence when interpreting B diffusivities.

Results of DFT calculations on boron diffusion have been reported by several authors in the last five years. Some results of these calculations are summarized in Table 4.1, where a spread of 1.3 eV in the calculated \( E_a^{B} \) values is revealed. It is also seen that the majority of the authors find that boron diffuses through the interstitialcy mechanism, which in general was ruled out by Nichols et al. [170].
Windl et al. [83] found that the boron diffusion is dominated by a neutral BI complex, which is consistent with a linear dependence of the B diffusivity on the concentration of free charge carriers. In contrast, Hakala et al. [84] found that the neutral complex is metastable, and B diffuses through a positively charged complex giving a quadratic dependence on $p$. Further, they also found that the migration energy is very small, in conflict with Watkins [85], who measured a migration energy of 0.6 eV.

The results by Windl et al. [83] are obtained by the same method as used by Liu et al. [79] for P diffusion. Thus, it is tempting to compare the two values for the obtained activation energies; $E_a^B$ is at least 0.25 eV larger than $E_a^P$, consistent with the experimentally determined activation energies in the present work.

### 4.2 Non-equilibrium diffusion

The scatter in the diffusivities observed in Figure 4.1 and Figure 4.7 indicates that some of the experiments have been performed under non-equilibrium conditions. Some conditions, which cause deviation from equilibrium, are already mentioned in Section 2.5. This section treats non-equilibrium diffusivities associated with surface reactions, ion implantation, and transient effects due to short time annealing.

#### 4.2.1 Surface reactions

Silicon is inclined to react with certain elements during annealing. Some of these reactions have been studied extensively, because they form insulating or electrically conductive layers used for integrated circuits on a chip.

Dry or wet oxidation, i.e., annealing of Si in O$_2$ or water vapor, respectively, form a layer of insulating SiO$_2$ on the surface, with a thickness depending on temperature and annealing duration [86]. Oxidation is associated with injection of Si self-interstitials. This was recognized from the fact that stacking faults of an interstitial kind were

<table>
<thead>
<tr>
<th>Ref</th>
<th>$E_{BI}^{bind}$</th>
<th>$\Delta H_{BI}^{m}$</th>
<th>$E_a^B$</th>
<th>Method</th>
<th>Mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>[19]</td>
<td>0</td>
<td>3.9</td>
<td>LDA</td>
<td>Kick out</td>
<td></td>
</tr>
<tr>
<td>[82]</td>
<td>0.9</td>
<td>0.7</td>
<td>3.0</td>
<td>LDA</td>
<td>Interstitialcy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.4</td>
<td>GGA</td>
<td></td>
</tr>
<tr>
<td>[20]</td>
<td>0.2</td>
<td>0.8</td>
<td>3.8</td>
<td>LDA</td>
<td>Kick out</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.3</td>
<td>GGA</td>
<td></td>
</tr>
<tr>
<td>[83]</td>
<td>0.5</td>
<td>0.4</td>
<td>3.5</td>
<td>LDA</td>
<td>Interstitialcy</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.2</td>
<td>3.8</td>
<td>GGA</td>
<td></td>
</tr>
<tr>
<td>[84]</td>
<td>0.2</td>
<td>0.2</td>
<td>4.2</td>
<td>LDA</td>
<td>Interstitialcy</td>
</tr>
</tbody>
</table>

Table 4.1 Summary of theoretical B diffusion energies for intrinsic material.
formed in the Si lattice \[87\] and also because an enhanced diffusion of boron and phosphorus was observed during oxidation \[3\]. The excess of self-interstitials occurs when O\(_2\) and Si react to form a surface layer of SiO\(_2\). The larger volume of SiO\(_2\) causes a strain at the interface, which is relieved by creating a Si self-interstitial. The enhancement of diffusion of interstitialcy mediated diffusers during oxidation is known as oxidation enhanced diffusion (OED). Sb diffusion, on the other hand, is retarded during oxidation, because the supersaturation of self-interstitials lead to an undersaturation of vacancies. This effect is called oxidation retarded diffusion (ORD).

Nitridation and silicidation of a Si surface are, like oxidation, frequently used in Si processing. They may cause a retardation of B diffusion in Si and an enhancement of Sb diffusion \[3, 88\]. This is consistent with a supersaturation of vacancies and an undersaturation of self-interstitials, so nitridation and some silicidations can be associated with an injection of vacancies. A nitride film even retards B diffusion after it has been formed, an effect which has been attributed to either depletion of self-interstitials or supersaturation of vacancies caused by stress exerted on the nitride film. This effect can be avoided with a SiO\(_2\)-layer in between the nitride and the silicon layers \[89\].

OED and ORD are signatures of diffusion via interstitials and vacancies, respectively, so oxidation can be used to determine mechanisms of diffusion in Si. However oxidation results can only determine the mechanism qualitatively. Quantitatively, the mechanism can be determined by combining oxidation and nitridation results with diffusion result obtained with inert annealing of several dopants. This was done by Ural et al. \[9\], who measured the diffusion of Si, P, B, As, and Sb in Si. For each element and annealing condition, an equation like \(2.26\) can be set up, and by solving this set of equations the interstitialcy fraction and excess point defect concentrations are deduced.

In the present work OED of P has been measured, and any effect of non-deliberately oxidation, which may take place if the inert N\(_2\)-annealing atmosphere contains a small concentration of O\(_2\) has been considered. Specifically, we have found that oxidation at 810°C for 3 hours caused an increase in the P diffusivity by a factor of 5. This oxidation resulted in a SiO\(_2\) layer of a thickness of 150 Å [Paper I]. The oxide thickness was measured also on the samples annealed in a N\(_2\) atmosphere before and after anneal, and it was found that the oxide increased by less than 15 Å. Compared with the enhancement observed during the intentional oxidation, an unintentional oxidation resulting in a 15 Å thick oxide will cause an enhancement in the diffusivity of approximately 5 %, which is within the experimental uncertainty of the determined diffusivities.
Vacuum annealing may also affect the concentration of Si self-interstitials. In Paper I a retardation of P and B diffusion compared to N\textsubscript{2} annealing is observed (see Figure 4.9). Gossmann \textit{et al.} \cite{90} have, similarly, observed that B diffusion is retarded and that the retardation increased with time. Retardation of B and P means that the self-interstitial concentration is depleted during the vacuum annealing. There is yet no good explanation of the mechanism for this depletion, but it was conjectured in Ref. \cite{90} that silicon atoms on the surface evaporate due to reactions in the native oxide or with residual gases. Gossmann \textit{et al.} also studied Sb diffusion, and they found that vacuum annealing did not affect the diffusion compared to inert annealing. This means that the process responsible for the I depletion does not influence the V concentration. So vacuum annealing is a situation where the mass action law \cite{288} is not valid. The vacuum annealing results can not be explained as an effect of different pressures as treated in Section 2.6 – with the value of \(\Delta V_a\) measured for boron the difference in boron diffusivity at 0 atm and 1 atm will only be 0.002 \% at 900 °C.

### 4.2.2 Ion implantation

As mentioned in the preceding chapter, the diffusion of implanted dopants has been shown to be a complicated matter. For example, annealing of an implanted boron profile shows an initial fast diffusion, with an enhancement of a factor of 100 or more compared to the equilibrium diffusivity. This phenomenon is called transient enhanced diffusion (TED). The fast initial diffusion is caused by the point defects created by collisions between the energetic ions and atoms in the host lattice. The excess concentration of point defects eventually disappears due to annihilation or recombination at the sample surface, and the dopants will ultimately reach their equilibrium diffusivity \cite{91}. 

Figure 4.9 The effect of annealing in vacuum on B and P diffusion compared to N\textsubscript{2} ambient annealing.
In general, TED depends on annealing temperature, on implantation dose and energy. Also, at high concentrations the implanted dopants may form clusters, which are not likely to participate in the diffusion process. The combination of all these effects means that modeling of TED is a difficult task. A controlled way to study the effect of the excess point defect concentration on dopant diffusion is by using high energy proton irradiation at elevated temperatures. The protons generate self-interstitial-vacancy pairs with a uniform distribution in the part of the sample where the diffusion is studied [Paper VI]. As will be discussed below, diffusion of some dopants is enhanced due to the irradiation. This phenomenon is referred to as radiation enhanced diffusion (RED).

RED of B δ-spikes in samples irradiated at 570 °C was found to be enhanced by as much as a factor of $3 \times 10^5$. The duration of the irradiation and the sample temperature were such that the boron diffused according to Cowern’s intermittent model (Section 2.2.1), which means that it is possible to extract the $g$ and $\lambda$ parameters [Paper IV]. $\lambda$ was found to decrease with increasing temperature. This is in accordance with the results of Cowern et al. [92]. However, the absolute $\lambda$ values are a factor of 4 larger than those determined for B diffusion during an oxidizing annealing. It is suggested that the difference is caused by a lowering of the migration energy of the BI complex as the proton irradiation may be responsible for an ionization enhanced diffusivity.

Both $g$ and $\lambda$ depend on the maximum concentration of the boron spike [Paper V]. If this concentration is larger than the intrinsic carrier concentration both parameters decrease, which according to equation (2.16) means that the effective B diffusivity decreases. This effect is not fully understood, but it is conjectured that the high boron concentration affects the charge state of the self-interstitials. The correlation with $n_i$ may at first sight be in contradiction with the Fermi-level effect discussed in Section 4.1.2 but it should be kept in mind that this discussion is based on the assumption of equilibrium. Proton irradiation does definitively not satisfy this assumption.

In addition to a concentration dependence, a small depth dependence of the diffusion is observed. The diffusion in the boron spikes close to the surface is larger than the diffusion in the deeper spikes. This depth dependence is attributed to a depletion of vacancies at the surface, which means that the self-interstitials are less likely to recombine with a vacancy close to the surface than in the bulk. Zangenberg [89] observed a similar depth dependence for boron spikes for a 30 min annealing in an inert gas at 850 °C. For annealing durations of 1 hour and more an opposite depth dependence was observed, indicating that spatial and temporal variations of point defect concentrations are crucial also for diffusion under inert annealing conditions.
The RED results described above were obtained for samples in which only one dopant (boron) interacted with the point defects created by the irradiation. RED experiments were also performed with samples containing spikes of both boron and antimony (see Figure 4.10), in order to investigate the effect of an interaction with two different dopants [Papers V and VI]. A surprising result is that for the irradiation conditions used Sb diffuses only if the sample also contains B spikes, while in a sample containing only a Sb spike no diffusion is observed. Also the boron diffusion is affected by the presence of Sb spikes, with a slight increase in samples with both dopants. However, when the maximum boron concentration exceeds the intrinsic carrier concentration the small increase is overwhelmed by the concentration dependent decrease.

The reactions, which govern the diffusion, are:

\[
\begin{align*}
I + V & \rightarrow \text{Si}_s, \quad (4.6a) \\
\text{Sb}_s + V & \leftrightarrow \text{SbV}, \quad (4.6b) \\
\text{SbV} + I & \rightarrow \text{Sb}_s, \quad (4.6c) \\
\text{B}_s + I & \leftrightarrow \text{BI}, \quad (4.6d) \\
\text{BI} + V & \rightarrow \text{B}_s. \quad (4.6e)
\end{align*}
\]

Reaction (4.6e) affects the recombination rate of the BI complex and thus causes an increasing $\lambda$ with decreasing vacancy concentration. However, Cowern et al. [92] showed that $\lambda$ for boron diffusion is the same for inert annealing and for oxidizing conditions, where an undersaturation of vacancies is established. Hence, in a first approximation reaction (4.6e) may be neglected, but the corresponding reaction for Sb
diffusion (4.6c) cannot be neglected with similar arguments. In fact, reaction (4.6c) may explain that Sb only diffuses when boron is present, while B is only slightly affected by the presence of antimony.

4.2.3 Time dependent diffusion

It is usually assumed that, when a Si sample is heated up, the concentration of self-interstitials reaches its equilibrium value instantly. This assumption is probably not valid when short annealing durations are considered, as indicated by the results of Fang et al. [29]. They found that the boron diffusion is initially suppressed at 850°C, a similar effect has been observed in the present work [Paper II]. The results of these two works are shown in Figure 4.11, where it is seen that the boron diffusion requires annealing durations on the order of hours to reach equilibrium. All the boron diffusivities have been measured in B doped epi-layers grown on different types of substrates. The type of substrate affects the dynamical behavior of the diffusivity, as also shown in [Figure 4.11], Fang et al. [29] demonstrated that the diffusivity reaches equilibrium faster if the epi-layer is grown on a Cz substrate than on a FZ substrate. Further, an initial enhancement of the B diffusion is observed if the Cz substrate is annealed 8 h at 750 °C prior to the growth of the epilayer. This type of sample is referred to as LH-Cz in [Figure 4.11].

The observations in [Figure 4.11] have, so far, not been explained quantitatively. However, a model involving the concentration of self-interstitials, may provide some idea about the value of the physical parameters involved in the generation and
recombination the self-interstitials. Let us assume that the barriers $\Delta H_1^f$ and $\Delta H_1^r$, as defined in Figure 4.12(a), control the concentration of self-interstitials $C_I$. The generation rate, $g$, depends on the total barrier height for the process $\text{Si}_i \rightarrow \text{Si}_I$ ($\Delta H_1^f + \Delta H_1^r$). The recombination rate, $r$, is given by the barrier for the reverse process combined with the probability that the substitutional site is vacant, which is proportional to $C_V$. Hence, we obtain

$$g = \nu \exp\left(-\frac{\Delta H_1^f + \Delta H_1^r}{k_B T}\right) \quad \text{and} \quad r = \nu \exp\left(-\frac{\Delta H_1^r}{k_B T}\right) \frac{C_V}{N}.$$  \hspace{1cm} (4.7)

For simplicity, $C_V$ is considered to be time-independent, so the $C_I$ and $C_V$ are uncoupled. This approximation means that $r$ represents an upper limit of the actual recombination rate. The $I$ concentration is given by the differential equation

$$\frac{\partial C_I}{\partial t} = gN - rC_I. \hspace{1cm} (4.8)$$

The solution of this equation (boundary conditions: $C_I(t=0) = 0$ and equilibrium value $C_I(t=\infty) = C_I^*$) becomes

$$\frac{C_I(t)}{C_I^*} = 1 - \exp(-rt), \text{ where } C_I^* = \frac{gN}{r}. \hspace{1cm} (4.9)$$

From equation (4.9) it is clear that the time evolution of the $I$-concentration only depends on the barrier for recombination, which is not related to the enthalpy of formation of I. In Figure 4.12(b) the normalized time averaged B diffusion coefficient calculated from (2.26) with $C_i(t)$ given by equation (4.8), assuming $\Delta H_1^f = 0.65$ eV,
and using the theoretical results of Ref. \cite{16} for $C_V$, is shown together with the experimental values.

The model does not consider any other sources of self-interstitials than bulk silicon, while in reality, the Si surface, impurities, and structural defects may act as sources or sinks. Despite this simplification, the model shows that a recombination barrier smaller than 1 eV can account for transients with a duration of hours. That such a barrier may exist was suggested by Antoniadis and Moskowitz \cite{93}, who found that their diffusivities of P, B, As and Sb for short oxidation times could be explained by assuming an energy barrier for IV recombination of the order of 1 eV.

The fact that a more accurate model should include the surface as a source of point defects is illustrated by Fang et al. \cite{29} (see Section 2.4). The dependence on substrate as indicated in Figure 4.11 is attributed to different concentrations of O indicating that the model should include interaction with impurities as well. This is further supported by Rücker et al. \cite{32}. They used silicon samples doped with carbon and found that P and B diffusion were retarded while that of As and Sb were enhanced when carbon was present in high concentration. The retardation of B and enhancement of Sb diffusion due to the high C concentration was found to be restricted to about 12 h at 900 °C, after this initial period the diffusivities reached their equilibrium values. Further, the initial retardation of the B diffusion increased with increasing C concentration (Rücker et al. \cite{32}). These results are explained by an initial trapping of self-interstitials by the substitutional carbon atoms.

The presence of C and O affect the concentration of self-interstitials in Si, and CVD grown epi-layers generally has a lower concentration of these impurities than Cz Si. However, P diffusion in an epi-layer and in a Cz substrate in a temperatures of 810 – 900 °C were found to be the same within the experimental accuracy \cite{Paper I}, suggesting that for these particular Cz samples the impurity concentrations were not sufficiently high to influence the trapping of self-interstitials.
Chapter 5

Dopant diffusion in SiGe

Up until now only B and Sb diffusion in SiGe have been thoroughly studied, and the intention of this work is to add P diffusion to that list. In this chapter the results on P diffusion will be presented as well as what is known about diffusion of B and Sb. First, the composition effect will be treated, and then the effect of an applied strain is described, before the diffusion mechanism and the segregation phenomenon in Si/SiGe heterostructures are discussed. Finally, results on P diffusion in SiGe samples with imperfect crystallinity are presented.

5.1 Composition effect on dopant diffusion

Phosphorus diffusion in relaxed $\text{Si}_{1-x}\text{Ge}_x$ has been studied in samples with $x = 0.11$ and $x = 0.19$. The results are shown in Figure 5.1 together with the results of Zangenberg et al. [10] for $x = 0.11, 0.19$ are from Paper III and $x = 0.07, 0.12, 0.24, 0.40$ are from Ref. [75].
al. [75], which, so far, represents the only other study of P diffusion in relaxed, epitaxially grown Si$_{1-x}$Ge$_x$. The results from Ref. [75] show an overall increase in diffusivity with Ge content in Si$_{0.88}$Ge$_{0.12}$ and Si$_{0.76}$Ge$_{0.24}$ but in Si$_{0.60}$Ge$_{0.40}$ a decrease in P diffusivity is observed.

The maximum concentration of the P profiles used as the initial profiles is $6 \times 10^{18}$ cm$^{-3}$ (see Figure 3.1 and Figure 3.2). Assuming that the intrinsic carrier concentration in SiGe is given by equation (2.32), using the change in band gap ($\Delta E_g \sim 0.05$ eV (Si$_{0.89}$Ge$_{0.11}$), $\Delta E_g \sim 0.08$ eV (Si$_{0.81}$Ge$_{0.19}$)), the value ($n/n_i$) will vary between 2.6 – 1.5 (Si$_{0.98}$Ge$_{0.11}$) and 2.4 – 1.4 (Si$_{0.81}$Ge$_{0.19}$) in the temperature interval studied. Thus, if P exhibits a diffusivity similar to the one given in equation (4.3), the extracted values should be corrected accordingly. Including these corrections will, however, not change the conclusions made below, and since there exists no measurements of the Fermi-level dependence of the P diffusion in SiGe, the diffusion is assumed to occur under intrinsic conditions.

From the results in Figure 5.1, the prefactor and the activation energy can be calculated, and these are shown in Figure 5.2 as a function of the Ge content. The prefactor for the diffusivity increases with the Ge content, and this can be modeled by assuming that it is given by a term similar to the entropy of mixing. The activation energy also increases with increasing Ge content, which is opposite to that expected from the discussion in section 2.7.1. There it is argued that the formation enthalpy of point defects decreases with increasing $x$. However, the activation energy consists of other contributions than the formation enthalpy of point defects. The overall increase in activation energy may be explained by the elastic properties of SiGe, as outlined in Paper III.

Figure 5.2 (a) Prefactors versus Ge content fitted to a entropy of mixing term dependence. (b) Activation energy for P diffusion versus Ge content shows a linear dependence.
Boron diffusion in relaxed SiGe has also been investigated in this study, although to a lesser extent than phosphorus diffusion. The boron diffusivities in SiGe, measured in a temperature interval of 850 – 950 °C, are shown in Figure 5.3 and it is found that the B diffusion decreases with increasing Ge content. This is in agreement with the observations by Kuo et al. [95], which are depicted in Figure 5.4. Kuo et al. only measured the diffusivity at 800 °C, so they were unable to determine prefactors and activation energies for the diffusion. Zangenberg et al. [75] measured B diffusion in relaxed Si\(_{0.88}\)Ge\(_{0.12}\) and Si\(_{0.76}\)Ge\(_{0.24}\) in a temperature interval of 800 – 925 °C. They found only a small Ge concentration dependence on the individual diffusivities in the temperature interval, Figure 5.3. However, they did find that adding 1% Ge increased the activation energy by almost 0.5 eV, which then remained essentially constant in Si\(_{0.88}\)Ge\(_{0.12}\) and Si\(_{0.76}\)Ge\(_{0.24}\).

As described in the previous chapter, phosphorus and boron diffuse by the same mechanism in Si and they respond to non-equilibrium conditions in a similar manner. Hence, it might be expected that the same would be the case in SiGe alloys with moderate Ge content. This is, however, not true as seen by comparing the results in Figure 5.1 with those in Figure 5.3. Lever et al. [96] suggested that the decrease was due to pairing of boron and germanium, as they observed a pileup of B in the SiGe layer in a Si/SiGe/Si heterostructure. Although Chen et al. [45] have shown that the pileup could just as well be explained by segregation without any difference in the diffusivities of B in Si and SiGe, the results obtained in relaxed SiGe do suggest some interaction between B and Ge atoms. Especially, since Zangenberg et al. found that in

![Figure 5.3 Arrhenius plot of the boron diffusivity in SiGe](image-url)
Si$_{0.99}$Ge$_{0.01}$ the activation energy for the boron diffusion increased substantially relative to that in pure Si.

The effect of the Ge content on both the P and B diffusion is summarized in Figure 5.4, where the P and B diffusivities at 800 °C versus Ge content are depicted. There is a general trend that the P diffusion increases and the B diffusion decreases when the Ge concentration is increased, although there is a disagreement in the magnitude of this effect. Zangenberg et al. [75] see only a minor decrease in the B diffusivity, while Kou et al. [95] found that the diffusivity decreased by a factor of 6. By extrapolating our results in Figure 5.3 to 800 °C we find a diffusivity in between the values of Ref. [95] and Ref. [75].

The increase in the P diffusivity with Ge content observed by Zangenberg et al. is also less pronounced than that observed in the present work, and there is even a decrease when going from Si$_{0.76}$Ge$_{0.24}$ to Si$_{0.60}$Ge$_{0.40}$. Interestingly, this decrease is inherent in our diffusion model, with an increasing prefactor due to an entropy of mixing term and an increasing activation energy. In Figure 5.4, the phosphorus diffusivity, according to the model based on the parameters in Figure 5.2, is represented by the thick dotted line, and it is seen that it follows the trend of the experimentally determined P diffusivities. According to Södervall and Friesel [43], the P diffusivity in pure Ge is $5 \times 10^{-13}$ cm$^2$/s at 800 °C, so the model is not valid in the whole range of Ge contents, which may be a consequence of a change in the diffusion mechanism, from a pure interstitialcy one in Si to a pure vacancy one in Ge [43].

![Figure 5.4 Comparison of the composition effect for boron and phosphorus diffusivities. Data from this work has been extrapolated to 800 °C.](image-url)
To complete this account of dopant diffusion in relaxed SiGe, it should be mentioned that for Sb the chemical effect has also been investigated using samples with Ge contents between 0% and 50%. The results showed a decrease in activation energy with increasing Ge fraction toward the value for Sb diffusion in pure Ge [97], and the prefactors exhibit an increase in accordance with the entropy of mixing term, as suggested in [Paper III], and similar to Figure 5.2(a).

5.2 Effect of strain

Results for P diffusion in strained SiGe layers in Si/SiGe/Si heterostructures are shown in Figure 5.5 for a temperature interval of 810 – 900 °C. It is seen that P diffuses faster in strained SiGe than in Si, but if the values in Figure 5.5 are compared with those in Figure 5.1, it is clear that compressive strain causes the P diffusivity to decrease. In order to minimize the relaxation of the strained layers, annealing of the samples has been restricted to a narrow temperature interval. As a consequence, the diffusivities do not show a clear Arrhenius dependence, which makes it difficult to deduce reliable activation energies for the diffusion in strained SiGe layers. However, activation energies have been estimated based on the results in Figure 5.5 by assuming that the strain does not affect the prefactors of the diffusivities, leaving only the activation energy as a free parameter when using the prefactors obtained for relaxed SiGe (Figure 5.2(a)). These fits are included in Figure 5.5, where it is seen that especially the 900 °C diffusivities are lower than expected. This is more pronounced the larger the Ge content and strain, and is probably an effect of an onset of strain relaxation. In fact, it has been observed that the average diffusivity decreases with increasing annealing time where more relaxation occurs.

Figure 5.5 Arrhenius plot of P diffusivities in compressively strained Si$_{1-x}$Ge$_x$. The full lines are fits to the diffusivities using only the activation energy as a free parameter (see text).

Figure 5.6 Activation energy for P diffusion in relaxed SiGe (□) and compressively strained SiGe (○). The dotted and dashed lines represent the apparent activation energy model (see text).
The activation energies obtained in this way are displayed in Figure 5.6, where the activation energies obtained in relaxed material are included as well. The full line in the figure represents a linear fit to the values for the relaxed material, and the dotted line represents fully strained layers using an apparent activation value of $-13$ eV per unit strain. In an attempt to estimate the uncertainty of the apparent activation value, the dotted lines represents $Q' = -13 \pm 4$ eV per unit strain. The activation energy can generally be expressed as

$$E_a^{\text{SiGe}}(x,s) = E_a^{\text{Si}} + \Delta E_a(x) + Q's,$$  \hspace{1cm} \text{(5.1)}

where, in the case of Si$_{1-x}$Ge$_x$ grown on a Si-substrate, $s = -0.0418x$. The composition effect is given by the fit in Figure 5.2(b), $\Delta E_a(x) = 0.97x$.

Even though boron diffusion in strained material has not been measured in the course of this work, it is illustrative to describe the current status according to literature for a comparison with the phosphorus results. Several results on B diffusion in strained SiGe have been reported, which all found that the boron diffusivity decreased with increasing Ge content [51, 98 - 100]. However, these results were all based on a comparison between boron diffusion in strained SiGe and in Si, and as shown in the previous section, the chemical effect is responsible for part of the decrease. The strain dependence can be determined directly by measuring the diffusion in a Si$_{1-x}$Ge$_x$ layer grown pseudomorphically on a relaxed Si$_{1-y}$Ge$_y$ template, and by varying $y$ the Si$_{1-x}$Ge$_x$ layer can be both compressively and tensile strained. Kou et al. [95] used this technique to measure the strain dependence at 800 °C. Also Zangenberg et al. employed this concept, but with a smaller $s$ range and a temperature interval (800 –

![Figure 5.7 Strain effect of B diffusion at 800 °C](image-url)
900 °C) in order to investigate the temperature dependence as well. The results from Ref. \[95\] and Ref. \[75\] at 800 °C are shown in Figure 5.7. They generally agree that the diffusivity increases with increasing strain, except for Kuo’s measurements of B diffusion in tensile strained Si, where the strain causes a decrease.

The apparent activation energy can be estimated directly from the figure using the relation: \( D(s) = D(0)\exp(-Q' s / k_B T) \). Kuo’s results correspond to an apparent activation value of approximately –6 eV per unit strain for the SiGe results, whereas \( Q' \approx 5 \) eV per unit strain for Si. Zangenberg et al. determine values of \( Q' \) between –3 and –64 eV per unit strain.

In the model by Aziz (equation (2.42)) the apparent activation energy is related to the activation volume, and from this equation we find that

\[
Q' = 15 \left[ -1 - \Delta V_a^\ast / \Omega + \left( V_m^a - V_m^\perp \right) / \Omega \right] \text{ eV/unit strain.} \tag{5.2}
\]

\( \Delta V_a^\ast = -0.13 \Omega \) for B diffusion in Si \[36\] and it may generally vary with the Ge content, but so far there exist no other data. Using the Si value we get the following estimate of the apparent activation energy for B diffusion:

\[
Q'_B = -13 + 15 \left( V_m^a - V_m^\perp \right) / \Omega \text{ eV/unit strain.} \tag{5.3}
\]

In compressively strained material, where the lattice constant in the direction parallel to the surface is smaller than the lattice constant in the direction perpendicular to the surface, the migrational relaxation volume in the latter direction is expected to be smaller than that in the former direction. For tensile strain the opposite is expected. Thus according to equation (5.3) \( Q' \) can be expected to be smaller than –13 eV for compressively strained material and larger than –13 eV for tensile strained material. The activation volume for phosphorus diffusion in Si and SiGe has, to the best of our knowledge, not been measured. However, if we assume \( \Delta V_a^\ast = -0.5 \Omega \) and \( \left( V_m^a - V_m^\perp \right) = -0.3 \Omega \), which seems as reasonable values, we find \( Q'_P \) values consistent with our results.

The strain dependence of antimony diffusion, has been determined by Kringhøj et al. \[101\] who found that compressive strain increased Sb diffusion in Si\(_{0.91}\)Ge\(_{0.09}\), with a \( Q' \) value of 13±3 eV per unit strain, and tensile strain decreased Sb diffusion in Si, \( Q' \approx 17\pm5 \) eV per unit strain. Kuznetsov et al. \[102\] also found that compressive strain enhanced Sb diffusion in Si\(_{0.90}\)Ge\(_{0.10}\) and Si\(_{0.80}\)Ge\(_{0.20}\), and their results were consistent with an apparent activation energy of 24±5 eV per unit strain. For Sb diffusion an activation volume of 0.07\( \Omega \) has been measured \[39\], which according to equation (2.42) will give the following apparent activation energy:
\[ Q'_{\text{Sb}} = 14 + 15\left(\nu_+^{m} - \nu_-^{m}\right)/\Omega \text{ eV/unit strain} \]  

(5.4)

This is consistent with the results of Kringhøj et al. if the anisotropy term is small or vanishing. According to the results of Kuznetsov et al., the anisotropy term is positive, which is plausible for compressive strain, as argued above.

The apparent activation energy may be a good monitor for a quantitative measure of the effect of an applied strain. However, in order to have any physical relevance, \( Q' \) has to be independent on temperature, which implies that the prefactor is strain independent. That the prefactor, indeed, may be regarded as strain independent can be tested on the results of Kringhøj et al. [101] and Zangenberg et al. [75], as they measured the temperature dependence of Sb and B diffusivities in both relaxed and strained layers. Based on the prefactors derived by fitting the diffusivities to an Arrhenius expression, Zangenberg asserted that they are not strain independent. However, a closer examination of their results shows that satisfactory fits to the Arrhenius expression can be obtained under the assumption of strain independent prefactors. Thus, their results do not contradict the assumption of strain independence.

In summary, B and P diffusion respond to compressive strain with a reduction, as manifested in the negative apparent activation energy, whereas Sb diffusion is increased by compressive strain.

5.3 Diffusion mechanism

It has already been mentioned in Chapter 2 that P and B diffuse in Si entirely via an interstitialcy mechanism. Likewise, it has been determined that Sb diffuses in Si via a vacancy mechanism [9]. There may, however, be a different situation in Si\textsubscript{1-x}Ge\textsubscript{x}.

There are no established techniques to inject point defects into SiGe, like oxidation and nitridation of Si surfaces inject point defects into Si. It has been shown, though, that the point defects generated by oxidation of a Si cap layer in a heterostructure will pass through a SiGe layer [58] and thus, it is possible to measure dopant diffusion in strained SiGe layers with a supersaturation of Si self-interstitials. Kuo et al. [58] found that the enhancement of boron diffusion, caused by oxidation of a Si-cap, in Si\textsubscript{0.90}Ge\textsubscript{0.10} and Si\textsubscript{0.82}Ge\textsubscript{0.18} is the same as that in Si. Similarly, Fang et al. [47] used both nitridation and oxidation, and found that the interstitialcy fraction of B diffusion in Si\textsubscript{0.80}Ge\textsubscript{0.20} is the same as the one for diffusion in Si.

In [Paper III], P diffusion in SiGe during injection of Si self-interstitials is discussed, and it is found that the P diffusion is generally increased but that the enhancement factor decreases with the Ge content. The decrease is attributed to a reduction of the
excess interstitials due to trapping by misfit dislocations at the Si/SiGe interfaces, rather than to a gradual change in mechanism.

Antimony diffusion in strained Si$_{1-x}$Ge$_x$ ($x = 0.1, 0.2$) during oxidation, studied by Kuznetsov et al. [103], was found to be retarded by factors of 0.2 – 0.4. The samples were annealed at different temperatures, and the results do show that the vacancy mechanism is dominating.

So, boron, phosphorus and antimony diffuse by the same kind of mechanism in Si$_{1-x}$Ge$_x$ ($x < 0.2$) as in Si, consistent with the signs of the apparent activation energies as discussed in the previous section.

### 5.4 P segregation in heterostructure interfaces

The segregation phenomena, as described in Section 2.7.2, was studied for phosphorus employing the Si/SiGe/Si heterostructures used to determine the P diffusion in strained SiGe. The samples were annealed for 5 and 10 hours at 900 °C, and the segregation coefficient was determined as the ratio of the P concentration at the Si side of the interface to the P concentration at the SiGe side. An example of a P profile in which the segregation is clearly seen is shown in Figure 5.8, where the P concentration is shown on a linear scale. The profile has been simulated by adding an additional P flux term at the interface to the Fick diffusion model, as described in section 2.7.2, and this model is clearly able to account for the main features of the segregation.

The segregation coefficients obtained for the three Si/Si$_{1-x}$Ge$_x$ samples studied ($x =$ 0.08, 0.13, and 0.18) are shown in Figure 5.9, together with the literature data from Hu et al. [46] and Kobayashi et al. [104]. The segregation coefficient is expected to be
temperature dependent, so since the values in Figure 5.9 have not been determined for the same annealing temperature, they can not be directly compared. Nevertheless, there is a clear increase with increasing Ge content, which, since the segregation coefficient is equal to the ratio between the solid solubilities of P in Si and SiGe, implies that the solubility of P in SiGe decreases with increasing Ge content.

5.5 P diffusion in SiGe with imperfect crystallinity

P diffusion in relaxed SiGe, as described in Section 5.1, was studied in samples which had received an annealing (7 min at 800°C) just after growing the graded SiGe buffer layer. P diffusion has also been studied in samples grown without this annealing step, and with a thickness of the SiGe layer of 1 µm (as opposed to 3 µm in the other samples).

Figure 5.10 shows the P profiles in these samples before and after annealing at 810 °C for 1 hour. Applying the Fick model with the P diffusivity values given in Figure 5.1 results in a profile represented by the thin dashed line in Figure 5.10, practically identical with the as-grown one and does not at all account for the experimental profile after annealing. In fact, the Fick diffusion model can not be used to describe the diffused profile. Instead, the intermittent diffusion model is found to give satisfactory fits to the measured profiles in the whole temperature interval used (700 – 900 °C). The parameters extracted from the intermittent model, \( g \) and \( \lambda \), are shown by the two Arrhenius plots in Figure 5.11.
The \( g \) values are approximately independent on the Ge content, and they are comparable to or smaller than those for boron diffusion in Si obtained by Cowern et al. [9]. The \( \lambda \) values, on the other hand, are very large compared to those measured for B diffusion in Si (\( \lambda \approx 10 \) nm at 625 °C during proton irradiation (Paper IV), and \( \lambda < 5 \) nm at 900 °C during inert and oxidizing anneal [9]). The effective diffusivity given by equation (2.16) is much larger than the diffusivity obtained in section 5.1, as shown in Figure 5.12. This suggests that P diffuses by another mechanism in these samples, and some hints as to the nature of this mechanism may be obtained from the values of \( g \) and \( \lambda \). The parameter \( g \) is a measure of the probability that a dopant atom will perform
a diffusion step, in which the atom travels an average distance of $\lambda$. Since $g$ is smaller than (or comparable to) that found for interstitial B diffusion in Si, it suggests that a pure interstitial mechanism is not the main reason for the large diffusion coefficient in these samples. The very large values of $\lambda$ (about 1/10 of the thickness of the relaxed SiGe layer), on the other hand, indicate a mechanism with is a low migration barrier for the diffusing species. However, the temperature dependence of $\lambda$, suggests that the migration barrier is larger than the barrier for the migrating species to “retransfer” into substitutional immobile phosphorus.

A tentative interpretation of the results is that threading dislocations act as channels for the rapid diffusion. As mentioned in section 3.1, the graded SiGe layer contains threading dislocations, which may extend into the relaxed Si$_{1-x}$Ge$_x$ layer where they can affect the dopant diffusion. In this interpretation, $g$ is the probability for a P atom to encounter a dislocation and, thus $g$ will depend on the density of dislocations. This density is expected to be larger in the 20% Ge sample but, on the other hand, the probability for a P atom to encounter such a dislocation is probably limited by the self-interstitials providing the mobility of P. The self-interstitials are also likely to be trapped by dislocations, so materials with a high density of dislocations will generally have a lower concentration of self-interstitials. Hence, there are competing effects which may account for the similar $g$ values in the $x = 0.1$ and $x = 0.2$ samples. That dislocations may indeed act as channels for rapid dopant diffusion have been observed previously in Ref. [105], where Si atoms were found to diffuse very fast along dislocations in GaAs layers grown epitaxially on a Si substrate.

In conclusion, the P diffusion in relaxed SiGe does obviously depend strongly on the growth conditions of the samples. The mechanism for the rapid diffusivity in SiGe layers with imperfect crystallinity is not fully established, but fast diffusion along threading dislocations appears to be a viable candidate.
Chapter 6

Summary of papers

Paper I
Phosphorus and boron diffusion in Si were studied with SIMS. It was found that P diffuse with an activation energy of 2.74 eV in the temperature range 810 – 1100 °C. This value is lower than previously reported ones. Vacuum diffusion known to retard B diffusion, was found also to retard P diffusion. The author performed sample annealing, SIMS measurements, analysis of the data, and wrote the manuscript.

Paper II
Equilibrium P and B diffusion were studied and the activation energies were found to be smaller than the wealth of the literature values. An initial slow transient was observed for B diffusion. The results were compared with theoretical results on the microscopic diffusion processes. The author performed sample annealing, SIMS measurements, analysis of the data, and wrote the manuscript.

Paper III
Phosphorus diffusion in relaxed strained SiGe was measured. P diffusion is enhanced in SiGe compared to Si. Compressive strain retards the P diffusivity. Results on P segregation at Si/SiGe interfaces as well as the effect of strain relaxation. The author performed sample annealing, SIMS and HRXRD measurements, analysis of the data, and wrote the manuscript.

Paper IV
RED of boron at elevated temperature was studied. B was shown to diffuse according to the intermittent diffusion model with an effective diffusivity enhanced by up to 10^7. B diffusion exhibited a concentration dependence, which was correlated with the intrinsic carrier concentration. The author performed SIMS measurements and participated in analysis of the data.

Paper V
RED of boron at elevated temperatures was studied. B diffusion was enhanced by a factor of 10 to 10^7 compared to equilibrium diffusion. Recombination of I and V at the
surface was believed to account for a small depth dependence of $D_B$ that was observed. Presence of Sb spikes increased the diffusion of B.

The author performed SIMS measurements and participated in analysis of the data

**Paper VI**

B and Sb diffusion in separate $\delta$-spikes during proton irradiation at elevated temperatures were studied. It was found that Sb only diffused when B spikes are present in the sample. Sb diffusion even depends on the maximum concentration of B in the B spike.

The author performed the SIMS measurements and participated in analysis of the data
Appendix A
Numerical simulation of diffusion

A.1 Second order differential equation

Fick’s diffusion equation is a second order partial differential equation and can not generally be solved analytically. The equation

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2} \quad (A.1)$$

is solved numerically by replacing the partial derivatives with their finite-difference approximations using the Crank-Nicholson method [106], where $u_j^0$ (corresponding to the dopant concentration at point $j$ and at time 0 ($j = 1, 2, \ldots, J$) is known. $u_j^{n+1}$ is given by

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{D}{2} \left( \frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{(\Delta x)^2} + \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2} \right), \quad (A.2)$$

where the $u_j^{n+1}$ on the right hand side is given by the set of linear equations:

$$-\alpha u_j^{n+1} + (1 + 2\alpha)u_j^n - \alpha u_{j+1}^{n+1} = u_j^n, \quad j = 2, 3 \ldots J - 1, \quad \text{where } \alpha = \frac{D \Delta t}{(\Delta x)^2}. \quad (A.3)$$

A.1.1 Temperature ramp

If the temperature varies through an annealing it has to be considered in the simulation. This can be done by realizing that the diffusion coefficient at any temperature relative to a diffusivity $D_0$ at a given temperature $T_0$ is

$$D(T) = D_0 \exp \left( -\frac{E_a}{k_B \left( \frac{1}{T} - \frac{1}{T_0} \right)} \right), \quad (A.4)$$

The drawback is that the activation energy is normally not known, so a few iterations may be necessary.
An equivalent way to include the temperature ramp is in an effective diffusion time. A
time for which the diffusion for constant $D$ is equivalent to the actual diffusion with a
temperature ramp up, a constant temperature and the temperature down ramp. To
calculate the effective diffusion time, one needs to know the temperature versus time
profile, the $t_{\text{eff}}$ is given by

$$t_{\text{eff}} = \int_{t_{\text{start}}}^{t_{\text{end}}} \exp \left( -\frac{E_a}{k_B} \left( \frac{1}{T(t)} - \frac{1}{T_0} \right) \right) dt.$$  \hfill (A.5)

A.1.2 Segregation

In a heterostructure there might be some segregation at the interfaces. Segregation can
be included by adding an extra term to the flux at the interfaces for each time step in
the numerical solution of Fick’s equation. The semi-empirical model of segregation at
heterostructure interfaces between material 1 and material 2 uses a flux term, which is
given by

$$J = v_{12} \left( \frac{C_1}{k_{12}} - C_2 \right),$$ \hfill (A.6)

where $v_{12}$ is a transport velocity, $k_{12}$ is the segregation coefficient, and $C_1$, $C_2$ are the
dopant concentrations in material 1 and 2, respectively. Comparing with equation \[2\]. \texttt{[107]} it is seen that the transport velocity depends on the temperature, the gradient of
the Ge profile (in the case of a Si/SiGe heterostructure), $\Delta x$, and $\Delta t$. The segregation
coefficient, on the other hand, is the ratio between the solubility in the two materials,
and, thus, for a given sample, depends only on annealing temperature.

A.2 Intermittent diffusion

In the model for intermittent diffusion the solution for a delta spike is given if $\lambda$ and $gt$
are known \[107]. This solution can be applied to any real dopant profile, by
convoluting the initial profile with the solution to the delta profile. The solution for the
delta spike distribution is, with the depth scale normalized to $\xi = x/\lambda$ and defining $\theta = gt$:

$$s(\xi, \theta) = \sum_{n=0}^{\infty} P_n(\theta) f_n(\xi),\ \text{where}$$

$$P_n(\theta) = \frac{\theta^n}{n!} \exp(-\theta) \text{ and}$$

$$f_{n=0}(\xi) = \delta(\xi),\ f_{n>0}(\xi) = \frac{\exp\left(-|\xi|\right)}{2^n(n-1)!} C_n(|\xi|),$$ \hfill (A.9)
The amount of terms necessary in the summation in equation (A.7) depends on the value of $\theta$. For $\theta < 5$, 10 terms are sufficient.

With the initial profile given by a vector $\text{INI}$, the result is given by a vector $\text{RES}$ as

$$RES^i(\lambda, \theta) = \sum_n \frac{\sum_j f_n^{i,j} \cdot INI^j}{\sum_j f_n^{i,j} P_n(\theta)}$$

where $f_n^{i,j} = f_n(\xi_i - \xi_j)$.

(A.11)
Appendix B
Density functional theory

Since the main part of the theoretical results on dopant diffusion in Si is based on density functional theory (DFT), a brief introduction to DFT, based on [108, 109], may be appropriate.

In DFT the ground-state total energy $E$ for a collection of electrons which interact with each other and an external potential is calculated. $E$ is expressed as a functional of the electron density $n(r)$, and $E(n(r))$ will be at a minimum when $n(r)$ represents the true electron density. DFT for solid materials usually only considers the energy of the valence electrons. The energy of the ion and core electrons are considered constant, and they constitute the external potential $V_{\text{ion}}(r)$ in which the valence electrons move.

This means that the single electron energy and wave functions can be calculated by solving the single-particle Schrödinger-like equation:

$$\left[-\frac{1}{2}\nabla^2 + V^{\text{eff}}(r)\right]\psi_i = \varepsilon_i\psi_i,$$

where the first term is the kinetic energy operator and $V^{\text{eff}}(r)$ is an effective potential given by the sum of the potential due to the ions, the electrostatic interaction between electrons and a so-called exchange-correlation term $V_{\text{XC}}$:

$$V^{\text{eff}}(r) = V_{\text{ion}}(r) + e^2 \int \frac{n(r')}{|r-r'|} \, dr' + V_{\text{XC}}.$$  \hspace{1cm} (B.2)

The exchange energy term comes from the Pauli principle, which requires that the $N$-electron wave function is antisymmetric, the correlation energy arises from the approximate representation of the electron-electron repulsion in the Hartree–Fock method [248]. There exists no analytical expression for $V_{\text{XC}}$, and therefore no exact solution of $n(r)$. The $V_{\text{XC}}$ term in equation (B.2) is approximated, and the two most common approximations, when concerned with semiconductor calculations, are the local density approximation (LDA) and the generalized gradient approximation (GGA). For diffusion calculations in silicon, LDA provides results which are 0.2 – 0.5 eV smaller than results obtained using GGA.
From equation (B.1) the electron wave functions are calculated and then $n(r)$ is given by

$$n(r) = \sum_{i=1}^{N} |\psi_i(r)|^2.$$  \hspace{1cm} (B.3)

However, the one-electron wave functions depend on the electron density through $V_{\text{eff}}$ in equation (B.2), so the system is usually solved iteratively, by first choosing a convenient $n(r)$, find $\psi_i$ by solving equation (B.1), calculate a new value of $n(r)$ and start all over again, until the electron density has converged. When the electron density has converged the energy of the system $E$ can be calculated.

In the case of silicon, DFT can predict lattice parameters and elastic parameters with a very good accuracy while the band gap is underestimated by more than 50%. The calculations assume a crystal temperature of 0 K, which means that no information on entropy associated with the processes can be obtained. Also, DFT does not consider the dynamics of the defects considered.
References

3. Fahey P. M., Griffin P. B., Plummer J. D. Point Defects and Dopant Diffusion in Silicon Rev. Mod. Phys. 61, 289 (1989)
13. Thurmond C. D. The standard thermodynamic functions for the formation of electrons and holes in Ge, Si, GaAs, and GaP J. Electrochem. Soc. 122 (8), 1133 (1975)
16. Tang M., Colombo L., Zhu J., de la Rubia D. Intrinsic Point Defects in Crystalline Silicon: Tight Binding Molecular Dynamics Studies of Self-Diffusion,
29. Fang W. T. C., Fang T. T., Griffin P. B., Plummer J. D. Surface and Bulk Point Defect Generation in Czochralski and Float Zone Type Silicon Wafers Appl. Phys. Lett. 68 (15), 2085 (1996)
41. People R. *Indirect band gap of coherently strained Ge<sub>x</sub>Si<sub>1-x</sub> bulk alloys on <001> silicon substrates* Phys. Rev. B 32 (2), 1405 (1985)
47. Fang T. T., Fang W. T. C., Griffin P. B., Plummer J. D. *Calculation of the Fractional Interstitial Component of Boron Diffusion and Segregation Coefficient of Boron in Si<sub>0.8</sub>Ge<sub>0.2</sub>* Appl. Phys. Lett. 68 (6) 791 (1996)
60. Fewster P. F. a High-Resolution Multiple-Crystal Multiple-Reflection Diffractometer J. Appl. Cryst. 22, 64 (1989)
64. Zalm P. C. Secondary ion mass spectrometry Vacuum 45 (6/7), 753 (1994)


72. Pelleg J., Ditchek B. M. *Diffusion of P in a novel three-dimensional device based on Si-TaSi2 eutectic* J. Appl. Phys. 73 (2), 699 (1993)


89. Zangenberg N. *Defect and diffusion studies in Si and SiGe* Ph. D. thesis


