

HREM-Bildanalyse von III-V-Halbleiter-Schichtstrukturen durch quantitativen Bildvergleich Experiment – Simulation

Dissertation

zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der

Mathematisch-Naturwissenschaftlich-Technischen Fakultät
(Mathematisch-Naturwissenschaftlicher Bereich)
der Martin-Luther-Universität Halle-Wittenberg

von

Dipl.-Phys. Holger Kirschner
geb. am 04.01.1969 in Freising

Gutachter:

1. Prof. Dr. U. Gösele (Max-Planck-Institut für Mikrostrukturphysik)
2. Prof. Dr. W. Hergert (Martin-Luther-Universität Halle-Wittenberg)
3. Prof. Dr. W. Jäger (Christian-Albrechts-Universität Kiel)

Halle a.d. Saale, den 19. Januar 2000

Inhaltsverzeichnis

Verwendete Abkürzungen	5
Einleitung	6
1 Kompositionsbestimmung in HREM-Aufnahmen	9
1.1 Beschreibung der Problemstellung	9
1.2 Übersicht über vorhandene Verfahren	10
1.2.1 Verfahren die Gitterverspannung auswerten	10
1.2.2 Kontrastorientierte Verfahren	11
2 Simulation von HREM-Aufnahmen	12
2.1 Wechselwirkung Elektronenstrahl – Probe	12
2.2 Abbildung der Austrittswellenfunktion	14
2.3 Das Programmpaket EMS	16
3 Amorphe Oberflächenschichten	17
3.1 Hintergrund	17
3.2 Simulation amorpher Oberflächenschichten	18
3.2.1 Übersicht über Modelle	18
3.2.2 Simulation eines Objekts mit zufälligen Atompositionen	21

<i>INHALTSVERZEICHNIS</i>	2
3.3 Bestimmung experimenteller Amorphisierungstiefen	25
3.3.1 Beschreibung der Dickenbestimmung	26
3.3.2 Diskussion der Dickenbestimmung	27
3.4 Auswirkungen der Amorphisierung	31
4 Bildvorverarbeitung	36
4.1 Warum Bildvorverarbeitung ?	36
4.2 Zelldetektion	36
4.3 Realraumbasierte Vorverarbeitung: get3	38
4.3.1 Beschreibung	38
4.3.2 Motivation für get3	39
4.3.3 Diskussion von get3	45
4.4 Vorverarbeitung: lokale Fouriertransformierte	48
4.4.1 Beschreibung	48
4.4.2 Diskussion der lokalen Fouriertransformierten	49
5 Mittlere experimentelle Parameter	51
5.1 Szenario: Kompositionsbestimmung durch HREM-Bildauswertung . . .	51
5.2 Vergleich von Experiment und Simulation	52
5.2.1 Lineare Anpassung bei get3	53
5.2.2 Lineare Anpassung bei lokaler Fouriertransformierter	55
5.3 Präzisieren von experimentellen Parametern	56
5.3.1 Beschreibung	56
5.3.2 Diskussion der Parametersuche	59
5.4 Einfluss von amorphen Deckschichten	60
5.5 Kompositions- und Dickenbestimmung ohne Rauschen	63

<i>INHALTSVERZEICHNIS</i>	3
6 Neuronale Netze	64
6.1 Verwendeter Netztyp	64
6.2 Lernverfahren	65
6.3 Vor- und Nachverarbeitung	65
6.3.1 Reskalierung von Eingabe und Ausgabe	65
6.3.2 Vorabwissen über Kompositionsgrenzen	66
6.4 Architekturanpassung	66
6.5 Experimentelle Muster	67
7 Unterdrückung von Rauschen	70
7.1 Beschreibung des Testszenarios	70
7.2 Vorgehensweise mit neuronalem Netz	71
7.3 Minimalabstand zu simulierten Mustern	71
7.4 Minimalabstand mit Kontrast- und Helligkeitsanpassung	72
7.5 Projektion orthogonal zur Hauptkomponente des Rauschens	73
7.6 Optimale Projektionsebene	74
7.7 Vergleich der beschriebenen Verfahren	75
8 Experimentelle Anwendung	78
8.1 AlGaAs Grenzschrift	78
8.1.1 Mittlere Template Parameter	78
8.1.2 Kompositions- und Dickenbestimmung	82
8.1.3 Diskussion der Kompositionsverteilung	83
8.1.4 Diskussion des Dickenverlaufs	87
8.2 AlGaAs-Bragg-Reflektor	88

<i>INHALTSVERZEICHNIS</i>	4
9 Zusammenfassung	94
A Atomare Streuamplituden nach Doyle und Turner	97
B Symmetrische Gitterbesetzung bei ternären Halbleitern	99
C Fouriertransformierte des atomaren Potentials	103
Literaturverzeichnis	106

Verzeichnis der verwendeten Abkürzungen

HREM	High-Resolution transmission Electron Microscopy Hochauflösende Transmissionselektronenmikroskopie
RDO	Random numberDensity Object amorphes Objekt mit gleichverteilten Atompositionen
HSM	modifiziertes Hardsphere Modell
RPO	Random Phase Object amorphes Objekt mit zufälliger Phasenvariation
TEM	TransmissionsElektronenMikroskopie
EMS	Programmpaket zur Simulation hochaufgelöster TEM-Aufnahmen von P. Stadelmann [42],[43]
fcc	face-centered cubic Kubisch flächenzentriert
Ez.	Dickenangabe in Elementarzellen (AlGaAs: 5.7Å)

Einleitung

Bei modernen elektronischen Bauelementen erlangt die Mikrostrukturierung von Halbleitern zunehmende Bedeutung. Grund hierfür ist einerseits die steigende Integrationsdichte der Bauelemente, andererseits lassen sich bei kleinen Abmessungen Quanteneffekte gewinnbringend einsetzen.

Die chemische Zusammensetzung von Halbleitern beeinflusst deren grundlegende physikalische Eigenschaften. Daher ist eine Mikrostrukturierung der Halbleiterkomposition und deren Überprüfung durch orts aufgelöste Kompositionsbestimmung von vorrangiger Wichtigkeit. Auch bei gewöhnlichen (vordergründig makroskopisch strukturierten) Halbleiterbauelementen erzeugt die Strukturierung der Komposition innere Grenzflächen deren Mikrostruktur sich empfindlich auf Bauteileigenschaften auswirkt. So sind beispielsweise die elektrischen Eigenschaften von Tunnelioden eng verbunden mit der Grenzflächenmikrostruktur [6] [14]. Beim Einsatz von AlGaAs Quantentrog-Strukturen ist eine genaue Charakterisierung der Grenzflächen notwendig. Zum Beispiel wird die Form von Photolumineszenzlinien beeinflusst von der Rauigkeit der Grenzflächen [51] [4] [2]. Eine Optimierung von Prozessparametern erfordert somit eine quantitative Beschreibung der jeweiligen Grenzflächen und macht orts aufgelöste Kompositionsbestimmung notwendig.

Speziell für die III-V-Halbleiter Kompositionsbestimmung gibt es daher eine Reihe von Methoden. Die höchste laterale Auflösung wird zur Zeit mit der quantitativen Auswertung von hoch aufgelösten transmissionselektronenmikroskopischen (HREM-) Aufnahmen erreicht. Hierfür gibt es eine Reihe von Ansätzen (siehe Abschnitt 1.2).

Ziel der Arbeit war die Optimierung der orts aufgelösten Kompositionsbestimmung für praxisrelevante experimentelle Situationen.

Wichtig dabei ist der Verzicht des Verfahrens auf eine spezielle Abbildungssituation. Auf diese Weise kann die Aufnahme in herkömmlicher Weise kontrastoptimiert erfolgen und die Durchführung der quantitativen Auswertung geht nicht auf Kosten der Qualität der HREM-Aufnahme. Beim Vorabwissen über experimentelle Parameter müssen

(bei vielen Verfahren unterschätzte) Fehler berücksichtigt werden. Daher ist eine Präzisierung von experimentellen Parametern unerlässlich. Störungen der Aufnahme durch Schwankungen der Probendicke sollten möglichst wenig Einfluss auf die Kompositionsbestimmung haben. Schliesslich müssen nicht zu vermeidende Abweichungen der HREM-Proben vom ideal kristallinen Zustand quantifiziert und deren Fehlereinfluss auf die Kompositionsbestimmung minimiert werden.

Unter Berücksichtigung aller dieser Forderungen soll ein neues Verfahren entwickelt werden, um die chemische Zusammensetzung unverspannter Halbleiterproben in (001)-Projektion zu bestimmen. Zudem soll das Verfahren in einfacher Weise auf die Auswertung verspannter Strukturen erweiterbar sein.

Gliederung der Arbeit

Die vorliegende Arbeit ist wie folgt gegliedert:

Das anschliessende Kapitel beschreibt bereits bekannte Verfahren, um durch quantitative HREM-Bildauswertung orts aufgelöste Kompositionsbestimmung durchzuführen.

Kapitel 2 skizziert die Grundlagen der HREM-Abbildungssimulation.

Kapitel 3 beschreibt die Oberflächenamorphisierung der Probe durch den Abdünnprozess bei der Probenpräparation. Es wird eine Methode beschrieben um amorphe Schichten numerisch effektiv zu simulieren. Die Amorphisierungstiefe von HREM-Proben wird mit einer dafür entwickelten Diffraktogrammanalyse bestimmt. Schliesslich wird der Einfluss der Amorphisierung auf den HREM-Bildkontrast für praktisch aufgetretene Amorphisierungstiefen untersucht.

Kapitel 4 erklärt die Verwendung von Vorabinformation über die HREM-Bildentstehung im Rahmen von Bildvorverarbeitung. Mit Symmetrieargumenten werden die zahlreichen Möglichkeiten der Bildvorverarbeitung auf zwei Verfahren eingeschränkt:

Die lokale Fouriertransformierte und ein realraumbasiertes Verfahren (get3).

Kapitel 5 betrachtet über Probenbereiche gemittelte experimentelle Parameter. Es wird ein Verfahren entwickelt um durch Vergleich Simulation – Experiment lateral gemittelte experimentelle Parameter zu finden bzw. bekannte Parameter zu präzisieren. Der Einfluss von amorphen Oberflächenschichten auf die Parameterbestimmung wird untersucht und Verfahren zur Bildvorverarbeitung an dieser Problemstellung getestet.

Kapitel 6 skizziert die Grundlagen der Anwendung neuronaler *feed-forward* Netzwerke, die zum weitem Verständnis der Arbeit notwendig sind.

Die Amorphisierung der Probe erzeugt in der HREM-Aufnahme korreliertes Rauschen. Kapitel 7 untersucht Verfahren um diese Störung bei der orts aufgelösten Kompositions- und Probendickenbestimmung zu unterdrücken. Neben einer Reihe von klassischen Verfahren wird auch eine Methode unter Verwendung neuronaler Netze an einem praxisorientierten Szenario getestet.

Die entwickelten Verfahren werden in Kapitel 8 zur quantitativen Analyse von HREM-Aufnahmen angewendet.

Die Arbeit schließt mit einer Zusammenfassung der Ergebnisse und einem Ausblick auf zukünftige weitere Verbesserung der lokalen Kompositions- und Probendickenbestimmung.

Kapitel 1

Kompositionsbestimmung in HREM-Aufnahmen

In diesem Kapitel wird die in der vorliegenden Arbeit untersuchte Problemstellung der Kompositionsbestimmung in HREM-Abbildungen von Halbleiterheterostrukturen vorgestellt und eine Übersicht über die aus der Literatur bekannten Verfahren gegeben.

1.1 Beschreibung der Problemstellung

Die Problemstellung der vorliegenden Arbeit lässt sich wie folgt kurz zusammenfassen:

In hochauflösenden transmissionselektronenmikroskopischen Aufnahmen soll die Komposition des untersuchten Halbleitermaterials mit hoher Ortsauflösung bestimmt werden. Die angestrebte laterale Auflösung entspricht dabei der Größe der Elementarzelle. Beispiele für praxisrelevante Materialsysteme, die mit den hier vorgestellten Methoden untersucht werden können, sind (jeweils $x \in [0, 1]$):



Die untersuchten Halbleitermaterialien haben ausnahmslos Sphaleritstruktur. Das Kristallgitter der Sphaleritstruktur (Abb. 1.1) besteht aus zwei fcc-Untergittern die um ein Viertel der fcc-Raumdiagonalen verschoben sind [28].

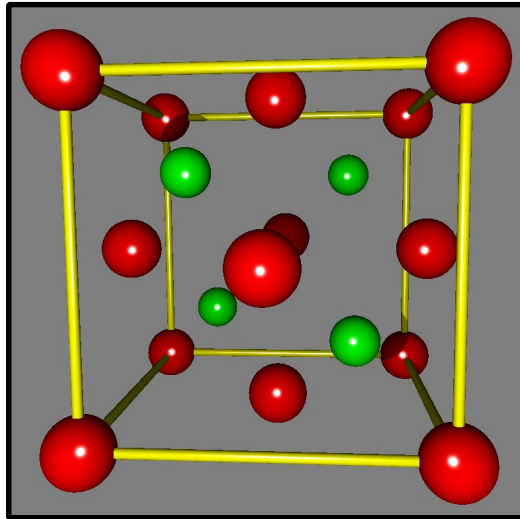


Abbildung 1.1: Elementarzelle der Sphaleritstruktur in (001)-Projektion: Die roten und grünen Kugeln repräsentieren jeweils ein fcc-Untergitter der Sphaleritstruktur.

Eine Kompositionsvariation (Veränderung von x in den Summenformeln 1.1-1.3) ändert bei den untersuchten Materialsystemen nur die Besetzung der Positionen eines Untergitters. Das jeweils variierende Untergitter wird dabei von Atomen zweier Elemente statistisch besetzt.

1.2 Übersicht über vorhandene Verfahren

Verfahren zur Kompositionsbestimmung in hochauflösenden TEM Aufnahmen kann man in zwei Kategorien einteilen:

Verfahren die Gitterverspannung auswerten und kontrastorientierte Verfahren.

1.2.1 Verfahren die Gitterverspannung auswerten

Eine Reihe von Verfahren (sog. *strain*-orientierte Verfahren) nutzen zur Kompositionsbestimmung die Tatsache, dass die Komposition von ternären Halbleitermaterialien deren Gitterkonstante beeinflusst (Vegardsches Gesetz [29]). Ein Beispiel hierfür ist das Computerprogramm DALI [39]. Probleme bereiten diesen Verfahren wenig verspannte Materialsysteme, so zum Beispiel AlGaAs. Die Probendicke beeinflusst die Oberflächenrelaxation der Gitterverspannungen und damit auch die Abstandsverhältnisse im Bild. Um eine Kompositionsbestimmung mit *strain*-orientierten Verfahren durchführen zu können, muss die absolute Dicke der Probe orts aufgelöst bekannt sein [40].

1.2.2 Kontrastorientierte Verfahren

Kontrastorientierte Verfahren nutzen im Gegensatz zu *strain*-orientierten Verfahren die unterschiedlichen Streueigenschaften von verschiedenen Elementen zur Kompositionsbestimmung. Dabei lassen sich wieder zwei Unterscheidungen treffen: fourierbasierte und realraumbasierte Verfahren.

Fourierbasierte Verfahren

Fourierbasierte Verfahren nutzen die chemische Sensitivität der 200-Fourierkoeffizienten des Bildes (vgl. Kap. 4.4.1), um die Komposition zu bestimmen. Beispiele hierfür sind Fourierfilterung [41] und lokale Fouriertransformierte [49] [50]. Basiert die Kompositionsbestimmung ausschließlich auf der Kompositionsabhängigkeit der 200-Reflexintensität, so führt dies in manchen Materialsystemen zu hohen Fehlern (siehe Anhang B).

In Materialsystemen mit Diamantstruktur (zum Beispiel $\text{Si}_x\text{Ge}_{1-x}$) lässt sich die Bildentstehung näherungsweise mit dem 220-Fünfstrahlfall beschreiben [45]. Trägt man lineare und nichtlineare Amplitude gegeneinander auf, so kann analytisch gezeigt werden [46], dass eine Variation der Komposition abhängig von Probendicke und Defokussierung verschiedene Ellipsen erzeugt. Fittet man experimentelle Daten von Dicken- und Kompositionsvariationen mit simulierten Ellipsen, so lassen sich mittlerer Defokus, lokale Probendicke und lokale Komposition bestimmen ([46]).

Das in dieser Arbeit vorgestellte realraumbasierte Verfahren wird in Abschnitt 5.4 mit der lokalen Fouriertransformierten (fourierbasiertes Verfahren) verglichen. Es stellt sich dabei heraus, dass die lokale Fouriertransformierte stärker durch amorphe Oberflächenschichten der Probe gestört wird als das realraumbasierte Verfahren.

Realraumbasierte Verfahren

Bei allen bisher bekannten realraumbasierten Verfahren wird die Bildinformation als Vektor betrachtet, dessen Komponenten die Intensitäten (Pixel) des untersuchten Bildausschnitts sind [35] [27] [23].

Diese Vorgehensweise verschlüsselt die zweidimensionalen Abstandsverhältnisse im Bild, so dass diese Information für die weitere Verarbeitung nicht mehr zur Verfügung steht. Dadurch sind diese Verfahren empfindlich gegenüber Verschiebung der Bildinformation um wenige Pixel (siehe Abschnitt 4.3.3). Eine optimale Fehlerunterdrückung des im Bild enthaltenen korrelierten Rauschens wird nur unter Berücksichtigung der zweidimensionalen Topologie des Bildes möglich (siehe Kapitel 7).

Kapitel 2

Simulation von HREM-Aufnahmen

In diesem Kapitel werden die theoretischen Grundlagen der Simulation von hochauflösenden transmissionselektronenmikroskopischen (HREM) Aufnahmen skizziert. Eine weitaus umfassendere Übersicht gibt z.B. [22].

Die Simulation von HREM-Aufnahmen unterteilt sich in zwei Schritte, die Wechselwirkung des Elektronenstrahls mit der zu untersuchenden Probe und die Abbildung der aus der Probe austretenden Elektronenwelle als Intensitätsbild in der Bildebene.

2.1 Wechselwirkung Elektronenstrahl – Probe

Die Wechselwirkung der Elektronenwelle mit der Probe wird durch die Theorie der dynamischen Elektronenbeugung beschrieben. In der Praxis wird zur Simulation dieses Prozesses meistens entweder das Bethe-Bloch-Verfahren [1] oder das *multislice*-Verfahren verwendet [8] [18].

Ausgangspunkt beim Bethe-Bloch-Verfahren ist die relativistisch korrigierte Schrödingergleichung für die Wellenfunktion des Elektrons im periodischen Kristallpotential der Probe. Diese Gleichung kann als Eigenwertproblem unter Ausnutzung der Kristallperiodizität mit einem Blochwellenansatz gelöst werden. Die Verwendung der Periodizität beschränkt die Anwendung des Verfahrens auf perfekt kristalline Proben. In dieser Arbeit werden teilamorphisierte HREM-Proben untersucht (siehe Kapitel 3). Für die Abbildung dieser aperiodischen Schichten ist das Bethe-Bloch-Verfahren ungeeignet.

Die in dieser Arbeit simulierten HREM-Abbildungen wurden daher mit dem *multislice*-Verfahren berechnet. Dabei wird die Probe (Dicke t) in dünne Teilschichten (*slices*) der Dicke t_s unterteilt. Innerhalb einer solchen Teilschicht wird die Wechselwirkung mit

dem Elektronenstrahl aufgeteilt in eine Phasenverschiebung der Elektronenwelle und eine anschließende ungestörte Propagation der Welle um die Dicke der Teilschicht.

Die Wellenfunktion des Elektrons wird beim Durchlaufen einer Teilschicht der Dicke t_s phasenverschoben:

$$\psi(x, y, t_s) = \psi_0(x, y) \exp\left(i\sigma \int_0^{t_s} V(x, y, z) dz\right) \quad (2.1)$$

Hierbei ist $V(x, y, z)$ das Potential in der Probe und σ die Wechselwirkungskonstante:

$$\sigma = \frac{2\pi m e \lambda}{h^2} \quad (2.2)$$

mit den Größen

$$\lambda = \frac{hc}{\sqrt{2EE_0 + E^2}} \quad \text{relativistisch korrigierte Wellenlänge} \quad (2.3)$$

$$U_b : \text{Beschleunigungsspannung} \quad (2.4)$$

$$E = eU_b \quad (2.5)$$

$$E_0 = m_0 c^2 \quad (2.6)$$

$$m = m_0 + \frac{E}{c^2} \quad \text{rel. korrigierte Masse des Elektrons} \quad (2.7)$$

$$m_0 : \text{Ruhemasse des Elektrons} \quad (2.8)$$

Der Phasenfaktor in (2.1) wird als Phasengitterfunktion $\text{pg}(x, y)$ bezeichnet.

$$\text{pg}(x, y) = \exp\left(i\sigma \int_0^{t_s} V(x, y, z) dz\right) \quad (2.9)$$

Abschnitt 3.2.2 beschreibt eine teilanalytische Berechnung der Phasengitterfunktion für amorphe Proben.

Nach der Phasenverschiebung durch das Potential der Teilschicht wird die Wellenfunktion des Elektrons um die Dicke der Teilschicht t_s propagiert (freie Ausbreitung). Dies entspricht einer Faltung der Wellenfunktion mit der Propagatorfunktion P (s. [48]):

$$P(x, y) = \frac{i}{\lambda t_s} \exp\left(\frac{\pi i(x^2 + y^2)}{\lambda t_s}\right) \quad (2.10)$$

Die Elektronenwelle durchläuft auf diese Weise Teilschicht für Teilschicht die gesamte Probe. Man erhält schließlich die Wellenfunktion an der Probenunterseite (Austrittswellenfunktion $\psi(x, y, t)$).

2.2 Abbildung der Austrittswellenfunktion

Die Austrittswellenfunktion beschreibt den Elektronenstrahl nach der Wechselwirkung mit der Probe. Der Elektronenstrahl durchläuft nach dem Austritt aus der Probe die elektromagnetische Optik des Mikroskops und erzeugt durch Interferenz in der Bildebene die HREM-Abbildung. Beobachtet wird bei der Abbildung die Intensität der Elektronenwelle.

Statt den Abbildungsvorgang direkt durch Faltung zu berechnen, wird die Faltung numerisch effizienter unter Ausnutzung des Faltungstheorems (siehe [5]) im Fourierraum ausgeführt. Der Intensitätsbildung im Realraum entspricht im Fourierraum eine Autokorrelation. Die Abbildungsfehler des Mikroskops, die aus Kontrastgründen notwendige Defokussierung Z und die Inkohärenz der Elektronenstrahlung lassen sich im Fourierraum mit dem Transmissionskoeffizienten T berücksichtigen. Die Fouriertransformierte $I(\vec{k})$ der Bildintensität $I(\vec{r})$ ergibt sich bei partiell kohärenter Beleuchtung als

$$I(\vec{k}) = \int F(\vec{g} + \vec{k}) F^*(\vec{g}) T(\vec{g} + \vec{k}, \vec{g}) d\vec{g} \quad (2.11)$$

wobei $F(\vec{k})$ die Fouriertransformierte der Austrittswellenfunktion $\psi(x, y, t)$ repräsentiert [15]. Bei dieser Formulierung geht man davon aus, dass die Kohärenzeigenschaften des Elektronenstrahls keinen Einfluss auf die Beugung in der Probe haben.

Der Transmissionskoeffizient T lässt sich analytisch nicht exakt berechnen. Es wird eine Reihenentwicklung bis zur ersten Ordnung in Defokus und Wellenvektor \vec{g} durchgeführt. In dieser Näherung kann T faktorisiert dargestellt werden:

$$T(\vec{u}, \vec{v}) = A(u) A(v) P(u, v) E_s(\vec{u}, \vec{v}) E_t(u, v) \quad (2.12)$$

Die Aperturfunktion A beschreibt die Objektivblende. Es ist

$$A(k) = \begin{cases} 1 & \text{für } k \leq g_a \\ 0 & \text{sonst} \end{cases} \quad (2.13)$$

Wobei g_a der Radius der Objektivblende im reziproken Raum ist.

P repräsentiert die kohärente Kontrastübertragung

$$P(u, v) = \exp\left(-2\pi i \left(\frac{1}{4} C_s \lambda^3 (u^4 - v^4) - \frac{1}{2} \lambda Z (u^2 - v^2)\right)\right) \quad (2.14)$$

Die Defokussierung Z wird hier als positiv für Unterfokussierung (Brennweite größer als für korrekte Abbildung erforderlich) definiert.

Die räumliche Inkohärenz wird durch E_s wiedergegeben

$$E_s(\vec{u}, \vec{v}) = \exp\left(-\pi^2 q_0^2 \left| \vec{u} (C_s \lambda^2 u^2 - Z) - \vec{v} (C_s \lambda^2 v^2 - Z) \right|^2\right) \quad (2.15)$$

und E_t beinhaltet die zeitliche Inkohärenz

$$E_t(u, v) = \exp\left(-\frac{\pi^2 d_f^2 \lambda^2}{4} (u^2 - v^2)^2\right) \quad (2.16)$$

Die Strahlhalbkonvergenz q_0 gibt eine Obergrenze für den Winkel der Teilstrahlen des konvergenten Elektronenstrahls (Beleuchtung) an. Dies ist ein Maß für die räumliche Kohärenz der Elektronenquelle. Die von der Elektronenquelle emittierten Elektronen sind nicht monochromatisch, sondern weisen eine Energieverteilung auf (partielle zeitliche Kohärenz). Da Elektronen verschiedener Energie durch die chromatische Aberration C_c der elektromagnetischen Linsen auf unterschiedliche Brennebenen fokussiert werden, führt dies zu einer endlichen Ausdehnung des Fokus (Defokusstreuung d_f , engl. *defocus spread*).

Die HREM-Simulationen dieser Arbeit wurden mit dem EMS Programmpaket durchgeführt (s. Kap. 2.3). Für die an der Abbildung beteiligten Parameter liegen den durchgeführten Simulationen folgende Werte zugrunde (JEOL JEM-4000EX UHRTEM configuration [24]):

Strahlhalbkonvergenz	q_0	0.7 mrad
Defokusstreuung	d_f	10nm
Sphärische Aberration	C_s	1mm
Objektivblende	$2g_a$	15nm^{-1}

Translationssymmetrische Austrittswellenfunktion

Eine Translationssymmetrie in der Austrittswellenfunktion $\psi(x, y)$ findet man auch in der Abbildung wieder. Dies begründet sich wie folgt:

Ist die Austrittswellenfunktion translationssymmetrisch, z.B.

$$\psi(x, y) = \psi(x + a, y) \quad (2.17)$$

so ist das Raumfrequenzspektrum in k_x Richtung diskretisiert:

$$F(k_x, k_y) = \sum_{l_x} C_{l_x}(k_y) \delta\left(k_x - \frac{l_x}{a}\right) \quad (2.18)$$

Das Abbildungsintegral (2.11) hat damit nur von Null verschiedene Werte wenn k_x ganzzahliges Vielfaches von $\frac{1}{a}$ ist. Da für $I(\vec{r})$ damit nur translationssymmetrische Funktionen überlagert werden, setzt sich die Translationssymmetrie der Austrittswellenfunktion auch in der Abbildung fort:

$$I(x, y) = I(x + A, y) \quad (2.19)$$

Die Periodizität A im Bild (z.B. Größenordnung mm) hat aufgrund der Vergrößerung des Mikroskops natürlich einen anderen Wert als die Periodizität a der Austrittswellenfunktion (z.B. $a = 5.66 \text{ \AA}$).

2.3 Das Programmpaket EMS

Zur Simulation von HREM-Abbildungen wurde das Programmpaket EMS von P. Stadelmann verwendet [42],[43]. EMS ermöglicht eine Berechnung der dynamischen Elektronenbeugung sowohl mit dem *multislice*-Verfahren als auch mit dem Bethe-Bloch-Verfahren. Die Absorption der Probe kann entweder proportional zum Streupotential [19] oder mit den absorptiven Formfaktoren von A. Weickenmeier und H. Kohl [52] berechnet werden. Die Abbildung wird im Rahmen der nichtlinearen Abbildungstheorie in partiell kohärenter Näherung berechnet. Aus diesen Gründen ist EMS das meistverwendete Simulationsprogramm für HREM-Abbildungen [34] [44] [33].

Kapitel 3

Amorphe Oberflächenschichten

Zur quantitativen Analyse von HREM-Aufnahmen ist es notwendig, die Oberflächenamorphisierung der Probe zu berücksichtigen. In diesem Kapitel wird diese Störung genauer betrachtet.

3.1 Hintergrund

Auf der Oberfläche der Probe findet man stets amorphe Schichten von bis zu 10 nm Dicke. Die unregelmäßige Struktur dieser Schichten führt zu Rauschen in der HREM-Aufnahme. Abbildung 3.1 veranschaulicht die beiden amorphen Oberflächenschichten.

Bei der HREM-Probenpräparation muss eine Probendicke von wesentlich unter 50nm erreicht werden. Dazu wird die Probe am Anfang mechanisch auf eine Dicke von $5\mu\text{m}$ abgeschliffen. Abschliessend wird die Probe bis zur Perforation mit Ionen abgesputtert (*ion-milling*). Das Ionenabdünnen bei den in dieser Arbeit ausgewerteten Proben erfolgte mit Ar^+ Ionen der Energie 4-6 keV. Die Probe wurde beim Ionenabdünnen mit flüssigem Stickstoff gekühlt, um dabei entstehende Wärme abzuführen.

Der Hauptgrund für die Oberflächenamorphisierung der Probe ist dieses Ionenabdünnen während der HREM-Probenpräparation. Dabei entstehen amorphe Schichten welche in ihrer Zusammensetzung dem ursprünglichen (kristallinen) Probenmaterial entsprechen. Die Abdünnung der Querschnittsproben erfolgt von beiden Seiten der Probe. Man erwartet daher eine vergleichbare Dicke von Deckschicht und Abschlusschicht.

Zur quantitativen Analyse von amorphen Oberflächenschichten werden im nächsten Abschnitt drei Näherungsmodelle betrachtet.

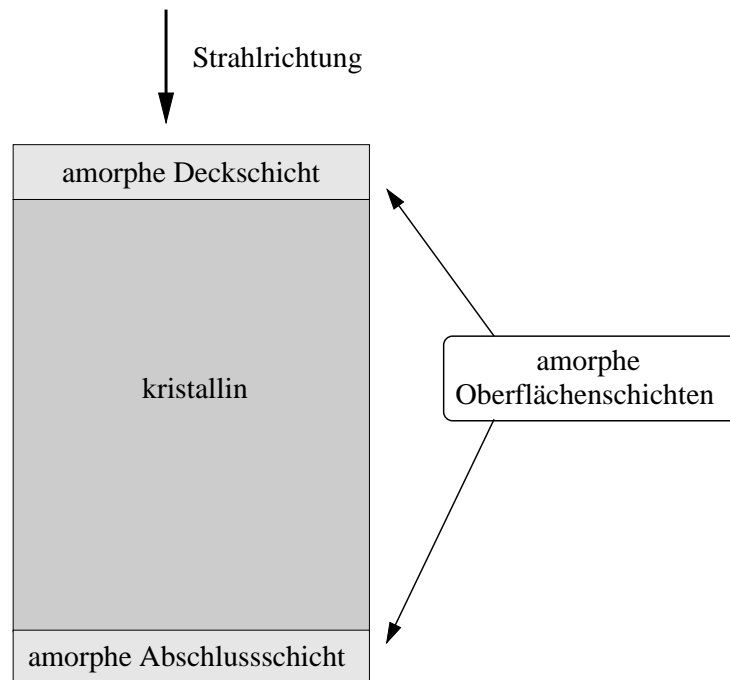


Abbildung 3.1: Amorphe Oberflächenschichten: Deckschicht und Abschlusschicht.

3.2 Simulation amorpher Oberflächenschichten

3.2.1 Übersicht über Modelle

Objekt mit zufälliger Phasenvariation (RPO)

In diesem Modell wird die Austrittswellenfunktion einer ungeordneten Oberflächenschicht durch ein Phasenobjekt mit zufälliger Phasenvariation genähert [37]. Dieses Modell ist sehr schnell berechenbar. Problematisch ist der fehlende Bezug zur Zusammensetzung des amorphen Materials, die fehlende Absorption und der fehlende Bezug zur Dicke der amorphen Schichten. Die Dicke der amorphen Schichten beeinflusst in komplexer Weise die Standardabweichung der Phasenvariation. Durch den fehlenden Bezug zur Probendicke ist mit diesem Modell auch keine Bestimmung der Dicke von experimentellen Oberflächenschichten möglich. Aus diesen Gründen wurde von einer Berechnung der amorphen Oberflächenschichten als Objekt mit zufälliger Phasenvariation abgesehen.

Objekt mit zufälligen Atompositionen (RDO)

Dieses Modell für amorphe Substanzen geht von statistisch unabhängigen Atompositionen aus. Die im amorphen Objekt vorliegende Paarverteilungsfunktion g wird so als konstant Eins genähert ($g(r)=1$). Mit dieser Annahme lässt sich die Phasengitterfunktion für das *multislice*-Verfahren numerisch effizient berechnen (siehe Abschnitt 3.2.2). Dieses Modell berücksichtigt sowohl die Zusammensetzung des Materials (atomare Streuamplituden) als auch Absorption der Schichten. Es kann über den Debye-Waller-Faktor die korrekte Temperaturabschwächung des Streupotentials berücksichtigt werden. Ausserdem hat dieses Modell einen direkten Bezug zur Dicke der amorphen Oberflächenschichten. Dadurch wird die in Abschnitt 3.3 vorgestellte Bestimmung der Dicke von amorphen Oberflächenschichten in experimentellen HREM-Aufnahmen möglich.

Modifiziertes Hardsphere Modell (HSM)

Eine Verfeinerung des RDO-Modells berücksichtigt auch die Paarverteilungsfunktion im amorphen Material.

Bei der Simulation von HREM-Abbildungen mit dem *multislice*-Verfahren wird die simulierte amorphe Schicht in mehrere Teilschichten unterteilt (s. Kap. 2.1). In diesem Abschnitt wird ein modifiziertes Hardsphere Modell untersucht, welches den Nächsten-Nachbar-Abstand innerhalb einer Teilschicht durch eine an experimentelle Daten gefittete Verteilung korrekt wiedergibt (Monte-Carlo-Simulation [21]).

Die Korrelation zwischen den Schichten wird vernachlässigt. Dieses Modell kann noch verfeinert werden, indem man auch die Abstände zu Nachbarn höherer Ordnung berücksichtigt. Selbst das einfache Modell, welches nur den Nächsten-Nachbar-Abstand verwendet, stellte sich für die untersuchten Bildfeldgrößen ($90 \times 90 \text{ \AA}$) als numerisch sehr aufwendig heraus.

Abbildung 3.2 ermöglicht einen Vergleich zwischen modifiziertem Hardsphere Modell und RDO. Dazu werden mit beiden Modellen 17 \AA (entspricht 3 Ez.) dicke amorphe GaAs Schichten berechnet und HREM-Abbildungen simuliert (Defokus 50 nm). Die zweidimensionalen Fouriertransformierten beider Abbildungen zeigen eine ringförmige Intensitätsverteilung, welche entlang konzentrischer Kreise (Betrag des rez. Vektors konstant) gemittelt wird (*radial averaging*). Abbildung 3.2 zeigt die so gemittelten Fourierintensitäten in Abhängigkeit des reziproken Abstands.

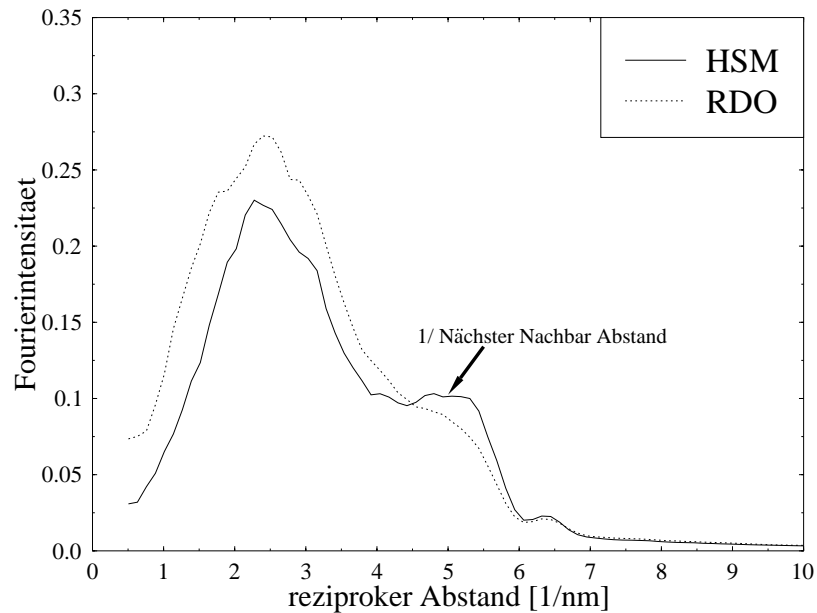


Abbildung 3.2: Vergleich von modifiziertem Hardsphere Modell (HSM) und Objekt mit zufälligen Atompositionen (RDO). Eingezeichnet ist die radial gemittelte fouriertransformierte Bildintensität zweier amorpher GaAs Schichten der Dicke 17\AA .

Das gröbere Modell (RDO, gepunktet) folgt im wesentlichen der Verteilung des exakteren modifizierten Hardsphere Modells. Die radiale Intensitätsverteilung wird vor allem durch die Kontrastübertragung des Mikroskops bestimmt. Diese wird vom RDO-Modell korrekt wiedergegeben. Das modifizierte Hardsphere Modell hat etwas geringere Intensität bei kleineren Raumfrequenzen (grosse Abstände) zu Gunsten einer Vergrößerung der Intensität bei der dem Nächsten-Nachbar-Abstand entsprechenden Raumfrequenz. Bei der Intensitätsverteilung über die Bildfrequenzen ist die Abweichung zwischen RDO-Modell und modifiziertem Hardsphere Modell gering.

Zur Berechnung der Phasengitterfunktion (1024×1024 Pixel) einer $90 \times 90 \times 17\text{\AA}$ GaAs Schicht (6144 Atome) benötigt die im nächsten Abschnitt vorgestellte Berechnung des RDO-Modells 2 Minuten im Vergleich zu 5 Stunden Zeitbedarf bei der Simulation mit modifiziertem Hardsphere-Modell. Mit dem numerischen Aufwand des modifizierten Hardsphere Modells wären die Simulationen in dieser Arbeit nicht möglich gewesen. Alle im weiteren Verlauf der Arbeit durchgeführten Simulationen von amorphen Oberflächenschichten verwenden das RDO-Modell.

3.2.2 Simulation eines Objekts mit zufälligen Atompositionen

Um in HREM-Bildsimulationen amorphe Oberflächenschichten (RDO-Modell) berücksichtigen zu können, berechnen wir in diesem Abschnitt die Phasengitterfunktion für das *multislice*-Verfahren (siehe (2.9)) teilsanalytisch.

Phasengitterfunktion

Im RDO-Modell haben die Atome in der amorphen Schicht voneinander statistisch unabhängige Positionen. Die amorphe Schicht der Dicke t wird in Unterschichten (*sub-slices*) aufgeteilt. Die Unterschichten sollen die Dicke t_s haben. Durch Propagation der Elektronenwelle durch mehrere solcher Unterschichten im Rahmen des *multislice*-Verfahrens wird die gesamte Dicke der amorphen Schicht aufgebaut.

Die Phasengitterfunktion einer Unterschicht der Dicke t_s ergibt sich aus dem Potential $V(\vec{r})$ wie folgt:

$$\text{pg}(x, y) = \exp \left(i\sigma \int_0^{t_s} V(x, y, z) dz \right) \quad (3.1)$$

(siehe (2.9)).

Wir berechnen das z-integrierte Potential als:

$$V_z = \int_0^{t_s} V(x, y, z) dz \quad (3.2)$$

wobei sich das Potential V als Faltung von atomarem Potential V_a und Dichtefunktion ρ zusammensetzt:

$$V(\vec{r}) = \int_{\vec{a} \in \mathbb{R}^3} V_a(\vec{a}) \rho(\vec{r} - \vec{a}) d\vec{a} \quad (3.3)$$

Aus 3.2 und 3.3 erhält man durch Umordnen der Integration:

$$V_z(x, y) = \int_{a_1, a_2 \in \mathbb{R}} \int_{a_3 \in \mathbb{R}} V_a \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \int_{z=0}^t \rho \begin{pmatrix} x - a_1 \\ y - a_2 \\ z - a_3 \end{pmatrix} dz da_3 da_1 da_2 \quad (3.4)$$

Wir verwenden einen allgemeinen Ansatz für die Anzahldichte. Die betrachtete Unterschicht soll im Raumvolumen $x \cdots x + dx$, $y \cdots y + dy$, $z = 0 \cdots t_s$ eine Anzahl

Atome $N(x, y)$ eines Elements enthalten. Die z-Koordinate der Atome ist dabei z_n für $n = 1 \cdots N(x, y)$. Die Dichte ist damit:

$$\rho \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \sum_{n=1}^{N(x,y)} \delta(z - z_n) \quad (3.5)$$

wobei die z_n von x und y abhängen und die Atome sich ausnahmslos innerhalb der betrachteten Unterschicht befinden sollen:

$$z_i \in [0, t_s] \quad \forall_{i=1 \cdots N(x,y)} \quad (3.6)$$

Die betrachtete Unterschicht wird durch (3.6) räumlich begrenzt. Die z-Integration (3.2) kann so genauer auf $z \in \mathbb{R}$ ausgewertet werden:

$$V_z(x, y) = \int_{a_1, a_2 \in \mathbb{R}} \int_{a_3 \in \mathbb{R}} V_a \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \int_{z=-\infty}^{\infty} \sum_{n=1}^{N(x-a_1, y-a_2)} \delta(z - a_3 - z_n) dz da_3 da_1 da_2 \quad (3.7)$$

Die z-Integration ergibt die Anzahl der Atome in der Säule:

$$V_z(x, y) = \int_{a_1, a_2 \in \mathbb{R}} N(x - a_1, y - a_2) \int_{a_3 \in \mathbb{R}} V_a \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} da_3 da_1 da_2 \quad (3.8)$$

Mit dem z-integrierten atomaren Potential V_{za}

$$V_{za}(a_1, a_2) = \int_{a_3 \in \mathbb{R}} V_a \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} da_3 \quad (3.9)$$

ergibt sich V_z als zweidimensionale Faltung:

$$V_z(x, y) = \int_{a_1, a_2 \in \mathbb{R}} N(x - a_1, y - a_2) V_{za}(a_1, a_2) da_1 da_2 \quad (3.10)$$

Numerisch effizient wertet man diese Faltung im Fourierraum aus, unter Ausnutzung des Faltungstheorems:

$$V_z(x, y) = FT^{-1} [FT [N(x, y)] FT [V_{za}(x, y)]] \quad (3.11)$$

Dies bringt Vorteile mit sich, weil der Zeitbedarf der schnellen Fouriertransformation (FFT) mit $(N_p \log_2(N_p))^2$, die direkt ausgeführte Faltung aber mit N_p^4 skaliert (N_p ist die Anzahl der Pixel im berechneten Bild). Bei der praktischen Berechnung des z-integrierten Potentials für eine Unterschicht verwenden wir einen rechteckigen Speicherbereich $N(x = 1 \cdots N_x, y = 1 \cdots N_y)$. Die Größe dieses Speicherbereichs entspricht der Größe des später berechneten Bildes ($N_x \times N_y$ Pixel). Die Anzahl der Atome in der Unterschicht N_a ergibt sich als Produkt aus Schichtvolumen und Anzahldichte des amorphen Materials. Die Speicherzellen des Speicherbereichs werden mit Null vorbelegt und N_a zufällig ausgewählte Speicherzellen jeweils um 1 inkrementiert. Schließlich fouriertransformiert man den Speicherbereich und multipliziert zu den Speicherzellen der Fouriertransformierten $FT[N(x, y)]$ jeweils den Wert $FT[V_{za}(x, y)]$ (Die Berechnung dieses Wertes wird später dargestellt). So erhält man nach (3.11) das integrierte Potential für die Unterschicht und nach (3.1) und (3.2) die Phasengitterfunktion für das *multislice*-Verfahren.

Zu berechnen ist noch die Fouriertransformierte des z-integrierten Potentials. Dies kann analytisch durchgeführt werden:

$$\begin{aligned}
 FT[V_{za}(x, y)](K_1, K_2) = \\
 \int_{R_x \in \mathbb{R}} \int_{R_y \in \mathbb{R}} \int_{a_3 \in \mathbb{R}} V_a \begin{pmatrix} R_x \\ R_y \\ a_3 \end{pmatrix} \exp(2\pi i \vec{R} \vec{K}) da_3 dR_y dR_x \quad (3.12)
 \end{aligned}$$

Mit der zweidimensionalen Fouriertransformierten des atomaren Potentials an der Stelle a_3 :

$$f(\vec{K}, a_3) = \int_{R_x \in \mathbb{R}} \int_{R_y \in \mathbb{R}} V_a \begin{pmatrix} R_x \\ R_y \\ a_3 \end{pmatrix} \exp(2\pi i \vec{R} \vec{K}) dR_y dR_x \quad (3.13)$$

ergibt sich:

$$FT[V_{za}(x, y)](K_1, K_2) = \int_{a_3 \in \mathbb{R}} f(\vec{K}, a_3) da_3 \quad (3.14)$$

Anhang C skizziert die Berechnung von f aus der atomaren Streuamplituden nach Doyle und Turner.

Als Ergebnis erhalten wir:

$$f(\vec{K}, z) = \frac{h^2}{m_0 e \sqrt{\pi}} \sum_{j=1}^4 \frac{\alpha_j}{\sqrt{\beta_j}} \exp\left(-\frac{K^2 \beta_j}{4}\right) \exp\left(-\frac{4\pi^2 z^2}{\beta_j}\right) \quad (3.15)$$

mit: $K_1, K_2 [\text{\AA}^{-1}]$

Damit ergibt sich aus Gleichung (3.14) :

$$FT[V_{za}(x, y)](K_1, K_2) = \frac{h^2}{2\pi m_0 e} \sum_{j=1}^4 \alpha_j \exp\left(-\frac{K^2 \beta_j}{4}\right) \quad (3.16)$$

Dies entspricht der ursprünglichen atomaren Streuamplitude von Doyle und Turner bei $k_3 = 0$ (reziproke Schnittebene senkrecht zur Strahlrichtung).

Ternäre amorphe Schicht

Mit der VCA-Näherungsmethode (*virtual crystal approximation*), welche auch in EMS realisiert ist, berechnen wir amorphe Schichten von komplizierterer Zusammensetzung. Dazu wird kompositionsabhängig zwischen den atomaren Streuamplituden der beteiligten Elemente gemittelt (vgl. Anhang B, Berechnung in Gleichung (B.2)). Die daraus resultierende Streuamplitude wird in der Berechnung der Phasengitterfunktion verwendet.

Temperaturabschwächung des Streuvermögens

Analog zur Röntgenbeugung führt auch bei der Beugung der Elektronenwelle die Wärmebewegung der Atome zu einer Verringerung der Beugungsintensität. Analog zu EMS tragen wir dem durch einen Zusatzfaktor zum Potential

$$V'_z(\vec{k}) = V_z(\vec{k}) \exp(-25 B_a k^2) \quad (3.17)$$

Rechnung. Dabei ist B_a der Debye-Waller-Faktor [nm^2]. Für die hier betrachteten Materialsysteme erhält man aus der Literatur Werte von

$$B_a = 0.005 \text{ nm}^2 \dots 0.006 \text{ nm}^2 \quad (3.18)$$

In den Rechnungen wurde ein Wert von $B_a = 0.005 \text{ nm}^2$ verwendet. Dieser entspricht dem in der EMS Dokumentation [43] vorgeschlagenen Wert.

Absorption der Probe

Die Absorption des amorphen Materials berücksichtigen wir analog dem EMS Teilprogrammen sc1 - sc4 durch Vorgabe einer Absorptionskonstante ζ . Die Absorption wird dabei als proportional zum Streupotential genähert. Man ersetzt das reelle z-integrierte Potential von (3.17) durch

$$V_z''(\vec{k}) = V_z'(\vec{k})(1 + i\zeta) \quad (3.19)$$

Nach [19] wurde der Wert

$$\zeta = 0.1 \quad (3.20)$$

verwendet.

Beim (dickeren) kristallinen Teil der Probe ist die Absorption entscheidender. Diese Simulationen wurden daher mit den absorptiven Formfaktoren von A. Weickenmeier und H. Kohl (EMS Programm sc5) [52] gerechnet.

Vorteile der Berechnung

Die hier vorgestellte teilanalytische Berechnung der Phasengitterfunktion reduziert die Rechenzeit der *multislice*-Simulation erheblich. Gegenüber einer direkten Berechnung mit EMS aus den Atomkoordinaten ergibt sich ein um den Faktor 12 geringerer Zeitbedarf. Der große Zeitvorteil rührt daher, dass bei der teilanalytischen Vorgehensweise die numerische Berechnung des Projektionsintegrals wegfällt.

3.3 Bestimmung experimenteller Amorphisierungstiefen

In diesem Abschnitt wird eine im Rahmen dieser Arbeit entwickelte Methode vorgestellt, um bei experimentellen HREM-Bildern die Dicke der amorphen Oberflächenschichten zu bestimmen.

3.3.1 Beschreibung der Dickenbestimmung

Betrachten wir einen quadratischen Bildausschnitt einer Probe von der die Dicke der amorphen Oberflächenschichten bestimmt werden soll. Solche Bildausschnitte entsprechen bei dem hier vorgestellten Verfahren typischerweise Probenbereichen von etwa $100 \times 100 \text{ \AA}$. Hierbei ist zu beachten, dass Bereiche ausgewählt werden, die homogenes Material enthalten (kaum Schwankungen der Komposition und der kristallinen Proben Dicke).

Der Bildausschnitt wird fouriertransformiert und das so erhaltene Diffraktogramm ausgewertet. Abbildung 3.3 zeigt einen experimentellen Bildausschnitt und das zugehörige Diffraktogramm. Man erkennt deutlich die Reflexe und das ringförmige (nicht richtungsabhängige) Rauschen im Diffraktogramm.

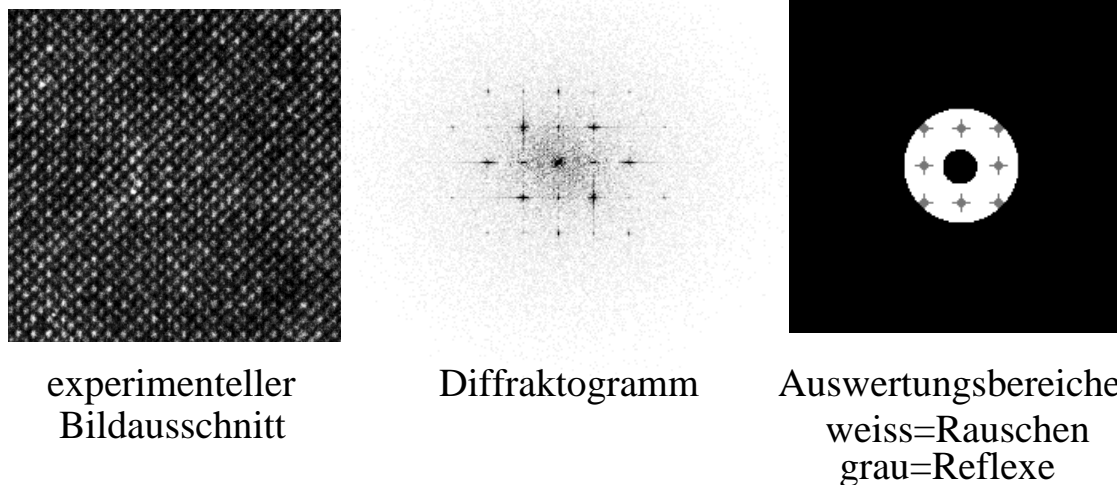


Abbildung 3.3: Skizze der amorphen Dickenbestimmung.

Wir wollen die Intensität des Rauschens mit der Intensität der Reflexe vergleichen.

Zunächst bestimmen wir die Intensität der Reflexe I_R . Dazu addieren wir die Intensität der hellsten Reflexe auf kreisförmigen Bereichen um das jeweilige Reflexzentrum auf. Der Nullreflex bleibt natürlich davon ausgenommen.

Durch die starke Konzentration der Reflexintensität kann der Radius der Auswertebereiche sehr klein gewählt werden. Es wurde hierfür ein Radius von 0.4 nm^{-1} verwendet. Durch Abbrucheffekte kommt es zu sogenannter Streakbildung bei den Reflexen. Dabei wird Reflexintensität entlang der x und y Richtung verteilt. Die Streakintensität wurde zur Reflexintensität gerechnet und innerhalb von 0.8 nm^{-1} ausgewertet.

Um die Intensität des Rauschens I_N abzuschätzen, betrachten wir die Intensität außerhalb der Reflexe. Bei dem auszuwertenden Bereich werden von der Gesamtfläche des Diffraktogramms die schon ausgewerteten Reflexbereiche ausgeschlossen.

Sind im Bildbereich Dicken- oder Kompositionsschwankungen vorhanden, so verändern diese vor allem die Intensität bei kleinen Raumfrequenzen. Diese Variationen verfälschen das Ergebnis für die Bestimmung der Amorphisierungstiefe und müssen daher unterdrückt werden. Darum werden für die Bestimmung der Amorphisierungstiefe nur Raumfrequenzen ausgewertet, die oberhalb der Frequenz liegen, welche der reziproken Gitterkonstante entspricht (z.B. GaAs $\frac{1}{a} = 1.77 \text{ nm}^{-1}$). Dies ist äquivalent zu der Annahme, dass Dicke und Komposition vor allem in Abständen variieren die größer sind als die Elementarzelle.

Das amorphe Rauschen ist nicht der einzige Beitrag zum Rauschen. Bei hohen Raumfrequenzen äußern sich vor allem Einflüsse wie Körnung des Photomaterials (bzw. Pixel-Schwankungen oder Ausfälle der CCD). Um diese Störung gering zu halten, wird daher die Intensität des amorphen Rauschens nur bei Raumfrequenzen bis zum $\frac{1}{e}$ Wert der atomaren Streuamplitude ausgewertet. Betrachtet man einen GaAs Bereich, so liegt der $\frac{1}{e}$ Wert der mittleren Streuamplitude bei 5.84 nm^{-1} (entspricht im Realraum 1.71 \AA).

Die Auswertungsbereiche für ein experimentelles Beispiel zeigt Abbildung 3.3.

Wertet man große Bereiche ($> 63 \times 63 \text{ \AA}$) mit hinreichender Samplingrate (hier 20-64 Pixel pro 0.57 nm) aus, so kann man experimentelle und simulierte Bilder vergleichen, und auf diese Weise die Dicke der amorphen Oberflächenschichten im Experiment bestimmen.

Dazu wird eine kristalline Probe mit schrittweise steigenden amorphen Schichtdicken simuliert und jeweils der Kennwert $\frac{I_N}{I_R}$ bestimmt. Abbildung 3.4 zeigt solche Kennwerte für einen GaAs Bereich (kristalline Dicke 70 \AA ; Defokussierung 58 nm). Die Punkte werden linear interpoliert. Die interpolierte Kurve ist umkehrbar. Somit kann aus dem Wert $\frac{I_N}{I_R}$ einer experimentellen Aufnahme, die amorphe Dicke bestimmt werden.

3.3.2 Diskussion der Dickenbestimmung

Die Auswertung experimenteller Aufnahmen erfordert eine Unabhängigkeit der Werte im Bereich der auftretenden Größen der Bildausschnitte und der Samplingraten.

Dies wurde gewährleistet durch eine Festlegung aller Radian in reziproken Raumvektoren (nm^{-1}). Die dafür notwendige Skalierung des Diffraktogramms (Pixel $\rightarrow \text{nm}^{-1}$) erreicht man durch interaktive Reflexauswahl und Raumfrequenzzuordnung.

Die Ergebnisse sind in weitem Bereich unabhängig von der Wahl des Reflexauswerteradius. Weil die Reflexintensität stark konzentriert ist, führt ein Vergrößern dieses Radius um 50 Prozent zu praktisch keiner Veränderung der ausgewerteten Reflexintensität I_R . Eine solche Vergrößerung verkleinert den großen Bereich in dem die Intensität des Rauschens ausgewertet wird nur unwesentlich. Daher ändert sich auch das Ergebnis für die Intensität des Rauschens nur unmaßgeblich. Der Einfluss des Reflexauswerteradius auf den Wert $\frac{I_N}{I_R}$ kann somit vernachlässigt werden.

Gleiches gilt für eine Variation der Größe der Streakbereiche.

Fordert man einen Durchmesser des Reflexauswerteradius von mindestens 5 Pixel (entspricht 0.4 nm^{-1}), so muss der ausgewertete Bildausschnitt einen Probenausschnitt von wenigstens 63 \AA abbilden. Dies war bei allen untersuchten Bildausschnitten der Fall. Zudem soll die oberste Raumfrequenz im Bild größer sein als die obere Frequenz des Bandpassfilters (GaAs 5.84 nm^{-1}). Daraus ergibt sich eine Samplingrate von mindestens 7 Pixel pro Elementarzelle (5.7 \AA). Die kleinste experimentell gegebene Samplingrate lag bei 20 Pixel pro Elementarzelle.

Im Gegensatz zu den in Abschnitt 5.4 dargestellten Untersuchungen an der lokalen Fouriertransformierten (Fouriertransformierte auf kleinen Bildausschnitten) sind die Abweichungen aufgrund der Aperiodizität hier vernachlässigbar. In Abschnitt 5.2 wird die Vergleichbarkeit von experimentellen und simulierten Bildern diskutiert. Bei dem hier betrachteten Wert $\frac{I_N}{I_R}$ kürzen sich eventuelle Kontrastunterschiede zwischen Experiment und Simulation heraus. Ein Unterschied in der Helligkeit von simuliertem und experimentellem Bild äußert sich nur in der Intensität des Nullreflexes und beeinflusst daher den Wert $\frac{I_N}{I_R}$ ebenfalls nicht.

Die hier vorgestellte Dickenbestimmung von amorphen Schichten ist abhängig von der kristallinen Dicke der Probe und vom Defokus im Abbildungsprozess. Daher ist für Proben bei denen diese Parameter unbekannt sind, eine Parameterbestimmung (siehe Abschnitt 5.3) durchzuführen.

Zur Bestimmung von amorphen Schichtdicken sind nur homogene Bereiche geeignet. Insbesondere sollten nur Bereiche mit binärem Halbleitermaterial verwendet werden. Proben von ternärem Halbleitermaterial zeigen in ihrer Komposition durch die statistische Besetzung eines Untergitters sogenannte Random-Alloy-Fluktuationen. Dadurch schwankt die Komposition auch innerhalb der Ausdehnungen der Sphalerit-Elementarzelle deutlich. Die durchgeführte Hochpassfilterung unterdrückt diesen Effekt nur unzureichend. Die Dicke der amorphen Oberflächenschichten wird daher bei der Anwendung des Verfahrens auf ternäre Bereiche überschätzt.

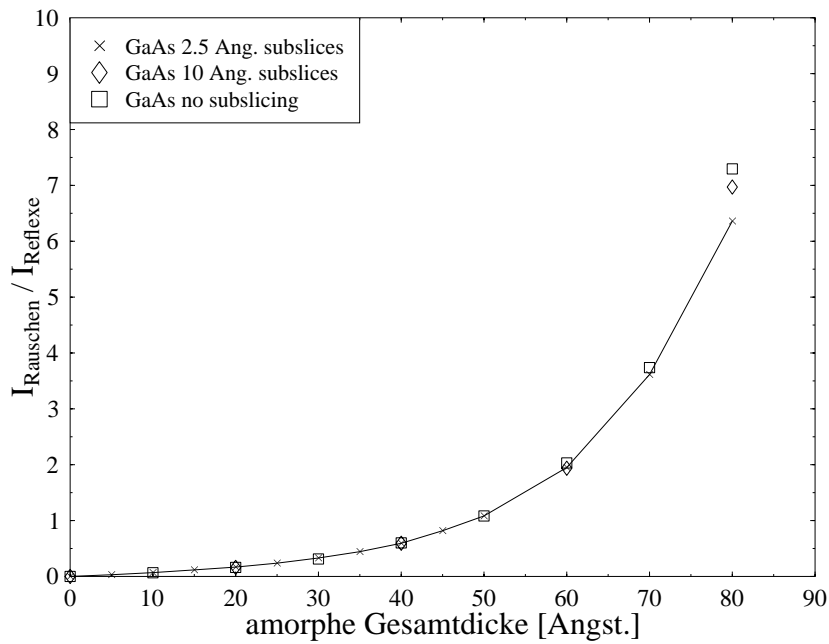


Abbildung 3.4: Vergleich von amorphen Schichten mit unterschiedlicher Aufteilung in Unterschichten (*subslices*).

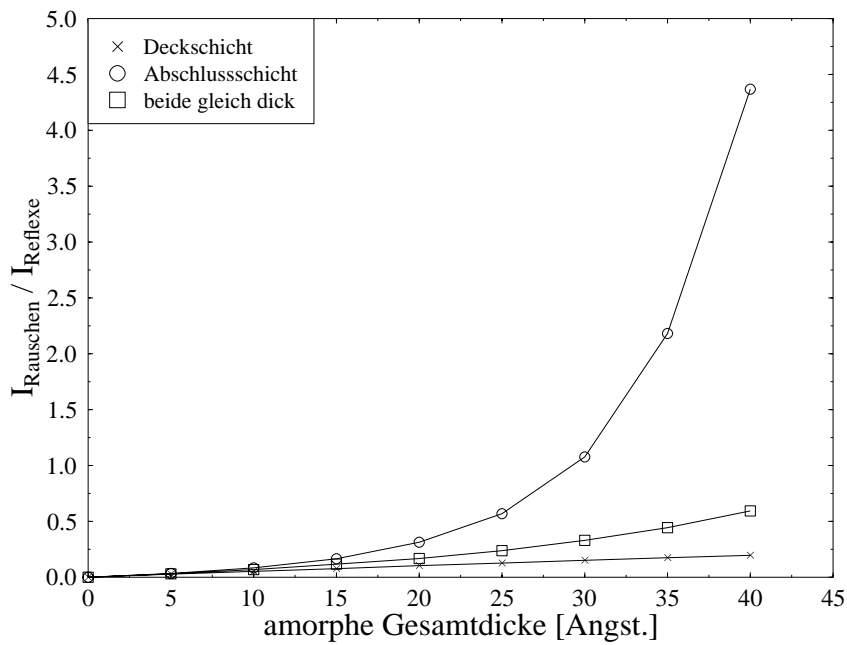


Abbildung 3.5: Vergleich von Deckschichttrauschen mit Abschlusschichttrauschen.

Abbildung 3.4 zeigt verschiedene Unterteilungen der amorphen Schichten in Unterschichten (*subslices*). Im Vergleich dazu ist das Ergebnis der Berechnung der amorphen

Oberflächenschichten ohne Aufteilung in Unterschichten (vgl. Abschnitt 3.2.2: $t_s = t$) dargestellt. Es zeigt sich, dass die Unterteilung erst ab einer Gesamtschichtdicke von 60Å deutliche Unterschiede erzeugt.

Bei den in dieser Arbeit untersuchten HREM-Proben erfolgte das Ionenabdünnen auf beiden Probenseiten. Daher wird man bei Deck- und Abschlusschicht vergleichbare Dicken erwarten. Dies wird im Folgenden durch experimentelle Auswertung belegt.

Wir betrachten die beiden Fälle einer einseitigen Amorphisierung (nur Deckschicht bzw. nur Abschlusschicht) im Vergleich zur symmetrischen Schichtbelegung (Deckschicht und Abschlusschicht gleich dick). Abbildung 3.5 erlaubt einen Vergleich der drei Fälle anhand von simulierten $\frac{I_N}{I_R}$ Kennwerten. Die Deckschicht erzeugt im Vergleich zur Abschlusschicht deutlich weniger Rauschen in der Abbildung. Die Ursache hierfür ist die Raumfrequenzfilterung der Deckschicht-Störung durch den kristallinen Teil der Probe.

Angenommen die Probe wäre mit einer Dicke von 30Å einseitig amorphisiert. Mit unserem Verfahren zur amorphen Schichtdickenbestimmung erhalten wir aus dem Vergleich mit einer symmetrischen Schichtbelegung folgende Ergebnisse:

Gegebene Situation	äquivalente Gesamtdicke bei sym. Schichten
30Å Deckschicht	18Å
30Å Abschlusschicht	50Å

Bei einer einseitigen Amorphisierung unterscheiden sich die beiden möglichen Orientierungen der amorphen Schicht (Deckschicht bzw. Abschlusschicht) deutlich im Ergebnis der Schichtdickenbestimmung. Bei den durchgeführten HREM-Aufnahmen wurde keine explizite Orientierung der amorphen Schichten durchgeführt. Welche der beiden Orientierungen einer einseitigen amorphen Schicht bei der HREM-Aufnahme vorliegt, ist somit Zufall.

Betrachten wir eine Serie von einseitig amorphisierten Proben vergleichbarer Amorphisierungstiefe. Bei der Anfertigung der HREM-Abbildung ist für jede Probe der Serie die Wahrscheinlichkeit für die beiden möglichen Orientierungen im Mikroskop gleich hoch. Führt man eine Bestimmung der amorphen Schichtdicke an den einzelnen Proben der Serie durch, wird man deshalb starke Schwankungen bei den Ergebnissen erwarten. Die Auswertung an fünf experimentellen Zweischichtsystemen (vergleichbare Probenpräparation) ergab im GaAs Teil der Proben folgende Werte¹:

¹Zur Bestimmung der Defokussierungen und der kristallinen Probendicken siehe Abschnitt 5.

Negativ	krist. Dicke	Defokus	amorphe Gesamtdicke
3d188	12 Ez.	58 nm	31.8Å
3d158	14 Ez.	50 nm	28.0Å
3d152	24 Ez.	50 nm	36.3Å
3d150	15 Ez.	51 nm	28.1Å
3d147	22 Ez.	53 nm	36.0Å

Die Ergebnisse für die amorphen Schichtdicken streuen nur zwischen 28Å und 36Å. Offensichtlich ist kein Wechsel zwischen Deckschicht und Abschlusschicht aufgetreten. Nimmt man eine einseitige Amorphisierung an, so ist die Wahrscheinlichkeit hierfür nur $\frac{1}{24}$. Eine einseitige Amorphisierung kann mit diesen fünf Experimenten daher mit 94 Prozent Signifikanz ausgeschlossen werden.

3.4 Auswirkungen der Amorphisierung

In diesem Abschnitt wird der Einfluss der amorphen Oberflächenschichten auf den Kontrast der HREM-Aufnahme eines simulierten Beispiels untersucht.

Der kristalline Teil der simulierten GaAs Probe hat eine Dicke von 12 Elementarzellen. Die Aufnahme wurde mit einer Defokussierung von 58nm simuliert. In einer Serie von HREM-Bildsimulationen wurde die Probe mit amorphen Oberflächenschichten schrittweise steigender Dicke symmetrisch belegt (Deck- und Abschlusschicht gleich dick).

Um die Auswirkung der amorphen Oberflächenschichten auf den HREM-Bildkontrast abzuschätzen, berechnen wir den Kontrastanteil Rauschen. Der Kontrastanteil Rauschen ergibt sich aus dem Verhältnis zwischen Kontrast durch Rauschen und dem Gesamtkontrast.

Der Gesamtkontrast K_g ist die Standardabweichung der Intensität in allen Bildpositionen (Pixels).

Den Kontrast durch Rauschen K_r berechnen wir wie folgt. Die simulierte HREM-Aufnahme läßt sich in Bereiche (Zellen) unterteilen, die der Größe der Elementarzelle entsprechen und jeweils gleichviele Bildpunkte (64×64) enthalten. Vergleicht man in einer ungestörten Aufnahme (homogene Probe !) den äquivalenten Bildpunkt in verschiedenen Zellen, so hat dieser immer den selben Helligkeitswert. Die Standardabweichung über einen äquivalenten Bildpunkt verschiedener Zellen hat ihre Ursache nur im Rauschen. Diese Standardabweichungen mitteln wir über die verschiedenen Bildpunkte

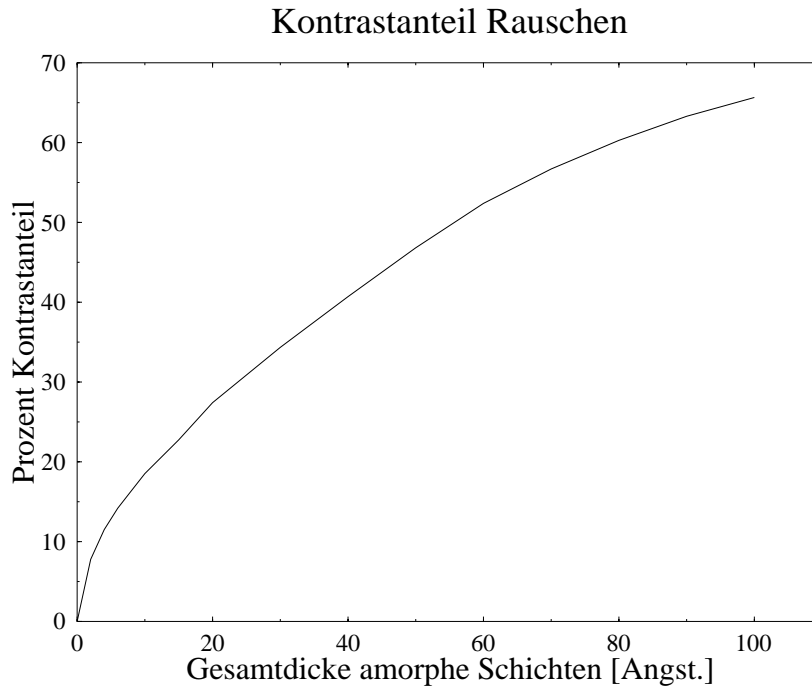


Abbildung 3.6: Kontrastanteil Rauschen in Abhängigkeit von der Gesamtdicke der amorphen Oberflächenschichten (simulierte GaAs Schicht, kristalline Dicke 68\AA , Defokus 58nm).

der Zelle. Abbildung 3.6 stellt den prozentualen Anteil des Rauschens am Bildkontrast für verschiedene Dicken der Oberflächenschichten dar.

Bei den zu dieser Arbeit untersuchten Aufnahmen von AlGaAs Proben traten in den GaAs Bereichen für die Gesamtdicke der amorphen Oberflächenschichten Werte von $20\text{-}40\text{\AA}$ auf. Dem entspricht ein Kontrastanteil des Rauschens von $25\text{-}40$ Prozent. Der Bildkontrast wird also zu etwa einem Drittel nur vom Rauschen bestimmt.

Abbildung 3.7 stellt den Kontrast durch Rauschen in Abhängigkeit von der amorphen Schichtdicke dar. Der Zusammenhang zwischen Kontrast durch Rauschen und Schichtdicke läßt sich in einem einfachen Modell verstehen.

Betrachten wir eine diskretisierte (in $N_p = N_x \times N_x$ Pixel aufgeteilte) quadratische amorphe Probe mit Dicke t und lateraler Größe $R \times R$. Die Probe soll N Atome enthalten. Die Atompositionen sind in unserm Modell statistisch unabhängig. Daher unterliegt die Anzahl der Atome, die sich auf den verschiedenen Pixel befinden, einer Binomialverteilung.

Setzt man zufällig ein Atom auf die Probe, so ist die Wahrscheinlichkeit einen bestimmten Pixel zu treffen $P = \frac{1}{N_p}$.

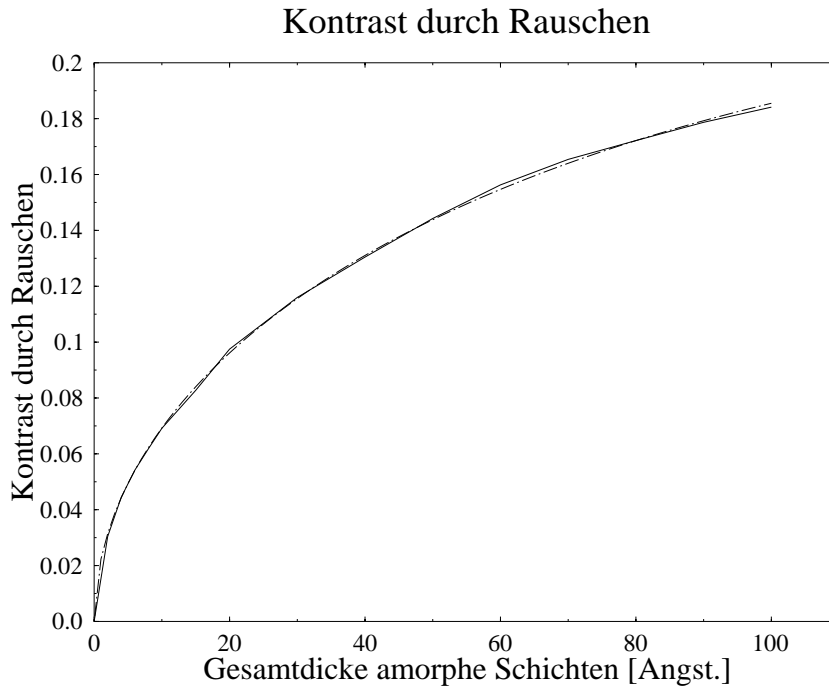


Abbildung 3.7: Kontrast durch Rauschen in Abhängigkeit von der Gesamtdicke der amorphen Schichten. Gestrichelt eingezeichnet ist das Ergebnis aus einem einfachen theoretischen Modell (3.23). Beschreibung der Simulation s. Abb. 3.6.

Die Varianz für die Anzahl von Atomen auf den verschiedenen Pixel ist

$$\text{Var} = P(1 - P)N \approx PN = \frac{\rho R^2 t}{N_p} \quad (3.21)$$

Die Näherung in (3.21) ist möglich, da man in der Praxis stets mehr als 1600 Pixel (mind. 40×40 Pixel Bild) verwendet. Die Standardabweichung für die Anzahl der Atome auf den verschiedenen Pixel ist somit proportional zu \sqrt{t} . Damit ist auch die Standardabweichung für das Streuvermögen auf den verschiedenen Pixel zu \sqrt{t} proportional. Der Kontrast im Bild hängt näherungsweise linear von der Standardabweichung des Streuvermögens auf den einzelnen Pixel ab. Für kleine Schichtdicken t nimmt der Kontrast durch Rauschen daher mit \sqrt{t} zu.

Für größere Schichtdicken wird die Gesamtintensität und damit auch der Kontrast durch die Absorption der Probe reduziert. An einer simulierten amorphen GaAs Schicht wurde der Absorptionskoeffizient β für das Beersche Gesetz

$$I(t) = I_0 \exp(-\beta t) \quad (3.22)$$

als $\beta = 0.001841 \text{ \AA}^{-1}$ bestimmt.

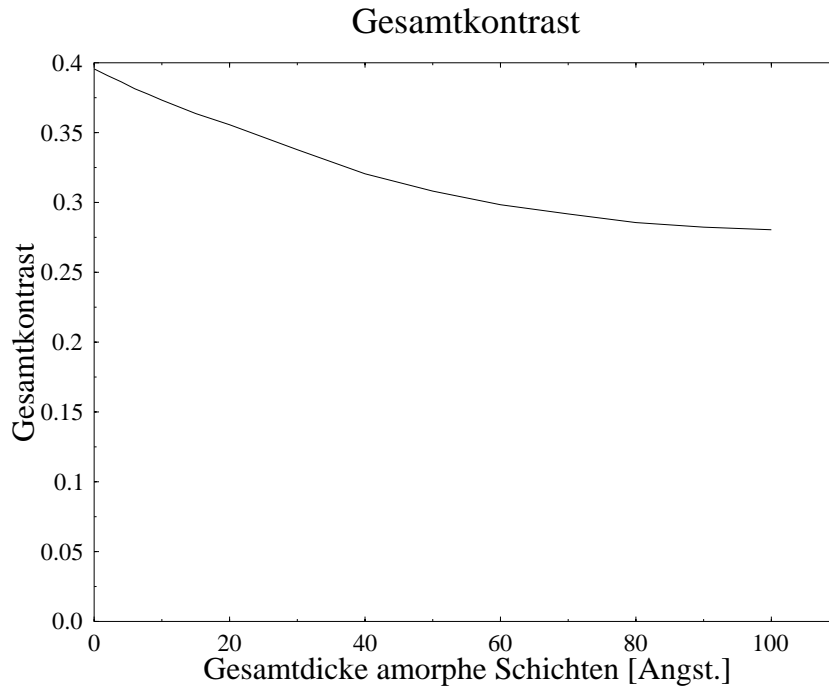


Abbildung 3.8: Gesamtkontrast einer GaAs Probe in Abhängigkeit von der Gesamtdicke der amorphen Schichten (simulierte GaAs Schicht, kristalline Dicke 68Å, Defokus 58nm).

In Abbildung 3.7 ist dem Kontrast durch Rauschen aus der Simulation das theoretische Modell

$$K_r(t) = a\sqrt{t} \exp(-\beta t) \quad (3.23)$$

gegenübergestellt. Der Parameter a wurde dabei gefittet. Es ergibt sich eine sehr gute Übereinstimmung. Im hier dargestellten (praxisrelevanten) Bereich für die Amorphisierungstiefe steigt der Kontrast des Rauschens mit der Dicke der amorphen Oberflächenschichten. Asymptotisch sinkt der Kontrast für zunehmende Oberflächenschichtdicke ab einer Dicke von $t > \frac{1}{2\beta}$ durch die Absorption des amorphen Materials. Allerdings ist die dafür notwendige Dicke der amorphen Oberflächenschichten nicht praxisrelevant (bei GaAs $t > 272\text{Å}$).

Abbildung 3.8 zeigt den Zusammenhang zwischen Gesamtkontrast im Bild und amorpher Gesamtdicke aus der Simulation. Bei den zu dieser Arbeit durchgeführten Auswertungen an experimentellen Aufnahmen ergaben sich für die amorphe Gesamtdicke Werte von 20-40Å. Bei diesen Amorphisierungstiefen verringert sich der Gesamtkontrast in unserem Beispiel zwischen 10 und 20 Prozent.

Der Zusammenhang zwischen Gesamtkontrast und amorpher Schichtdicke ist nicht durch ein einfaches Modell anzunähern. Probleme bereitet dabei der Kontrast aus dem

kristallinen Teil der Probe. Der kristalline Kontrast ist auf wenige Raumfrequenzen verteilt (Reflexe) und hängt deshalb von den Details der Kontrastübertragung ab. Asymptotisch nähert sich der Gesamtkontrast dem Zusammenhang $K_g(t) \approx K_r(t)$ (im Grenzfall wird der Kontrast nur noch durch die amorphe Substanz bestimmt).

Kapitel 4

Bildvorverarbeitung

4.1 Warum Bildvorverarbeitung ?

Bilddaten bestehen aus Bildpunkten (sog. Pixels) in einer Matrixanordnung. Typischerweise enthalten die zahlreichen Pixel des Bildes nicht alle voneinander unabhängige Information. Die Bildpunkte sind korreliert, das Bild enthält Redundanz. Die Information die man aus dem Bild erhalten möchte, hat in der Regel sehr geringe Dimension. Bei der hier angestrebten Kompositionsbestimmung ist die gesuchte Information eindimensional. Die hohe Dimension (entspricht der Anzahl der Pixel) der Bilddaten muss reduziert werden. Dabei soll die Redundanz der Daten verkleinert werden und möglichst keine Information verlorengehen. Dies geschieht unter Verwendung von Vorabinformation (Information zusätzlich zu den Bilddaten) über die Bildentstehung [3]. Diese Vorabinformation wird im folgenden Kapitel verwendet, um zwei Methoden zur Bildvorverarbeitung zu entwerfen, die in Abschnitt 5.4 miteinander verglichen werden.

4.2 Zelldetektion

Die im weiteren vorzustellenden Verfahren ziehen Information zu Komposition und Probendicke aus Bildbereichen, die der Größe der Elementarzelle entsprechen. Ein HREM-Bild wird daher in einem ersten Schritt in solche Bereiche unterteilt.

Bei dem hier entwickelten Verfahren wählt man interaktiv ein Helligkeitsmaximum (*blob*) und jeweils ein horizontal und ein vertikal benachbartes Maximum. Das Verfahren sucht die drei Maxima mit Hilfe der Schwerpunktmethod. Eine so erhaltene Position für ein Maximum wird durch einen Fit mit einem elliptischen Paraboloid zweiter

Ordnung weiter präzisiert. Der Fit wird mit einem elliptischen Paraboloid durchgeführt, obwohl die Form der Helligkeitsmaxima (200- bzw. 220-Fünfstrahlfall) in zweiter Ordnung entwickelt ein Rotationsparaboloid darstellt. Der Grund hierfür ist, dass sich bei dem elliptischen Paraboloid im Gegensatz zum Rotationsparaboloid die Position des Paraboloids analytisch optimieren lässt. Dadurch werden Schwierigkeiten mit iterativen numerischen Methoden (Konvergenzprobleme, hoher Zeitbedarf) vermieden. Aus den so präzisierten Positionen der drei eingegebenen Maxima werden erste Werte für die horizontale und vertikale Gitterkonstante berechnet.

Das in der Aufnahme zu bestimmende Gitter wird als Baumstruktur (dynamische Speicherstrukturen siehe [53]) gespeichert. So ist auch die Erfassung von verzerrten Gittern und Versetzungen möglich. Jeder Knoten des Baumes (Gitterpunkt) kennt dabei seine Koordinaten und seine nächsten Nachbarn. Daneben ist noch ein Parameter gespeichert, der die Zuverlässigkeit der Position angibt (1 entspricht sicher, 0 entspricht unbestimmt).

Das Gitter wird nun an seinem Rand erweitert, bis entweder der Rand des Bildes oder eine vorgegebene Grenze für die Auswertung erreicht ist. Für die Erweiterung des Gitters werden alle schon vorhandenen sicheren Gitterpositionen ausgewertet. Die Abstandsinformation der weiter entfernt liegenden Positionen kann über einen vorgegebenden Parameter exponentiell gedämpft werden. Durch diesen Parameter kann die Steifigkeit des Gitters reguliert werden. Für sehr verspannte Proben ist das Gitter flexibler zu wählen als im Fall der vorliegenden AlGaAs Strukturen.

Der Parameter, der die Sicherheit einer Position angibt, wird wie folgt gewählt: Interaktiv gewählte Positionen gelten als sicher. Minima werden als unsicher angesehen. Die Präzisierung von Minima durch Schwerpunktmethode und/oder Paraboloidfit hat sich als nicht sinnvoll herausgestellt. Maxima deren Position eine im Vergleich zu einem Achtel der Elementarzelle kleine Korrektur erfordert, werden als sicher betrachtet.

Für wenig verspannte Strukturen kann auch ein optimiertes starres Gitter (*rigid grid*) aus den vorhandenen Positionen berechnet werden. Bei diesem können interaktiv Bereiche, deren Maxima als zuverlässig gelten, ausgewählt werden. Anhand dieser Bereiche kann für alle anderen (unsicheren) Positionen die Gittersuche wie oben beschrieben erneut erfolgen.

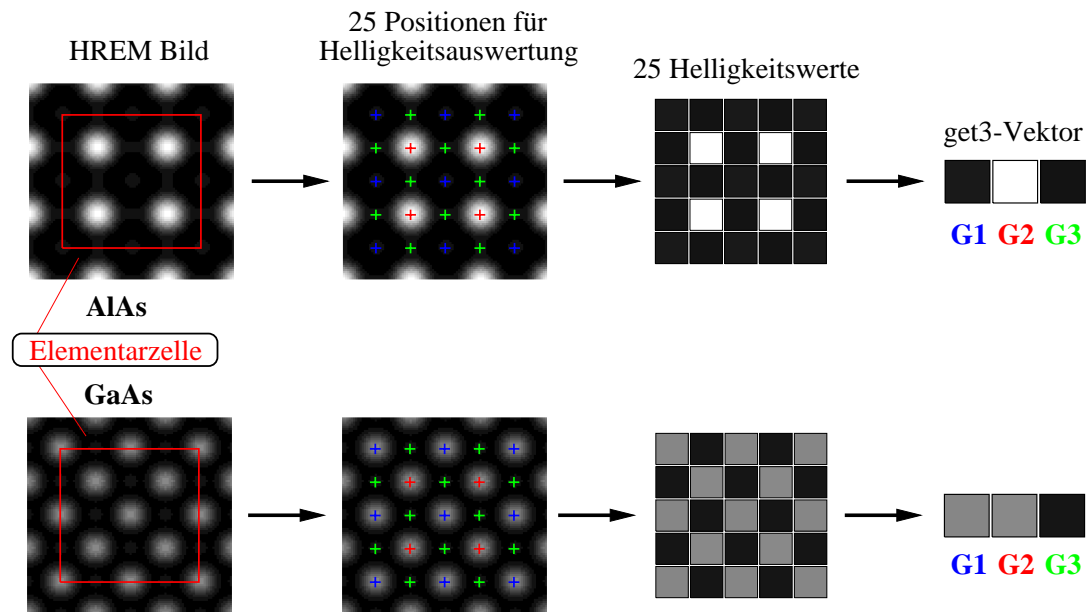


Abbildung 4.1: Bildvorverarbeitung für zwei simulierte Beispiele: AlAs und GaAs.

4.3 Realraumbasierte Vorverarbeitung: get3

Im weiteren wird eine im Rahmen dieser Arbeit entwickelte Bildvorverarbeitung für TEM Hochauflösungsaufnahmen von Materialien mit Sphaleritstruktur in (001)-Projektion vorgestellt.

4.3.1 Beschreibung

Wie in Abschnitt 4.2 dargestellt, wird das Gesamtbild in Bereiche unterteilt, welche der Größe der Elementarzelle entsprechen. In Abbildung 4.1 ist die Vorverarbeitung für simulierte HREM-Bilder von AlAs und GaAs skizziert. Die erste Spalte der Abbildung zeigt die simulierten HREM-Bilder. Der einer Elementarzelle entsprechende Bildbereich ist jeweils rot markiert.

Zunächst werden auf dem Bildbereich einer Elementarzelle 25 Positionen bestimmt, an denen sich Helligkeitsmaxima befinden, oder aufgrund der Kristallsymmetrie und des Abbildungsvorgangs prinzipiell möglich wären. In der zweiten Spalte von Abb. 4.1 sind die 25 Positionen jeweils markiert.

An diesen Positionen wird die Intensität durch Anfitzen eines Rotationsparaboloids 4. Ordnung bestimmt. Zur Helligkeitsbestimmung wird hier im Gegensatz zur Positionsbestimmung ein Rotationsparaboloid zum Fit verwendet. Dies ermöglicht durch

1	3	1	3	1
3	2	3	2	3
1	3	1	3	1
3	2	3	2	3
1	3	1	3	1

Abbildung 4.2: Letzter Schritt der Bildvorverarbeitung: Die Zahlen repräsentieren äquivalente Positionen für die get3-Mittelung.

seine höhere numerische Stabilität, auch in Helligkeitsminima die Intensität durch Fit zu bestimmen. Tests mit Pixelrauschen ergaben einen Vorteil des Fits mit einem Paraboloid vierter Ordnung gegenüber dem Fit mit zweiter Ordnung. Die dritte Spalte von Abbildung 4.1 zeigt die 25 Helligkeitswerte in einer Graustufendarstellung. Die Fläche der Graustufendarstellung ist etwas größer als die Elementarzelle. Für die Bestimmung der Helligkeit an den Positionen am Rand der Elementarzelle werden auch Bildpunkte ausserhalb des Bereichs der Elementarzelle verwendet (laterale Auflösung siehe Abschnitte 4.3.3 und 8.2).

Die 25 Positionen kann man aufgrund der Kristallsymmetrie zu drei Gruppen zusammenfassen (s. Kap. 4.3.2). Abbildung 4.2 repräsentiert diese Gruppen (G1, G2, G3) durch die Zahlen eins bis drei. Wir bilden den Mittelwert innerhalb jeder der drei Gruppen und erhalten einen dreidimensionalen Vektor. Im folgenden wird dieser Vektor als get3-Vektor bezeichnet. Die rechte Spalte von Abbildung 4.1 zeigt get3-Vektoren für die zwei simulierten Beispiele in Graustufendarstellung.

4.3.2 Motivation für get3

Motivation für get3: Vergleich Gitter – Bild

Für ein ideales Gitter und spezielle (idealisierte) Abbildungsbedingungen kann man Kristallgitter und Bild direkt miteinander vergleichen. Den beiden Untergittern der Sphaleritstruktur entsprechen die Gruppen G1 und G2. Die Gruppe G3 entspricht den Zwischengitterpositionen. Die Kompositionsinformation wird sich vor allem in den Werten der Gruppen G1 und G2 widerspiegeln. Um später in der Lage zu sein, Kompositionsvariation von Variation in der Probendicke zu unterscheiden, wird hier die Gruppe G3 auch mit ausgewertet.

Motivation für get3: Gittersymmetrie

Die Sphaleritstruktur hat die Raumgruppe $F\bar{4}3m$. Nach dem Neumannschen Prinzip muss die (äußere) Symmetriegruppe die Symmetriegruppe des Kristalls enthalten [29].

Gehn wir von einer Abbildungssituation ohne Astigmatismus und ohne Strahlverkipfung aus, so sind die elektromagnetischen Felder des Mikroskops orthogonal zur Strahlrichtung rotationssymmetrisch. Daher führt der Abbildungsprozess zu keiner Verringerung der Punktsymmetrie. Die Abbildung hat mindestens die Punktsymmetrie des Kristalls.

Der Kristall hat in jedem Fall die Translationssymmetrie $T(a)$ also Invarianz bei Verschiebung um die Gitterkonstante a entlang der Kristallachsen. Betrachten wir die dynamische Beugung, welche mit dem *multislice*-Verfahren berechnet wird. Die Phasengitterfunktionen (2.9), die aus der Kristallstruktur berechnet werden, haben ebenfalls die Translationssymmetrie $T(a)$. Die Phasenverschiebung der Wellenfunktion erfolgt daher auch mit dieser Translationssymmetrie. Die resultierende Wellenfunktion ist somit ebenfalls translationssymmetrisch. Die Faltung mit der Propagatorfunktion erhält diese Translationssymmetrie (Eine Faltung kann die Translationssymmetrie in keinem Fall reduzieren). Die mit dem *multislice*-Zyklus berechnete Austrittswellenfunktion ist also ebenfalls invariant gegenüber Verschiebung um den einer Elementarzelle entsprechenden Abstand. Wie in Abschnitt 2.2 dargestellt, setzt sich die Translationssymmetrie der Austrittswellenfunktion auch in der HREM-Abbildung fort.

Abbildung 4.3 zeigt eine Sphaleritstruktur in (001)-Projektion. Die beiden Untergitter sind jeweils durch schwarze und weiße Kreise dargestellt. Betrachten wir das schwarze Untergitter, so liegen zwei unterschiedliche Sorten von Atomsäulen vor. Die eine Atomsäule (kleine schwarze Kreise) ist gegenüber der anderen Atomsäule (große schwarze Kreise) um die Hälfte der Gitterkonstanten in Projektionsrichtung verschoben. In Abb. 4.3 sind drei Bildfrequenzen gezeichnet (110-, 200-, 220-Bildfrequenz; Funktionswerte als Graustufendarstellung mit Höhenlinien). Einer Unterscheidung der beiden Sorten von Atomsäulen eines Untergitters (Unterscheide kleine von großen schwarzen Kreisen) entspricht der 110-Bildfrequenz. Dies lässt sich in Abbildung 4.3 durch Vergleich der Bildfrequenz mit dem Modell der Sphaleritstruktur verifizieren.

Die Sphaleritstruktur hat in (001)-Projektion nur geradzahlig indizierte Reflexe (Auslöschungsregel für flächenzentrierte Gitter). Aus diesen geradzahlig indizierten Reflexen kann auch durch die Autokorrelation der Intensitätsbildung keine 110-Bildfrequenz entstehen. Durch das Fehlen der 110-Bildfrequenz können die beiden Typen von Atomsäulen

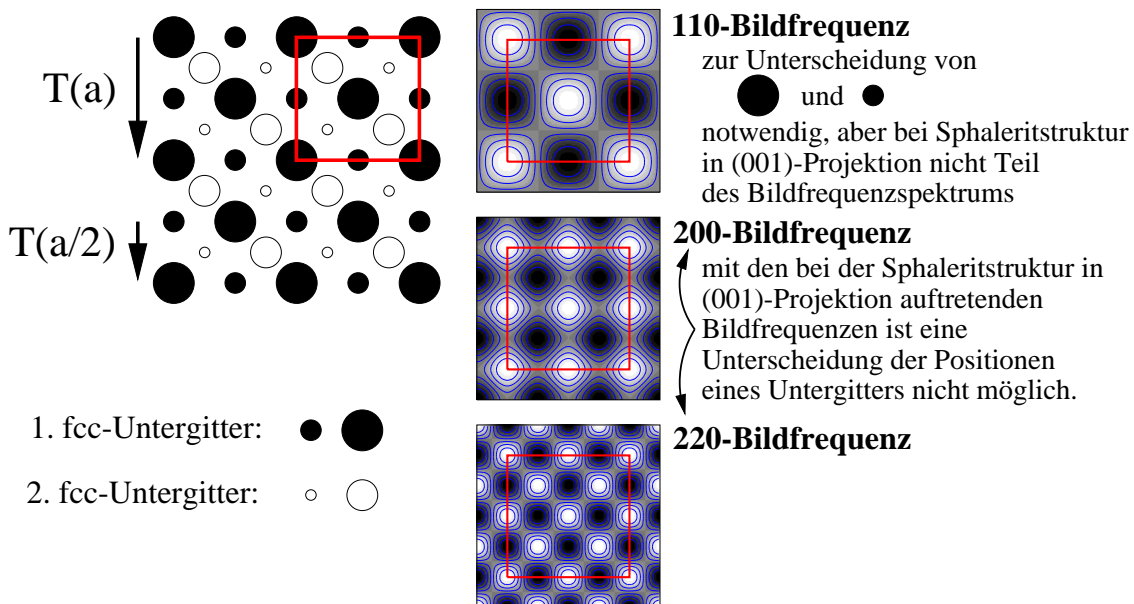


Abbildung 4.3: Sphaleritstruktur in (001)-Projektion und 110-, 200-, 220-Bildfrequenz. Für die Bildfrequenzen ist die Graustufendarstellung der Werte mit Höhenlinien (blau) unterlegt. Am Strukturmodell ist die Translationssymmetrie $T(a)$ und die zusätzliche Translationssymmetrie $T(\frac{a}{2})$ des HREM-Bildes dargestellt.

im Bild nicht unterschieden werden. In der Abbildung tritt eine zusätzliche Translationssymmetrie von der Hälfte der Gitterkonstanten auf ($T(\frac{A}{2})$).

Betrachten wir eine Fläche der Abbildung, die der Größe einer Elementarzelle entspricht. Jede der beiden Koordinatenachsen soll durch N Unterteilungen äquidistant aufgeteilt werden. Die Aufteilung erzeugt Ausschnitte der Abbildung in denen die Helligkeit bestimmt wird (Der genaue Kontrastverlauf innerhalb eines solchen Ausschnitts wird dabei vernachlässigt).

Der Gitterparameter (in der Abbildung) sei A . Die Zelle hat also eine Ausdehnung von $0 \dots A$ in beiden Koordinatenrichtungen. Die erste Unterteilung ist bei $\frac{A}{N}$.

Die Translationssymmetrie $T(\frac{A}{2})$ erzeugt eine weitere Unterteilung bei $\frac{A}{N} + \frac{A}{2}$. Diese muss mit einer beliebigen Unterteilung (Nummer x) zusammenfallen. Die Unterteilung der Nummer x ist bei $x\frac{A}{N}$ daher gilt wegen Translationssymmetrie:

$$\frac{A}{N} + \frac{A}{2} \stackrel{!}{=} x \frac{A}{N} \Leftrightarrow x \stackrel{!}{=} \frac{N}{2} \tag{4.1}$$

Da x eine ganze Zahl ist, muss die Zahl der Unterteilungen N eine gerade Zahl sein.

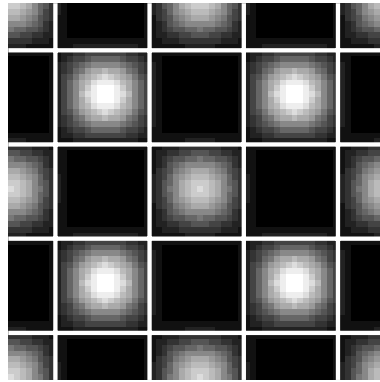


Abbildung 4.4: Die abgebildete Elementarzelle wird entlang jeder Koordinatenachse vierfach unterteilt. Dies führt zu 25 Positionen an denen die Helligkeit ausgewertet werden muss.

Sind die beiden Untergitter der Elementarzelle mit Atomen (etwa) gleichen Streuvermögens besetzt, so erhöht sich die Symmetrie weiter. Bezüglich der Streueigenschaft auf Elektronen läge dann eine Diamantstruktur vor. Bei dieser Struktur sind die 200-Reflexe verboten (Auslöschungsregel). Dadurch können die beiden Untergitter der Sphaleritstruktur auch in der Abbildung nicht unterschieden werden. Es kommt eine Invarianz gegenüber Verschiebungen um

$$\left(\frac{A}{4}, \frac{A}{4}\right) \quad (4.2)$$

hinzu (Translationssymmetrie $T(\frac{A}{4}, \frac{A}{4})$).

Die Unterteilung des Bildbereichs einer Elementarzelle muss auch mit dieser Symmetrie verträglich sein. Sonst würde bei hochsymmetrischen Strukturen (z.B. GaAs) über nichtäquivalente Bereiche gemittelt.

Aus der ersten Unterteilung bei $\frac{A}{N}$ wird durch die Translationssymmetrie $T(\frac{A}{4}, \frac{A}{4})$ eine weitere Unterteilung bei $\frac{A}{N} + \frac{A}{4}$. Wiederum muss diese mit einer beliebigen Unterteilung (Nummer x) zusammenfallen. Daher gilt:

$$\begin{aligned} \frac{A}{N} + \frac{A}{4} &\stackrel{!}{=} x \frac{A}{N} \Leftrightarrow \\ x &\stackrel{!}{=} \frac{N}{4} \end{aligned} \quad (4.3)$$

Da x eine ganze Zahl ist, muss die Zahl der Unterteilungen N ein Vielfaches von Vier sein.

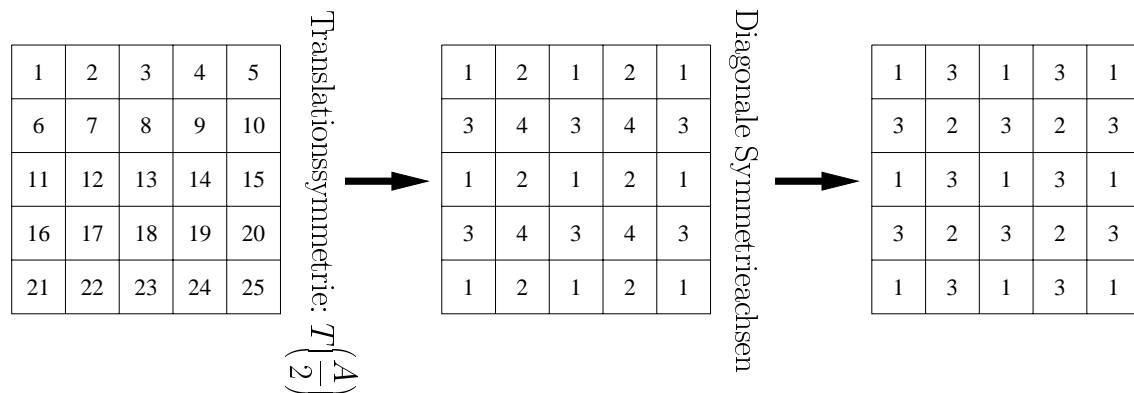


Abbildung 4.5: Reduktion der 25 Helligkeitspositionen auf die drei get3-Gruppen.

Die Punktauflösung des verwendeten Mikroskops ist 1.7 \AA . Die größte mögliche Unterteilung einer projizierten Elementarzelle ist $N = 4$. Hier ist der Abstand zwischen zwei Unterteilungen $d = 1.4 \text{ \AA}$. Die nächstfeinere Unterteilung wäre $N = 8$. Hier ist $d = 0.7 \text{ \AA}$. Eine feinere Auflösung als $N = 4$ ist offenbar nicht sinnvoll.

Abbildung 4.4 zeigt eine simulierte Abbildung einer GaAs Elementarzelle. Eingezeichnet ist die vierfache Unterteilung entlang der beiden Koordinatenachsen. Diese führt zu 25 Positionen an denen die Intensität bestimmt wird.

Im linken Teil der Abbildung 4.5 sind die 25 Positionen der Helligkeitsbestimmung dargestellt. Die Translationssymmetrie $T\left(\frac{A}{2}\right)$ reduziert die Positionen auf vier verschiedene Gruppen (Abb. 4.5 Mitte). Die Berücksichtigung der Punktsymmetrie reduziert die Positionen weiter auf drei verschiedene Gruppen (Abb. 4.5 Rechts). Es ergibt sich die schon in Abb. 4.2 dargestellte get3-Gruppeneinteilung.

Motivation für get3: Beteiligte Reflexe

Das Auftreten von chemisch sensitiven Reflexen im HREM-Beugungsbild kann über die Gittersymmetrie begründet werden. Die atomare Streuamplitude fällt für große Beugungswinkel stark ab, daher soll es hier genügen, Raumfrequenzen bis zur 220-Frequenz zu betrachten. Die Auslöschungsgesetze für flächenzentrierte Gitter erlauben in diesem Bereich nur 200- und 220-Reflexe.

Unterscheiden sich die beiden Untergitter der Sphaleritstruktur nur sehr wenig in ihrem Streuvermögen, so reduziert sich die Struktur bezüglich dieser Eigenschaft auf eine Diamantstruktur $Fd\bar{3}m$. Die Auslöschungsgesetze dieser Struktur verbieten auch die 200-Reflexe. Deshalb wird in (100)-Projektion nur 220-basierter Kontrast abgebildet. Anhang B berechnet für verschiedene Materialsysteme die Komposition mit

der höchsten Symmetrie (minimale 200-Reflexintensität). Abbildung 4.6 zeigt die 200-Reflexintensität für zwei Materialsysteme (simulierte Probe, Dicke 100\AA).

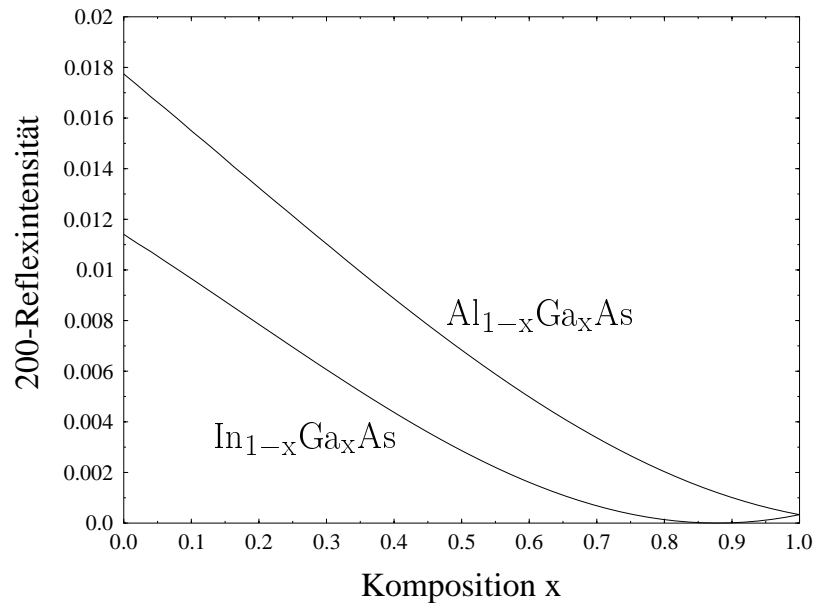


Abbildung 4.6: Intensität der 200-Reflexe in Abhängigkeit der Komposition für zwei Materialsysteme. Die simulierte Probe hatte eine Dicke von 100\AA .

Im Materialsystem $\text{Al}_{1-x}\text{Ga}_x\text{As}$ wird der minimale Unterschied im Streuvermögen der beiden Untergitter bei der Komposition $x=1$ (GaAs) erreicht. Daher ist die 200-Reflexintensität beim GaAs am schwächsten. Die 200-Reflexintensität verändert sich bei $\text{Al}_{1-x}\text{Ga}_x\text{As}$ monoton mit der Komposition x . Für andere Materialsysteme ist das nicht gegeben. Zum Beispiel liegt bei $\text{In}_{1-x}\text{Ga}_x\text{As}$ das Minimum der 200-Reflexintensität bei $x=0.86$. Der Zusammenhang zwischen 200-Reflexintensität und Komposition ist bei diesem Materialsystem nicht monoton. Dies führt bei ausschliesslicher Auswertung der 200-Reflexintensität zu großen systematischen Fehlern (siehe Anhang B).

Zwischen der vorgestellten get3 Vorverarbeitung und den einzelnen an der Abbildung beteiligten Reflexen besteht kein direkter Zusammenhang. Die einzelnen Reflexe beeinflussen stets den gesamten Bildbereich, während die Auswertung mit get3 im Realraum streng lokalisiert ist.

Trotzdem lassen sich natürlich über die gemeinsame Ursache (Kompositionsänderung) einer Signalveränderung Zusammenhänge aufzeigen. Ein Ansteigen der Intensität der 200-Reflexe ist mit einem Ansteigen des Wertes in einer der beiden Gruppen G1 oder G2 und jeweils mit einer Verringerung des Wertes in der anderen Gruppe verbunden. Welcher der beiden Gruppenwerte steigt und welcher fällt, hängt von der Phasenlage der Reflexe zum Nullstrahl ab. Eine Steigerung der 220-Intensität verändert G1 und

a	b	c	d																																																																																																				
<table border="1" style="border-collapse: collapse; width: 100%; text-align: center;"> <tr><td>1</td><td>3</td><td>1</td><td>3</td><td>1</td></tr> <tr><td>3</td><td>2</td><td>3</td><td>2</td><td>3</td></tr> <tr><td>1</td><td>3</td><td>1</td><td>3</td><td>1</td></tr> <tr><td>3</td><td>2</td><td>3</td><td>2</td><td>3</td></tr> <tr><td>1</td><td>3</td><td>1</td><td>3</td><td>1</td></tr> </table>	1	3	1	3	1	3	2	3	2	3	1	3	1	3	1	3	2	3	2	3	1	3	1	3	1	<table border="1" style="border-collapse: collapse; width: 100%; text-align: center;"> <tr><td>2</td><td>3</td><td>2</td><td>3</td><td>2</td></tr> <tr><td>3</td><td>1</td><td>3</td><td>1</td><td>3</td></tr> <tr><td>2</td><td>3</td><td>2</td><td>3</td><td>2</td></tr> <tr><td>3</td><td>1</td><td>3</td><td>1</td><td>3</td></tr> <tr><td>2</td><td>3</td><td>2</td><td>3</td><td>2</td></tr> </table>	2	3	2	3	2	3	1	3	1	3	2	3	2	3	2	3	1	3	1	3	2	3	2	3	2	<table border="1" style="border-collapse: collapse; width: 100%; text-align: center;"> <tr><td>3</td><td>2</td><td>3</td><td>2</td><td>3</td></tr> <tr><td>1</td><td>3</td><td>1</td><td>3</td><td>1</td></tr> <tr><td>3</td><td>2</td><td>3</td><td>2</td><td>3</td></tr> <tr><td>1</td><td>3</td><td>1</td><td>3</td><td>1</td></tr> <tr><td>3</td><td>2</td><td>3</td><td>2</td><td>3</td></tr> </table>	3	2	3	2	3	1	3	1	3	1	3	2	3	2	3	1	3	1	3	1	3	2	3	2	3	<table border="1" style="border-collapse: collapse; width: 100%; text-align: center;"> <tr><td>3</td><td>1</td><td>3</td><td>1</td><td>3</td></tr> <tr><td>2</td><td>3</td><td>2</td><td>3</td><td>2</td></tr> <tr><td>3</td><td>1</td><td>3</td><td>1</td><td>3</td></tr> <tr><td>2</td><td>3</td><td>2</td><td>3</td><td>2</td></tr> <tr><td>3</td><td>1</td><td>3</td><td>1</td><td>3</td></tr> </table>	3	1	3	1	3	2	3	2	3	2	3	1	3	1	3	2	3	2	3	2	3	1	3	1	3
1	3	1	3	1																																																																																																			
3	2	3	2	3																																																																																																			
1	3	1	3	1																																																																																																			
3	2	3	2	3																																																																																																			
1	3	1	3	1																																																																																																			
2	3	2	3	2																																																																																																			
3	1	3	1	3																																																																																																			
2	3	2	3	2																																																																																																			
3	1	3	1	3																																																																																																			
2	3	2	3	2																																																																																																			
3	2	3	2	3																																																																																																			
1	3	1	3	1																																																																																																			
3	2	3	2	3																																																																																																			
1	3	1	3	1																																																																																																			
3	2	3	2	3																																																																																																			
3	1	3	1	3																																																																																																			
2	3	2	3	2																																																																																																			
3	1	3	1	3																																																																																																			
2	3	2	3	2																																																																																																			
3	1	3	1	3																																																																																																			

Abbildung 4.7: Vier gleichwertige Möglichkeiten für get3-Einteilung.

G2 parallel. Ob eine Steigerung der 220-Intensität einen Anstieg oder ein Abfallen der Werte von G1 und G2 zur Folge hat, ist von der Phasenlage der Reflexe zum Nullstrahl abhängig. Die Gruppe G3 verändert sich unter dem Einfluss der 220-Intensität gerade umgekehrt wie G1 und G2.

4.3.3 Diskussion von get3

Mehrdeutigkeit bei der get3-Gruppeneinteilung

Die Gruppeneinteilung get3 ist bei simulierten Mustern eindeutig. Bei experimentellen Mustern ist diese Eindeutigkeit nicht mehr gegeben. Insgesamt gibt es vier Möglichkeiten die Gruppeneinteilung unter den 25 Positionen durchzuführen. Diese sind in Abb. 4.7 zusammengefasst.

Bei den praxisrelevanten Abbildungssituationen können wir pathologische Fälle vernachlässigen, bei denen im gesamten ausgewerteten Bild keine 200-Bildfrequenz auftritt. Betrachtet man einen Bildausschnitt, der der Größe der Elementarzelle entspricht und der die 200-Bildfrequenz enthält, so ist die Wahl der Gruppe G3 eindeutig. Von den zwei Möglichkeiten die Positionen von G3 zu wählen, wird diejenige verwendet, welche die minimale Varianz innerhalb der Gruppenpositionen G3 aufweist. Also kann in Abb. 4.7 unterschieden werden, ob entweder einer der Fälle a,b oder einer der Fälle c,d vorliegt. Dabei setzt sich die an einer Zelle durchgeführte Gruppeneinteilung auf das gesamte Bild fort. Tritt im Bild die 200-Bildfrequenz auf, so unterscheiden sich die Werte von G1 und G2. Vergleicht man Simulation mit Experiment, so kann aufgrund dieser Verschiedenheit eine Gruppeneinteilung eindeutig erfolgen.

Laterale Auflösung

Die Mittelung über die Positionen innerhalb der jeweiligen Gruppe erfolgt bei get3 ohne Gewichtung. Möchte man in der Auflösung des Verfahrens die Größe der Ele-

mentarzelle erreichen, so ist die Mittelung mit den entsprechenden Wichtungsfaktoren durchzuführen:

Betrachtet man den Wert zu einer Position als Kompositionsinformation einer Atomsäule (bzw. eines Zwischenplatzes im Gitter), so gewichtet man den Wert entsprechend dem Anteil der Atomsäule an der Elementarzelle:

$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$
$\frac{1}{2}$	1	1	1	$\frac{1}{2}$
$\frac{1}{2}$	1	1	1	$\frac{1}{2}$
$\frac{1}{2}$	1	1	1	$\frac{1}{2}$
$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$

Die Argumentation gilt natürlich nur *cum grano salis*. Eine korrekte Gewichtung berücksichtigt z.B. die Raumfrequenzfilterung der Kontrastübertragung und damit die Korrelation in Abhängigkeit des Abstandes der Gruppenpositionen. Hierfür wären die exakten Abbildungsparameter erforderlich, welche von vornherein nicht bekannt sind.

Bei der gewichteten Mittelung erhöht sich die laterale Auflösung auf Kosten des Signal-Rausch-Verhältnis. Bei dem in Kapitel 8 ausgewerteten HREM-Aufnahmen treten bei der Kompositionsbestimmung (Vorverarbeitung: ungewichtete get3-Mittelung) *worst-case*-Fehler von bis zu 7 Prozent auf. Versuche mit gewichteter get3-Mittelung ergaben Fehler über 10 Prozent (die laterale Einschränkung wirkt sich auf tiefpassgefiltertes Rauschen, im Vergleich zu unkorreliertem Rauschen, drastischer aus). Für HREM-Aufnahmen der hier untersuchten Amorphisierungstiefe (s. Kap. 3) und Dickenvariation (s. Kap. 8) ist daher die gewichtete get3-Mittelung für eine quantitative Auswertung nicht sinnvoll.

Die laterale Auflösung von get3 ist abhängig von den Abbildungsbedingungen des Mikroskops. In Abschnitt 8.2 wird die laterale Auflösung der Kompositionsbestimmung (Vorverarbeitung: ungewichtete get3-Mittelung) für eine experimentelle Anwendung untersucht. Das Ergebnis 5.9\AA ist wie erwartet etwas größer als die GaAlAs Elementarzelle (5.7\AA).

Vorteile von get3

Wie in Abschnitt 1.2 aufgezeigt, lassen sich die derzeit bekannten Verfahren zur Kompositionsbestimmung aus HREM-Aufnahmen in zwei Gruppen unterteilen: realraumba-

sierte Verfahren und Fourierverfahren. In diesem Abschnitt wird die get3-Vorverarbeitung zunächst mit anderen realraumbasierten Verfahren verglichen. Der Vergleich zu fourierbasierten Verfahren wird später in Abschnitt 5.4 untersucht.

Im Gegensatz zur get3-Vorverarbeitung, betrachten konventionelle realraumbasierte Verfahren die Bildinformation als hochdimensionalen (Pixel-) Vektor. In Abbildung 4.8 ist die Umsetzung der zweidimensionalen Bildinformation in den eindimensionalen Pixelvektor anhand eines simulierten HREM-Bildes (220-Fünfstrahlfall, dargestellt ist ein Viertel der Elementarzelle) veranschaulicht. Bei solchen Verfahren wird die Information über die zweidimensionalen Abstandsverhältnisse stark verschlüsselt. Benachbarte Punkte im Bild werden im Pixelvektor getrennt und auf weite Bereiche von Vektor-komponenten verteilt. Die Information über die Abstandsverhältnisse im Bild ist so für die weitere Verarbeitung nicht mehr zugänglich.

zweidimensionales

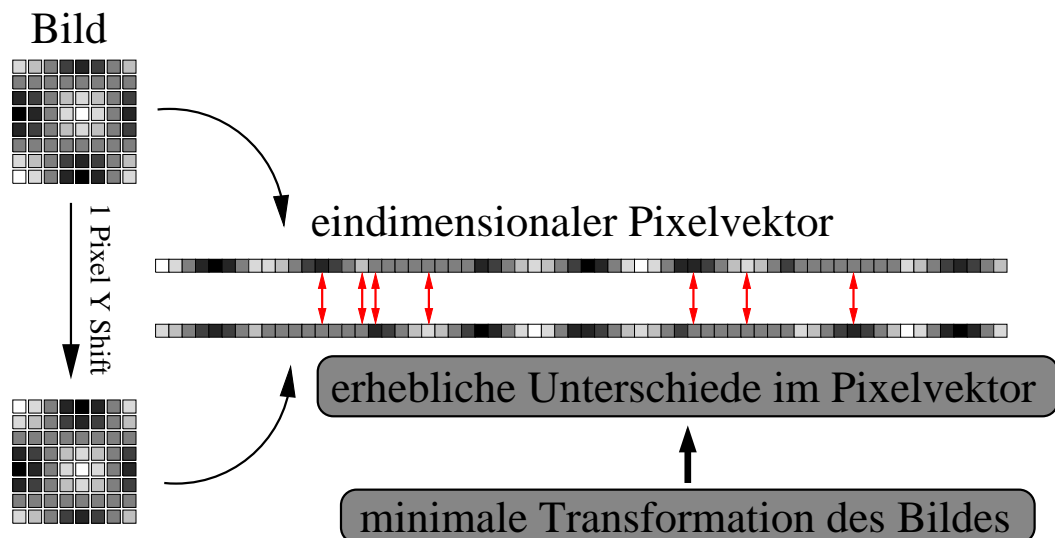


Abbildung 4.8: Nachteil einer Umsetzung der Bildinformation in einen eindimensionalen Pixelvektor. Aufgrund des fehlenden Bezugs zur Bildtopologie, führen minimale Transformationen des Bildinhalts zu erheblichen Veränderungen des Pixelvektors.

Wie in Abb. 4.8 dargestellt, führt schon eine Translation des Bildausschnittes um einen Pixel zu erheblichen Veränderungen im Pixelvektor. Solche minimalen Translationen lassen sich bei der Auswertung experimenteller HREM-Abbildungen nicht vermeiden. Die Aufteilung des Bildes in Pixel steht bei experimentellen Abbildungen in keinem Bezug zur Translationssymmetrie im Bild. Daher kann auch bei idealer, fehlerfreier Zelldetektion der Bildinhalt verschiedener Bildzellen um bis zu einen Pixel translatiert sein.

Im Vergleich dazu ist eine Translation des Bildinhaltes für die get3-Vorverarbeitung

kein Problem. Die Schwerpunktmethode findet die Orte der Helligkeitsmaxima bei den untersuchten Bildauflösungen wenigstens bis auf zwei Pixel Genauigkeit. Daraus berechnet der Fit mit elliptischem Paraboloid die Position auf Gleitkommagenauigkeit. Die reellwertige Position für die abschließende Helligkeitsbestimmung (Fit mit Rotationsparaboloid) ist somit von der Translation des Bildinhaltes unabhängig. Für die Durchführung der Paraboloidfits werden kreisförmige Bildbereiche um die Positionen der Helligkeitsauswertung verwendet. Das Ergebnis der Helligkeitsbestimmung ist unabhängig von einer quadratischen oder rechteckigen Aufteilung der Bilddaten. Eine beliebige Rotation des Bildinhaltes ändert das Ergebnis der Helligkeitsauswertung nicht, weil das Koordinatensystem des elliptischen Paraboloids beim Fit mitrotiert. Das Bild braucht für die Auswertung also nicht vororientiert werden.

Um bei den Paraboloidfits ausreichende numerische Stabilität zu erreichen, ist eine Bildauflösung von wenigstens 20×20 Pixel pro Elementarzelle notwendig¹. Ist dieses Kriterium erfüllt, so verbessert eine Erhöhung der Auflösung nur die Stabilität gegenüber Rauschen, verändert aber nicht den Wert oder die Dimension des Resultats der Vorverarbeitung.

4.4 Vorverarbeitung: lokale Fouriertransformierte

4.4.1 Beschreibung

Eine zweite, während dieser Arbeit detailliert ausgearbeitete Bildvorverarbeitung, ist die lokale Fouriertransformierte. Diese Methode analysiert einen Bildausschnitt welcher der Größe der Elementarzelle entspricht. Der Bildausschnitt wird fouriertransformiert und die Intensitäten der 200- und 220-Bildfrequenzen bestimmt. Man überlagert die Intensitäten I (inkohärent) und erhält so aus einem Bildausschnitt zwei Parameterwerte F_{200} und F_{220} :

$$F_{200} = I_{200} + I_{\bar{2}00} + I_{020} + I_{0\bar{2}0} \quad (4.4)$$

$$F_{220} = I_{220} + I_{\bar{2}\bar{2}0} + I_{2\bar{2}0} + I_{\bar{2}20} \quad (4.5)$$

¹Der optimale Radius der Auswertebereiche wurde als $\frac{1}{10}$ Gitterkonstante bestimmt. Wird die Grenze für die Auflösung des Bildes eingehalten, so stehen für jeden Fit mindestens 12 Werte (Pixel) zur Verfügung.

4.4.2 Diskussion der lokalen Fouriertransformierten

Bei HREM-Aufnahmen von Sphaleritstrukturen in (100)-Projektion sind an der Bildentstehung vor allem 200- und 220-Reflexe beteiligt (s. Kap. 4.3.2). Ebenfalls in Abschnitt 4.3.2 wurde die chemische Sensitivität der 200-Reflexe, d.h. die Abhängigkeit der Reflexintensität von Unterschieden im Streuverhalten der beiden Untergitter, begründet. Die 220-Reflexintensität wird indirekt ebenfalls durch die Komposition beeinflusst.

Die hier beschriebene Bildvorverarbeitung wertet die Fouriertransformierte einer Intensitätsverteilung (HREM-Bild) aus. Der Intensitätsbildung im Realraum entspricht eine Autokorrelation im reziproken Raum. Daher sind auch beliebige Vielfache und Kombinationen der 200- und 220-Objektfrequenzen durch nichtlineare Bildentstehung im Spektrum des Bildes vorhanden. So wird zum Beispiel ein Teil der Intensität von F_{220} durch nichtlineare Überlagerung von 200- und 020-Reflexen verursacht. Die Information aus den 200- und 220-Reflexen wird von den beiden Parametern der lokalen Fouriertransformierten vermischt wiedergegeben.

Die Abbildung der Austrittswellenfunktion durch das Elektronenmikroskop entspricht einer Raumfrequenzfilterung (vgl. Abschnitt 2.2). Diese Tatsache verändert im allgemeinen das Verhältniss der beiden Werte F_{200} und F_{220} . Trotz dieser Verfälschung der Reflexinformation bei der Abbildung, setzt sich die chemische Sensitivität der 200- und 220-Reflexe in der Intensität der 200- und 220-Bildfrequenzen fort. Die Parameter F_{200} und F_{220} ändern ihre Werte in Abhängigkeit von der Komposition.

Ausgewertet werden bei der hier vorgestellten Vorverarbeitung sowohl 200- als auch 220-Bildfrequenzen. Dies ist wegen der indirekten chemischen Sensitivität des 220-Reflexes auch sinnvoll. Wird nur die 200-Bildfrequenz ausgewertet, so kommt es in Materialsystemen wie z.B. InGaAs zu Fehlern bei der Kompositionsbestimmung (siehe Anhang B). Zudem sollen im Verlauf der Arbeit Schwankungen der kristallinen Probendicke von Kompositionsveränderungen unterschieden werden. Dies ist mit einer Bildvorverarbeitung, die ein eindimensionales Ergebnis liefert, nicht möglich.

Es werden bei beiden Reflexgruppen der lokalen Fouriertransformierten nur die Fourierintensitäten ausgewertet. Die zusätzlich vorhandene Phaseninformation wird nicht berücksichtigt. Die Phasen bei der lokalen Fouriertransformierten werden vor allem durch die translative Position des ausgewerteten Bereichs zum abgebildeten Gitter bestimmt. Wie in Abschnitt 4.3.3 dargestellt, muss man bei den verschiedenen ausgewerteten Bildbereichen von experimentellen Abbildungen mit einer Translation des Bildinhalts von wenigstens einem Pixel rechnen. Dies entspricht bei einer Auflösung

von 40 Pixel pro Elementarzelle einem Fehler der 200-Bildfrequenzphase von 10 Prozent (bezogen auf die maximale absolute Phasendifferenz π). Zu diesem Fehler addiert sich noch der Fehlerbeitrag vom Rauschen und von der endlichen Genauigkeit bei der Zelldetektion. Die Phase der Bildfrequenzen ist bei experimentellen Auswertungen stark fehlerbehaftet und wird daher bei der Bildvorverarbeitung nicht berücksichtigt.

Bei der lokalen Fouriertransformierten werden die Amplituden der einzelnen Fourierkoeffizienten (A_{hk0}) inkohärent überlagert:

$$F_{200} = |A_{200}|^2 + |A_{\bar{2}00}|^2 + |A_{020}|^2 + |A_{0\bar{2}0}|^2 \quad (4.6)$$

$$F_{\bar{2}20} = |A_{220}|^2 + |A_{\bar{2}\bar{2}0}|^2 + |A_{2\bar{2}0}|^2 + |A_{\bar{2}20}|^2 \quad (4.7)$$

Eine kohärente Überlagerung, also zum Beispiel

$$K_{200} = |A_{200} + A_{\bar{2}00} + A_{020} + A_{0\bar{2}0}|^2 \quad (4.8)$$

stellt nur einen Spezialfall einer Phasenauswertung dar. Dies wird klar, wenn man bedenkt, dass wegen der reellen Bildinformation (Intensitäten) für die Fouriertransformierte gilt:

$$A_{\bar{h}\bar{k}0} = A_{hk0}^* \quad (4.9)$$

(Der Stern repräsentiert die Konjugation komplexer Zahlen). Daher ist

$$K_{200} = 4 [\operatorname{Re}(A_{200}) + \operatorname{Re}(A_{020})]^2 \quad (4.10)$$

dies stellt also eine phasenabhängige Projektion der Intensität dar und ist aus oben genannten Gründen zumindestens für unverspannte Systeme nicht sinnvoll.

Die Intensität der 000-Bildfrequenz wurde bei der lokalen Fouriertransformierten nicht berücksichtigt. Diese Intensität wird vom Nullstrahl dominiert und hat typischerweise eine um mindestens eine Größenordnung höhere Amplitude als alle anderen Raumfrequenzen. In der Intensität der 000-Bildfrequenz spiegelt sich die Gesamthelligkeit des zugrundeliegenden Bildes wieder. Diese wird bestimmt von der Absorptionseigenschaft des Materials. Man wird also in dieser Intensität Information über die Komposition des Materials erwarten. Wie in Abschnitt 5.1 dargestellt werden wird, muss man bei dem in dieser Arbeit zugrundegelegten Szenario für die Kompositionsbestimmung experimentelle HREM-Aufnahmen mit simulierten Abbildungen vergleichen. Dabei hat der Vergleich unabhängig von der Gesamthelligkeit zu erfolgen (s.a. Kap 5.2.1). Daher wird auf die Auswertung der 000-Bildfrequenz verzichtet. Die Elimination von Gesamthelligkeit und Gesamtkontrast wird beim Vergleich von Experiment und Simulation in den Abschnitten 5.2.1 und 5.2.2 genauer diskutiert.

Kapitel 5

Mittlere experimentelle Parameter

5.1 Szenario: Kompositionsbestimmung durch HREM-Bildauswertung

In hochaufgelösten TEM Aufnahmen von Halbleiterheterostrukturen lassen sich stets Bereiche finden, an denen die Probe in Dicke und Zusammensetzung nur wenig schwankt. Solche Bereiche werden im weiteren als Templatebereiche bezeichnet. Typischerweise lassen sich in einer HREM-Aufnahme einer Grenzfläche mindestens zwei solcher Bereiche finden. Wendet man Bildvorverarbeitung auf einen Templatebereich an, so lassen sich die extrahierten Werte (Parametervektoren) über den Templatebereich mitteln: Man erhält den mittleren Parametervektor des Templatebereichs.

Es wurde bewusst nicht vorausgesetzt, dass Komposition und Dicke der Probe im Templatebereich konstant sein müssen. Dies ist aufgrund von Random-Alloy-Fluktuationen (vgl. Kap. 8) und von Oberflächenamorphisierung (vgl. Kap. 3) nie der Fall. Wir mitteln bewusst über Dicken und Kompositionsschwankungen in den Templatebereichen. Dabei sollten diese Schwankungen allerdings nicht so groß sein, dass sich der mittlere experimentelle Parametervektor eines Templatebereichs von einem Parametervektor zur mittleren Probendicke und Komposition des Templatebereichs signifikant unterscheidet.

Betrachten wir eine HREM-Aufnahme mit zwei Template Bereichen ($\mathcal{T}_a, \mathcal{T}_b$). Für die Vorverarbeitung get3 haben wir nach der Templatemittelung zwei dreidimensionale mittlere experimentelle get3-Vektoren. Für die angestrebte Kompositionsbestimmung fehlt noch der Zusammenhang zwischen get3-Vektor \vec{G} und der Komposition x :

$$x(\vec{G}) \quad \text{mit} \quad \vec{G} \in \mathcal{G} \quad (5.1)$$

Dabei müssen die in Frage kommenden get3-Vektoren auf einen Bereich \mathcal{G} eingeschränkt werden, weil nicht jeder dreidimensionale Vektor einer Kompositionsinformation zuzuordnen ist.

Wir haben durch die Mittelwerte über die beiden Templatebereiche zwei Punkte der Funktion (5.1). Für sehr eingeschränkte experimentelle Bedingungen kann man die beiden Punkte linear interpolieren analog der Vorgehensweise bei [35]. Davon wurde hier abgesehen.

In dieser Arbeit wird die Interpolation durch simulierte HREM-Aufnahmen unterstützt. Dadurch erreicht man eine optimale Unabhängigkeit von speziellen Defokus- und Dickenbereichen bzw. speziellen Materialsystemen. Die Simulation hochauflöster TEM Bilder mit anschließender Bildvorverarbeitung erlaubt uns Punkte

$$\vec{G}(x) \tag{5.2}$$

zu berechnen. Diese Punkte können interpoliert und auf einem (möglicherweise eingeschränktem) Intervall umgekehrt werden. So erhalten wir die Funktion (5.1).

5.2 Vergleich von Experiment und Simulation

Wir vergleichen experimentelle HREM-Bilder quantitativ mit simulierten Aufnahmen.

Die hier vorliegenden experimentellen digitalen HREM-Bilder wurden auf dem Weg über Photonegativ, Photo und Einscannen des Photos erhalten. Dabei kann die Bildhelligkeit und der Gesamtkontrast des Bildes (Standardabweichung der Helligkeit) nicht interpretiert werden. Das liegt daran, dass beim Intensitätstransfer unbekannte Parameter, wie Sensitivität des Negativs und des Photopositivs, eine Rolle spielen¹. Es wurde darauf geachtet, dass die Dynamik des Photomaterials nur in kleinem Bereich ausgenutzt wurde. Somit kann zumindestens ein linearer Intensitätstransfer garantiert werden.

Die Gesamthelligkeit (Mittelwert der Bildintensitäten) und der Gesamtkontrast (Standardabweichung der Bildintensitäten) wurden hier bei der Simulation angepasst. Nur nach einer solchen linearen Anpassung können simulierte HREM-Abbildungen mit den untersuchten experimentellen Aufnahmen verglichen werden.

¹Bei CCD-Aufnahmen lassen sich die Intensitäten direkt interpretieren.

5.2.1 Lineare Anpassung bei get3

Aus den beiden Templatebereichen der experimentellen Aufnahmen \mathcal{T}_a und \mathcal{T}_b werden die get3-Vektoren gemittelt. So erhält man zwei mittlere get3-Vektoren

$$\mathbf{G}_a \quad \mathbf{G}_b \quad ,$$

die wir zu einem für das Experiment charakteristischen sechsdimensionalen Vektor

$$\mathbf{X} := \begin{pmatrix} G_{a1} \\ G_{a2} \\ G_{a3} \\ G_{b1} \\ G_{b2} \\ G_{b3} \end{pmatrix} \quad (5.3)$$

zusammenfassen. Gibt man einen Satz experimenteller Parameter vor, so kann man zu den beiden Templatebereichen HREM-Aufnahmen simulieren, und die beiden get3-Vektoren der Simulation s_a, s_b entsprechend 5.3 anordnen.

$$\mathbf{S} := \begin{pmatrix} s_{a1} \\ s_{a2} \\ s_{a3} \\ s_{b1} \\ s_{b2} \\ s_{b3} \end{pmatrix} \quad (5.4)$$

Aus eingangs genannten Gründen, kann man Simulation und Experiment nur nach einer Kontrast- und Helligkeitsanpassung vergleichen. Man sucht also zwei reelle Parameter m , t und vergleicht dann

$$m\mathbf{S} + t \leftrightarrow \mathbf{X} \quad (5.5)$$

Die Parameter m und t werden so gewählt, dass Simulation und Experiment optimal übereinstimmen. Dazu minimiert man die quadratische Abweichung

$$E = \sum_{i=1}^6 (mS_i + t - X_i)^2 \quad (5.6)$$

Dies kann man mit einer *least squares* Optimierung analytisch erreichen. Für m und t erhält man damit:

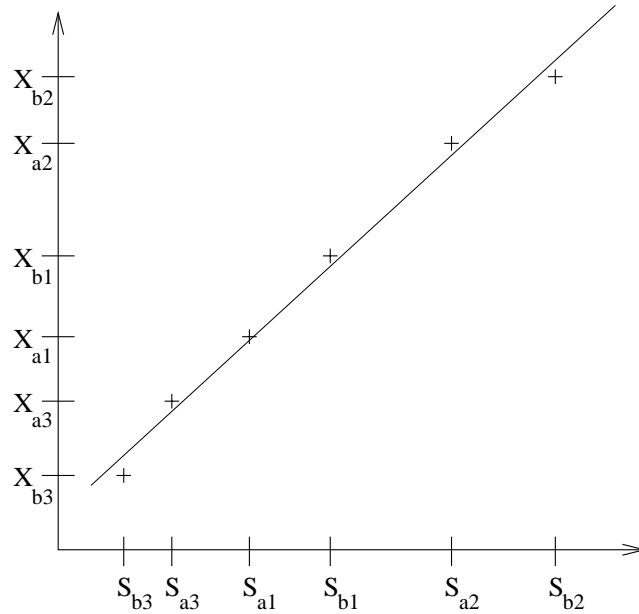


Abbildung 5.1: Veranschaulichung der linearen Anpassung von Simulation (S) an das Experiment (X). Die Regressionsgerade $s = mx + t$ der Punkte (S,X) liefert die gesuchten linearen Anpassungsparameter m und t in (5.5).

$$m = \frac{S_{11} - \frac{S_{01}S_{10}}{6}}{S_{20} - \frac{S_{10}^2}{6}} \quad (5.7)$$

$$t = \frac{S_{01} - aS_{10}}{6} \quad (5.8)$$

$$\text{mit: } S_{\mu\nu} = \sum_{i=1}^6 S_i^\mu X_i^\nu \quad (5.9)$$

Die Optimierung der Parameter m und t kann man sich veranschaulichen:

In Abbildung 5.1 sind horizontal die Komponenten der simulierten get3-Vektoren angetragen. Die Vertikalachse entspricht den experimentellen Werten. Man trägt die sich entsprechenden experimentellen und simulatorischen Werte als Punkte

$$(S_g, X_g) \quad (5.10)$$

in das Koordinatensystem ein. Der geforderte lineare Zusammenhang (5.5) zwischen Experiment und Simulation entspricht nun der Forderung, dass die Punkte auf einer Geraden liegen. Das Finden der optimalen Anpassungsparameter m und t entspricht der Bestimmung der Regressionsgeraden im Diagramm.

Vermeidung negativer Kontrastanpassung

Mit der so berechneten linearen Anpassung hat man nun die Möglichkeit, Experiment und Simulation unabhängig von den Werten für Gesamthelligkeit und Gesamtkontrast zu vergleichen. Dabei kann es allerdings zu Problemen kommen, weil in (5.5) eine negative Kontrastanpassung ($m < 0$) nicht ausgeschlossen wurde. Dies führt dazu, dass Bild und Inverse des Bildes als sehr ähnlich eingestuft werden. Dies ist aber physikalisch nicht sinnvoll. Also fordert man in einer Nebenbedingung zu der Optimierung, dass das gesuchte m positiv sein muss.

Angenommen bei einer Anpassung erhält man als „optimalen“ Wert für m einen negativen Wert m_0 :

$$\min_m(E) = E(m_0) \quad m_0 < 0 \quad (5.11)$$

Die Fehlerfunktion E (Gl. (5.6)) hat in m und t ein eindeutiges Minimum. Daher ist sie für $m > m_0$ monoton steigend:

$$m_0 < m_1 < m_2 \Rightarrow E(m_1) < E(m_2) \quad (5.12)$$

Damit ergibt sich das nicht negative m mit dem kleinsten Wert der Fehlerfunktion als

$$m_p = 0 \quad (5.13)$$

5.2.2 Lineare Anpassung bei lokaler Fouriertransformierter

Bei der lokalen Fouriertransformierten (siehe Abschnitt 4.4) gehen wir bei der linearen Anpassung in gleicher Weise vor wie in Abschnitt 5.2.1. Aus den Templatebereichen \mathcal{T}_a und \mathcal{T}_b erhalten wir jetzt einen vierdimensionalen Vektor

$$\mathbf{X} := \begin{pmatrix} F_{a200} \\ F_{a220} \\ F_{b200} \\ F_{b220} \end{pmatrix} \quad (5.14)$$

Gibt man einen Satz experimenteller Parameter vor, so kann man zu den beiden Templatebereichen HREM-Aufnahmen simulieren, und die beiden Wertepaare der Simulation s_a, s_b entsprechend (5.14) anordnen.

$$\mathbf{S} := \begin{pmatrix} s_{a200} \\ s_{a220} \\ s_{b200} \\ s_{b220} \end{pmatrix} \quad (5.15)$$

Betrachten wir die Fouriertransformierte eines linear angepassten Bildes:

Führt man bei einem $N_x \times N_y$ Bild \bar{I} eine Kontrast- und Helligkeitsanpassung durch, so ergibt sich für die diskrete Fouriertransformierte:

$$\text{FT}(m\bar{I} + t\bar{E}) = m \text{FT}(\bar{I}) + t \text{FT}(\bar{E}) \quad (5.16)$$

mit $E_{ij} = 1$ für $i = 1 \cdots N_y$ $j = 1 \cdots N_x$

Für die diskrete Fouriertransformierte der konstanten Matrix \bar{E} gilt:

$$\text{FT}(\bar{E})_{ij} = \begin{cases} N_x \cdot N_y & \text{für } i = j = 1 \\ 0 & \text{sonst} \end{cases} \quad (5.17)$$

Die Helligkeitsanpassung t wirkt sich also nur auf den Nullkoeffizienten der Fouriertransformation aus. Für die bei der lokale Fouriertransformierten verwendeten Koeffizienten F_{200} und F_{220} genügt daher eine Anpassung des Kontrastes:

$$m\mathbf{S} \leftrightarrow \mathbf{X} \quad (5.18)$$

Wie in Abschnitt 5.2.1 erhält man m aus einem *least squares* Fit:

$$m = \frac{S_{11}}{S_{20}} \quad (5.19)$$

$$\text{mit: } S_{\mu\nu} = \sum_{i=1}^4 S_i^\mu X_i^\nu \quad (5.20)$$

Wiederum muss beachtet werden, dass keine negative Kontrastanpassung erfolgt (analog zu 5.2.1).

5.3 Präzisieren von experimentellen Parametern

5.3.1 Beschreibung

Sind die experimentellen Parameter für die Templatreionen bekannt, so kann wie in den vorangegangenen Abschnitten beschrieben, eine lineare Anpassung der Simulation bestimmt werden. Dadurch werden Simulation und Experiment direkt vergleichbar. Zur Simulation von zwei Templatebereichen \mathcal{T}_a und \mathcal{T}_b benötigt man die folgenden experimentellen Parameter:

Mittlere Zusammensetzung beider Templatebereiche	x_a, x_b
Mittlere Probendicke beider Templatebereiche	t_a, t_b
Defokussierung der Abbildung	Z

Leider hat man in der Regel für diese Parameter nur sehr ungenaue Werte. Es soll nun ein Verfahren dargestellt werden, um solche mittleren experimentellen Parameter zu finden bzw. zu präzisieren.

Man teilt die relevanten Bereiche für die experimentellen Parameter in äquidistante Werte auf. Anschliessend simuliert man für jede Parameterkombination ein HREM-Bild, bestimmt den get3- (bzw. lokale FT-) Vektor und speichert diesen in einer Bibliothek. Der Kompositionsbereich $x \in [0, 1]$ wurde dabei in 32 Schritte unterteilt. Die Defokussierung wurde im Bereich von 5-104nm in Schritten von 1nm und die Probendicke im Bereich 5-53 Ez. (28-300Å) in Schritten von einer Elementarzelle simuliert. Für ein Materialsystem ergibt sich so eine typische Bibliotheksgröße von etwa 3 Mbyte. Die Bibliothek wird für ein Materialsystem nur einmal gerechnet und erspart viel Simulationszeit. Würde man statt der get3-Werte die simulierten 156800 HREM-Bilder speichern, so entspräche dies einem Speicherbedarf von 4 Gbyte.

Man berechnet die quadratische Abweichung der mittleren Templatevektoren des Experiments von den angepassten Werten aus der Bibliothek. Von den zwei möglichen get3-Gruppeneinteilungen des Experiments (vgl. Abschnitt 4.3.3) wird dabei immer diejenige mit der kleineren Abweichung betrachtet.

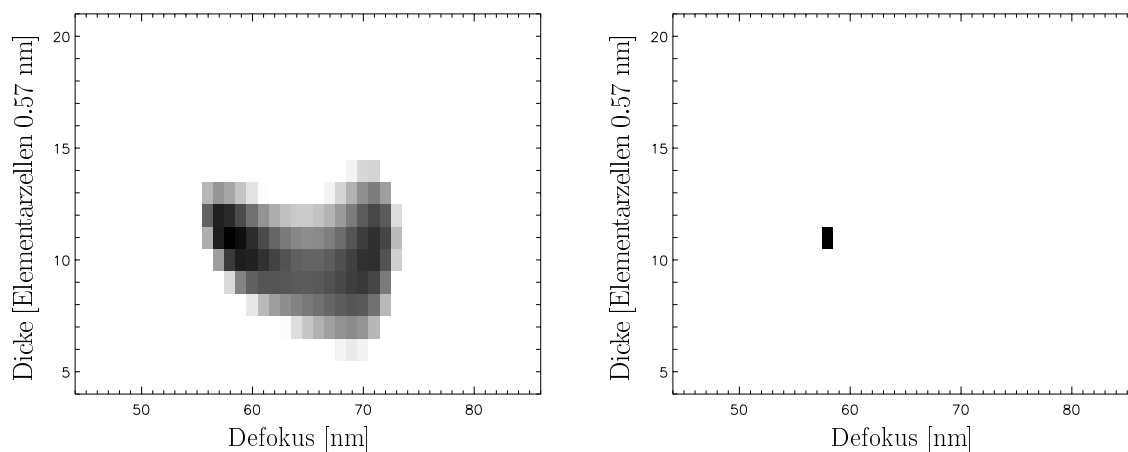


Abbildung 5.2: Quadratische Abweichungen von Experiment und angepassten Simulationen für verschiedene Defokus- und Dickenwerte. Rechts sind nur die Werte eingetragene Mittelwerte, die unterhalb des Fehlers für die experimentellen Mittelwerte lagen

Der linke Teil der Abbildung 5.2 illustriert eine solche Fehlerfläche (Vorverarbeitung: get3) in Abhängigkeit von der Probendicke in einem Templatebereich und der Defokus-

sierung. Bei allen anderen beteiligten Parametern wurde jeweils der minimale Fehler eingetragen. Dunkle Werte entsprechen kleinen Fehlern.

Um zu entscheiden, welche der dargestellten Werte mögliche Parameterkombinationen der experimentellen Templatebereiche sind, muss eine Fehlerschranke eingeführt werden. Diese ergibt sich aus dem Fehler der experimentellen Mittelwerte.

Am Beispiel der Vorverarbeitung get3 und einer Aufnahme mit zwei Templatebereichen soll dies erläutert werden:

Aus den beiden Templatebereichen \mathcal{T}_a und \mathcal{T}_b mitteln wir zwei get3-Vektoren: G_a , G_b . Geht man von statistisch unabhängigen Schwankungen aus, so addieren sich die quadratischen Fehler der Einzelkomponenten des sechsdimensionalen Vektors (5.3). Somit erhält man für die mittlere quadratische Abweichung vom mittleren Vektor den Wert:

$$E_{\min} = \frac{1}{N_a} \sum_{i=1}^3 \text{var}(G_{ai}) + \frac{1}{N_b} \sum_{i=1}^3 \text{var}(G_{bi}) \quad (5.21)$$

Wobei N_a und N_b jeweils die Anzahl der statistisch unabhängigen Beispiele in den Templatebereichen sind. Parameterkombinationen die eine kleinere quadratische Abweichung vom Experiment als E_{\min} haben, sind mögliche mittlere Parameter der Templatebereiche.

Der rechte Teil der Abbildung 5.2 zeigt den Teil der Fehlerfläche welcher den Fehler der experimentellen Mittelwerte (E_{\min}) unterschreitet. Das Sampling der Bibliothek ist hier hinreichend fein ($\Delta t = 1$ Elementarzelle, $\Delta Z = 1\text{nm}$). Damit ergibt sich hier ein eindeutiges Ergebnis in der Parametersuche.

Anzahl der statistisch unabhängigen get3-Vektoren im Templatebereich

HREM-Aufnahmen von (homogenen) Sphaleritstrukturen in (001)-Projektion besitzen Translationssymmetrie bei der Verschiebung um die Hälfte der Gitterkonstanten entlang der lateralen Kristallachsen ($T(\frac{A}{2})$ vgl. Abschnitt 4.3.2).

Man bestimmt den get3-Vektor zu einem Bildbereich, welcher der Größe der Elementarzelle entspricht. Auf einem Bereich, der um die Hälfte der (Bild-) Gitterkonstanten in eine der lateralen Kristallachsen verschoben ist, lässt sich wieder ein get3-Vektor berechnen. Ein Bereich einer get3-Auswertung überlappt auf diese Weise jeweils mit vier anderen Bereichen.

Typischerweise kann die Anzahl der am Rand des Templatebereichs bestimmten get3-Vektoren im Vergleich zur Gesamtzahl der get3-Vektoren im Templatebereich vernachlässigt werden (clusterförmige Templatebereiche). In diesem Fall sind also nur ein Viertel der in einem Templatebereich gemittelten get3-Vektoren statistisch unabhängig. Dies muss beim Berechnen der Fehlergrenze (5.21) beachtet werden.

Vorabwissen über die Probenparameter

In der Regel liegt Vorabwissen über die experimentellen Parameter vor. Die Kompositionen der Templatebereiche kennt man aus den Herstellerangaben auf etwa 5–10 Prozent genau. Bei den binären Materialien (z.B. GaAs oder AlAs) kann man von einer exakten Kenntnis ausgehen. Bei der Defokussierung ist zumindestens die Defokussierungshalbperiode bekannt, in der die Aufnahme gemacht wurde. Solchem Vorabwissen kann Rechnung getragen werden, indem bei der Parametersuche die Parameterbibliothek nur zum Teil abgesucht wird. Dadurch kann bei sehr verrauschten Abbildungen (hohe Fehlerschranke (5.21)) der Bereich der möglichen experimentellen Parameter eingeschränkt werden.

5.3.2 Diskussion der Parametersuche

Bei *least squares* Anpassungen ist es üblich die Einzelsummanden des quadratischen Fehlers (5.6) mit dem Reziprokwert der Varianz zu gewichten [36]. Auf diese Weise legt man beim Vergleich weniger Gewicht auf Gruppen die mehr Rauschen enthalten.

Bei der Gruppe G_3 sind bei der Mittelung über die Gruppenpositionen im Vergleich zu G_1 und G_2 mehr Werte beteiligt. Daher ist die Varianz in dieser Gruppe vergleichsweise niedrig. Mit der beschriebenen Vorgehensweise wird diese Gruppe stärker als die beiden anderen gewichtet.

In Abschnitt 4.3.2 wurde verdeutlicht, daß sich speziell die Information über die Komposition vor allem in den Gruppen G_1 und G_2 wiederfindet. Daher wäre es sinnvoller die Gruppe G_3 weniger zu gewichten. Dies würde allerdings voraussetzen, dass man die Abbildungsparameter schon *vor* der Parametersuche genau kennt.

Aus diesen Gründen wurde eine Gewichtung der Fehlerfunktion nicht durchgeführt.

Bei der Anpassung von Simulationen an experimentelle Aufnahmen wurde der Aspekt der negativen Kontrastanpassung besprochen (Abschnitt 5.2.1). Bei der Parametersuche können Parameterkombinationen vernachlässigt werden, deren simulierte get3-Vektoren eine negative Kontrastanpassung zum Experiment erfordern. Diese erreichen

nur den Fehler (Gleichung (5.6)) für $m=0$. Dieser ist für alle Simulationen mit negativer Anpassung gleich. Eine Simulation mit positiver Anpassung ist im Fehler immer besser als dieser Wert.

5.4 Einfluss von amorphen Deckschichten

In diesem Abschnitt wird der Einfluss von amorphen Deckschichten (s. Kap. 3) auf die Parametersuche (s. Kap. 5.3) untersucht. Dazu wurde eine $\text{Al}_{0.4}\text{Ga}_{0.6}\text{As}/\text{GaAs}$ Grenzschicht simuliert. Kompositionen und Defokus entsprechen dabei der in Abschnitt 8.1 vorgestellten experimentellen Anwendung. Davon abweichend wurde bei beiden Templatebereichen die gleiche kristalline Probedicke ($t=12$ Elementarzellen = 68\AA) simuliert. Durch diese Vorgehensweise ist eine zweidimensionale Darstellung der Ergebnisse möglich.

Die Parametersuche wurde einerseits mit get3 (Kapitel 4.3), andererseits mit der lokalen Fouriertransformierten (Kapitel 4.4) als Bildvorverarbeitung durchgeführt. Abbildung 5.3 zeigt das Ergebnis ohne amorphe Oberflächenschichten, sowie für die amorphen Gesamtdicken von jeweils 10\AA und 30\AA Dicke. Innerhalb der Abbildung ist für jede amorphe Schichtdicke jeweils ein Dicken-Defokusdiagramm dargestellt.

Der rote Punkt markiert dabei die Ausgangsparameter der simulierten ungestörten kristallinen Probe. Grüne Punkte entsprechen den mit der get3 Vorverarbeitung gefundenen Parameterkombinationen die unter dem Fehlerlimit (5.21) sind. Das Ergebnis einer Parametersuche mit der lokalen Fouriertransformierten und äquivalenter Fehler-schranke² ist mit blauen Punkten gekennzeichnet.

Ohne amorphe Oberflächenschichten finden sowohl die Parametersuche mit get3 als auch die Parametersuche mit lokaler Fouriertransformierter die exakten Parameterwerte der Probe (Abb. 5.3 oben). Eingezeichnet ist hier nur die Parameterkombination mit dem minimalen Fehler. Die Fehlergrenze (5.21) scheidet als Auswahlkriterium aus, da sie ohne den Störeinfluss der Oberflächenschichten verschwindet.

Eine amorphe Schichtdicke von 10 Angström (Abb. 5.3 unten links) verursacht Rauschen in der Abbildung und damit einen endlichen Wert der Fehlergrenze (5.21). Die erhöhte Fehlerschranke führt zu Mehrdeutigkeit bei der Parameterbestimmung bei beiden Verfahren. Die get3 Vorverarbeitung schränkt dabei den möglichen Parameterbereich wesentlich weiter ein als die lokale Fouriertransformierte. Die Ausgangsparameter werden von get3 im Gegensatz zur lokalen Fouriertransformierten noch gefunden.

²Das bedeutet nicht, dass die Fehlerschranke bei beiden Verfahren den gleichen Wert hat. Sie wird aus der Variation der Parameterwerte in beiden Fällen nach Gleichung (5.21) berechnet.

Wie in Abschnitt 3.3 an einigen experimentellen Proben untersucht wurde, treten in der Praxis amorphe Oberflächenschichten von etwa 30\AA Gesamtdicke auf. In Abbildung 5.3 rechts unten wird die Parametersuche mit dieser amorphen Schichtdicke durchgeführt. Um auch bei dieser Dicke der Oberflächenschichten eine sinnvolle Parametersuche durchführen zu können, muss Vorabwissen aus der Probenpräparation verwendet werden. Bei den später untersuchten experimentellen Proben ist die mittlere Komposition der Templatebereiche mit 5–10 Prozent Fehler bekannt. Bei der hier durchgeführten Untersuchung wurde diese Einschränkung pessimistisch als zehn Prozent angenommen. Bei der Parametersuche werden also nur Parameterkombinationen akzeptiert, welche um weniger als 10 Prozent von den angegebenen Kompositionen der Templatebereiche abweichen.

Das Verfahren mit get3 Vorverarbeitung findet einen eindeutigen Parameterwert in der Nähe der Parameter der ungestörten Probe. Im Gegensatz dazu liegen die vom Verfah-

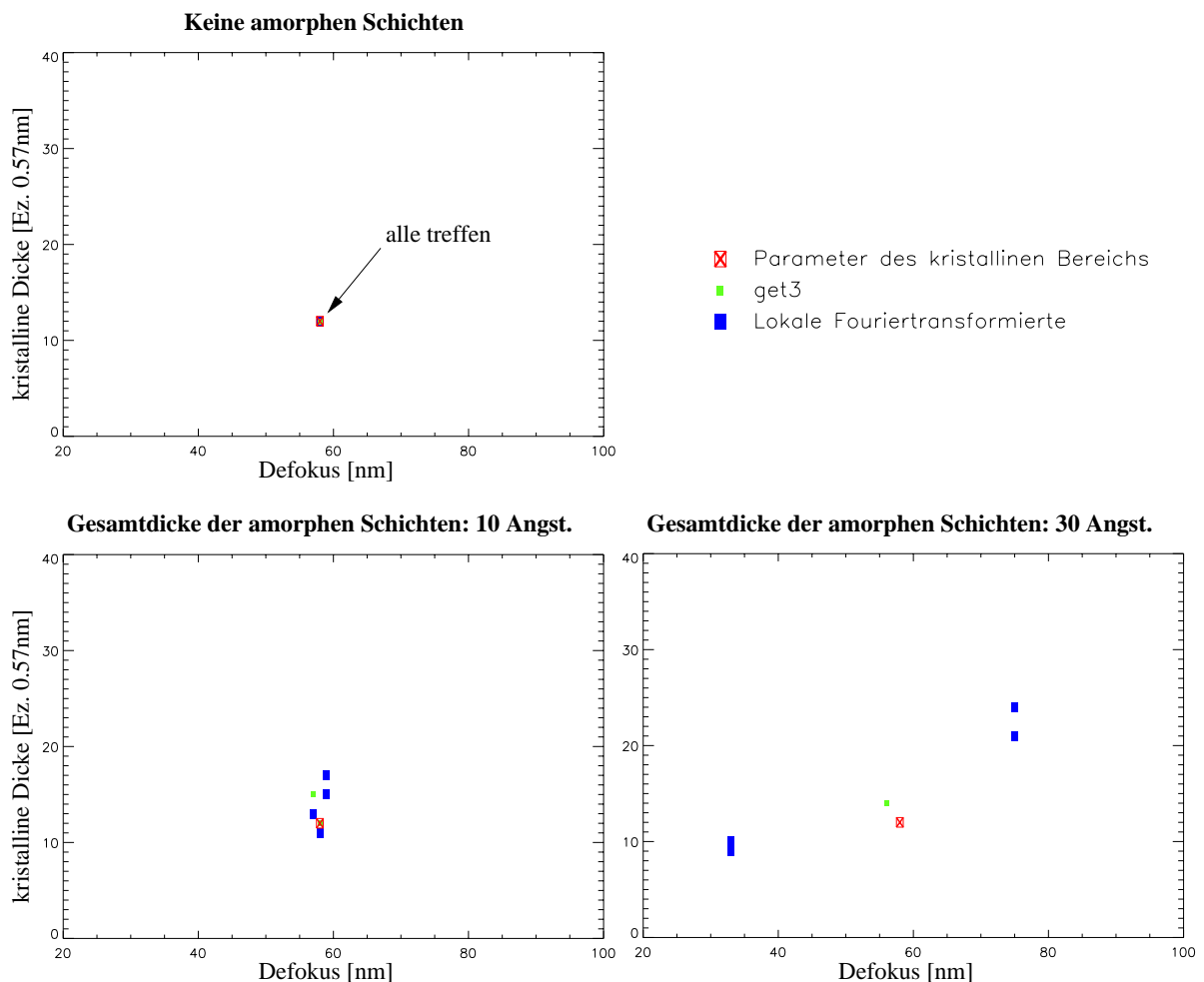


Abbildung 5.3: Ergebnis der Parametersuche mit amorphen Oberflächenschichten von jeweils 0, 10, 30 Ångström Gesamtdicke.

ren mit lokaler Fouriertransformierter bestimmten Werte weit von den Parametern der ungestörten kristallinen Probe entfernt. Hier kann nicht einmal die Defokussierungshalbperiode (bei GaAs $t=68 \text{ \AA}$ $Z \in [40 \text{ nm}, 64 \text{ nm}]$) eindeutig zugeordnet werden. Der gefundene Dickenbereich erlaubt kaum eine Aussage über die Dicke der Templatebereiche der Probe. Für die hier angestrebte Bestimmung der experimentellen Parameter ist die lokale Fouriertransformierte daher völlig ungeeignet. Weil schon das Finden der ortsgemittelten Probendicke Probleme bereitet, kann die im weiteren Verlauf der Arbeit angestrebte lokale Dickenbestimmung mit der lokalen Fouriertransformierten nicht durchgeführt werden.

Bleibt noch zu klären, ob die lokale Fouriertransformierte evtl. als Vorverarbeitung für die lokale Kompositionsbestimmung von Vorteil ist. Abbildung 5.4 zeigt das Ergebnis der Parametersuche an der simulierten Testprobe wiederum mit amorphen Oberflächenschichten von 30 \AA Dicke. Gegeneinander aufgetragen sind die Kompositionen der beiden Templatebereiche, welche unter der Fehlerschranke (5.21) liegen.

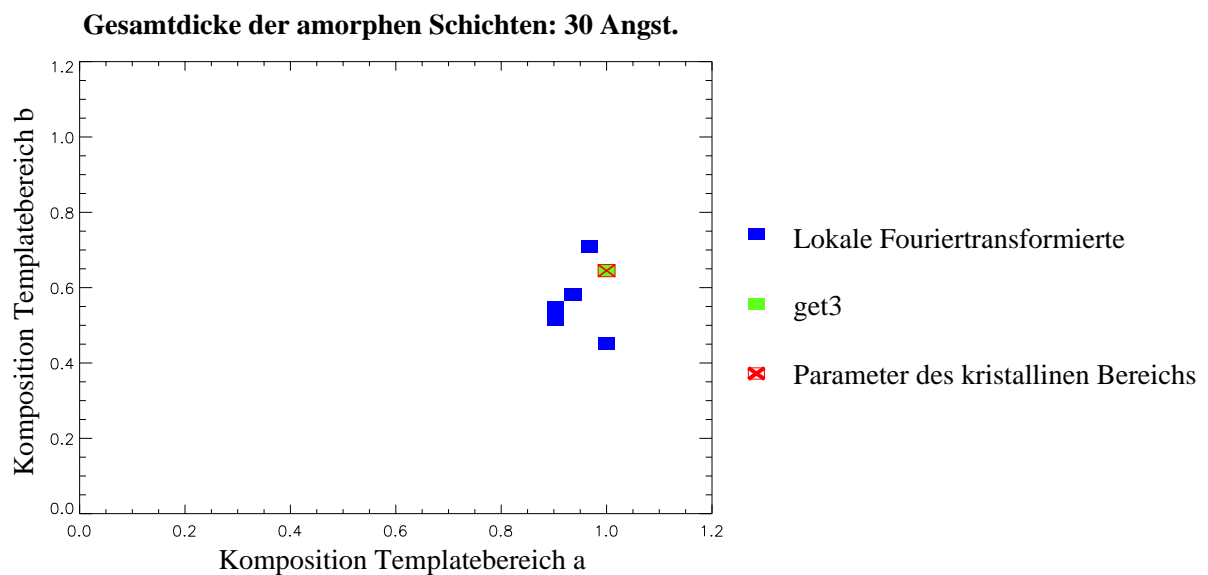


Abbildung 5.4: Bestimmung der mittleren Komposition in den Templatebereichen mit amorphen Oberflächenschichten von 30 \AA Gesamtstärke.

Das Verfahren mit `get3` Vorverarbeitung findet den exakten Wert der ungestörten kristallinen Probe in eindeutiger Weise. Dagegen sind die vom Verfahren mit lokaler Fouriertransformierter akzeptierten Parameterkombinationen mit einem Fehler von bis zu 20 Prozent behaftet. Der exakte Wert wird von diesem Verfahren nicht gefunden. Die lokale Fouriertransformierte ist also ungeeignet um in lateral aperiodisch gestörten Objekten eine großflächig ortsgemittelte Kompositionsbestimmung durchzuführen. Daher wird auch bei der später durchgeführten lokalen Kompositionsbestimmung nur die `get3` Vorverarbeitung verwendet.

5.5 Kompositions- und Dickenbestimmung ohne Rauschen

Wir sind nun in der Lage, aus HREM-Aufnahmen deren mittlere experimentelle Parameter zu bestimmen. Mithilfe dieser Parameter können angepasste HREM-Aufnahmen simuliert werden, die direkt mit dem Experiment vergleichbar sind. Vernachlässigt man zunächst einmal das Rauschen in den experimentellen Bildern, so gestaltet sich eine Kompositionsbestimmung sehr einfach.

Der Dicken und Kompositionsbereich wird dicht mit simulierten Beispielen gesampelt. Aus den simulierten HREM-Bildern erhält man Stützstellen für die Beziehung $\vec{G}(x, t)$ zwischen Komposition (und Dicke) und get3-Vektor. Diese können interpoliert werden. Daraus erhält man die Funktion $\vec{G}(x, t)$. Diese Funktion kehrt man auf einem Intervall für Komposition x und Dicke t um. So erhält man die gesuchte Beziehungen $x(\vec{G})$ und $t(\vec{G})$ und damit die Kompositions- und Dickenbestimmung.

Im Gegensatz zu obiger Annahme sind HREM-Aufnahmen stark verrauscht. In Abschnitt 3.3 haben wir den Kontrastanteil Rauschen durch amorphe Deckschichten für typische HREM-Aufnahmen als 25-40 Prozent abgeschätzt. Um mit quantitativer HREM sinnvolle Fehlerwerte zu erreichen, muss der Einfluss des Rauschens daher unterdrückt werden.

Zur Unterdrückung des Rauschens werden in Kapitel 7 verschiedene Methoden vorgestellt und verglichen. Die für das Verständnis dieses Kapitels notwendigen Grundlagen des Einsatzes neuronaler Netze werden im folgenden Kapitel skizziert.

Kapitel 6

Neuronale Netze

Ein wesentlicher Teil der in dieser Arbeit entwickelten Methoden zur lokalen Kompositions- und Dickenbestimmung verwendet neuronale Netze zur Unterdrückung von Störeinflüssen. Die zum weiteren Verständnis der Arbeit notwendigen Grundlagen sind in diesem Kapitel dargestellt. Eine detailliertere Darstellung wurde in [26] gegeben. Details zum biologischen Hintergrund der hier vorgestellten Methoden beschreibt umfassend [12].

6.1 Verwendeter Netztyp

Die hier zum Einsatz kommenden neuronalen Netze entsprechen dem in der Arbeit [26] vorgestellten Netzwerktyp (analoge *feed-forward* Netze). Das neuronale Netz errechnet aus einem Eingabevektor \vec{G} einen eindimensionalen Ausgabewert f . Die vom neuronalen Netz dabei berechnete Funktion ist

$$f(\vec{G}) = W_0 + \sum_{i=1}^{N_h} W_i \tanh\left(w_{i0} + \sum_{j=1}^3 w_{ij} G_j\right) \quad (6.1)$$

Die Netzgewichte \vec{W} und \bar{w} sind dabei vor der Verwendung des neuronalen Netzes in einem sogenannten Lernprozess (siehe 6.2) zu finden. Die Anzahl N_h der Neuronen in der Zwischenschicht bestimmt man im Rahmen eines Architekturadaptationsprozesses (siehe 6.4).

Die eindimensionale Ausgabe des neuronalen Netzes ist keine Einschränkung in der Architekturwahl, da durch paralleles Verwenden von Netzwerken dieses Typs auch beliebig hochdimensionale Ausgaben realisiert werden können.

6.2 Lernverfahren

Das neuronale Netz soll einen gewünschten Zusammenhang zwischen Eingabevektor \vec{G} und Netzausgabe f aus vorgegebenen Beispielen lernen (*supervised learning*). Dazu gibt man in einer Datei (Trainingsatz) Beispiele (Muster)

$$(\vec{G}_\mu, f_\mu) \quad \mu = 1 \cdots N_m$$

für gewünschte Eingabe-Ausgabe Kombinationen vor. In einem nichtlinearen Optimierungsprozess berechnet man die Netzgewichte \vec{W} und \bar{w} so, dass Trainingsmuster und Netzausgabe eine minimale quadratische Abweichung haben:

$$\min_{\vec{W}, \bar{w}} \left(\sum_{i=1}^{N_m} (f_\mu - f(\vec{G}_\mu))^2 \right) \quad (6.2)$$

In einer vorangegangenen Arbeit [26] wurden mehrere Trainingsverfahren (Methoden zur Optimierung der Netzgewichte) an Problemstellungen der Funktionsapproximation getestet. Die besten Ergebnisse erreichte der RPROP Algorithmus [38] mit einem geeigneten Vortraining. In dieser Arbeit wurden mehrere der Verfahren auch an der Problemstellung Kompositions- und Dickenbestimmung getestet. Dabei stellte sich heraus, dass ein Vortraining zum RPROP Algorithmus für diese Problemstellung keine Verbesserung bietet. Daher wurde als Trainingsverfahren im weiteren der RPROP Algorithmus verwendet.

6.3 Vor- und Nachverarbeitung

6.3.1 Reskalierung von Eingabe und Ausgabe

Die Netzgewichte werden vor dem Netztraining mit gleichverteilten Zufallswerten vorgelegt. Die optimale Wahl des Vorbelegungsintervalls ist abhängig von der Größenordnung von Eingabe- und Ausgabewerten.

In einer von Bishop [3] vorgeschlagenen Vorgehensweise wird diese Abhängigkeit von den Trainingsdaten durch eine lineare Reskalierung des Trainingsatzes eliminiert. Die lineare Umrechnung der Eingabe- und Ausgabedaten erfolgt ausserhalb des Netztrainings. Dem neuronalen Netz werden im Trainingsprozess die linear umgerechneten

Eingaben und Ausgaben präsentiert. Die lineare Umrechnung erfolgt so, dass der Trainingssatz jeweils in Eingabe und Ausgabe den Mittelwert Null und die Standardabweichung Eins besitzt. Die Vorbelegung aller Netzgewichte kann dann gleichverteilt im Intervall $[-1, 1]$ erfolgen [3]. Bei der Verwendung des so trainierten neuronalen Netzes für die Kompositions- bzw. Dickenbestimmung müssen die Eingabe- und Ausgabedaten entsprechend der im Training durchgeführten Skalierung umgerechnet werden.

6.3.2 Vorwissen über Kompositionsgrenzen

Bei der Kompositionsbestimmung mit neuronalem Netz liefert das Netz zum Teil Ergebnisse ausserhalb des physikalisch sinnvollen Intervalls $x \in [0, 1]$ (Die Abweichungen sind nur wenige Prozent). Hierin äussert sich die Tatsache, dass das neuronale Netz mit sinnvoller Architekturwahl die Trainingsmuster nur approximiert, eine Eigenschaft die zum Unterdrücken von Rauschen wesentlich ist [20] [16].

Die Berücksichtigung des physikalisch sinnvollen Ausgabeintervalls kann mit einer Vergrößerung der Netzarchitektur (siehe 6.4) unter gleichzeitigem Mehraufwand an Trainingsmustern mit beliebiger Genauigkeit erzwungen werden. Für den Rechenzeitaufwand günstiger ist aber ein einfaches Abschneiden des Netzergebnisses an den Kompositionsgrenzen. Daher werden Netzausgaben größer als Eins auf die Komposition Eins und negative Netzausgaben auf die Komposition Null abgebildet.

6.4 Architektur Anpassung

Die Wahlfreiheit in der Architektur beschränkt sich bei dem hier verwendeten Netztyp auf die Anzahl von Zwischenschichtneuronen N_h (vgl. (6.1)). Die Architekturwahl bestimmt die Qualität der Ausgabe des trainierten neuronalen Netzes für nicht im Trainingssatz enthaltene Eingabevektoren (Generalisierungsfähigkeit). Man möchte die Architektur soweit reduzieren, dass das Netz auch nicht im Training gezeigte Muster korrekt beantwortet. Dazu teilt man die vorhandenen Muster auf, in einen Trainingssatz (zum Trainieren des Netzes) und einen Validierungssatz. Die Qualität des trainierten neuronalen Netzes wird mit der Wurzel der quadratischen Abweichung auf dem Validierungssatz (Validierungsfehler) beurteilt. Abbildung 6.1 stellt für eine Kompositionsbestimmung die Abhängigkeit des Validierungsfehlers von der Anzahl der verwendeten Zwischenschichtneuronen dar.

Die Validierungsfehlerkurve in Abhängigkeit von N_h verläuft ab einer minimalen Architekturgröße sehr flach. Grund hierfür ist die hohe Anzahl an Trainingsmustern die

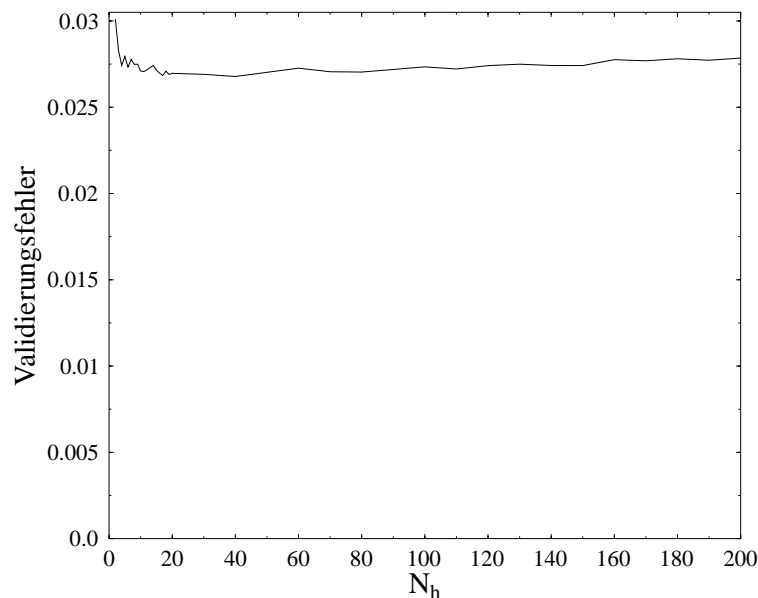


Abbildung 6.1: Abhängigkeit des Validierungsfehlers von der Anzahl der Zwischenschichtneuronen N_h

präsentiert wurden. Aus der Simulation kann man sich beliebig viele Trainingsmuster erzeugen (limitierend ist nur die Rechenzeit). Typischerweise wurden Trainingssätze von wenigstens 1000 Mustern verwendet, das entspricht bei einer Architekturgröße von $N_h = 20$ einer Netzlast von über 10 Mustern pro freiem Parameter (Netzgewicht).

Aufgrund der Tatsache, dass die Architekturwahl hier sehr unkritisch ist, wurde auf ausgefeiltere Architektur Anpassungsverfahren (siehe [26],[32],[10],[31]) verzichtet. Das Netztraining fällt von seinem Zeitbedarf im Vergleich zum Zeitaufwand zum Erzeugen der Trainingsmuster kaum ins Gewicht. Daher kann sich die Architektur Anpassung hier darauf beschränken, Architekturen bis zu einer bestimmten Obergrenze ($N_h = 30$) systematisch durchzuprobieren. Die Architekturgröße mit minimalem Validierungsfehler wird dann zur lateralen Kompositions- bzw. Dickenbestimmung verwendet.

6.5 Experimentelle Muster

Extrahiert man die get3-Vektoren aus einem experimentellen Templatebereich, so variieren diese. Abbildung 6.2 zeigt die Variation in zwei solcher Templatebereiche einer experimentellen Aufnahme. Gezeichnet sind jeweils drei Projektionen der dreidimensionalen Daten auf die Ebenen senkrecht zu jeweils einer Koordinatenachse. Man erkennt deutlich die beiden Punktwolken der beiden Templatebereiche (Kompositionen $x = 0$ und $x = 1$ in $\text{Al}_{1-x}\text{Ga}_x\text{As}$).

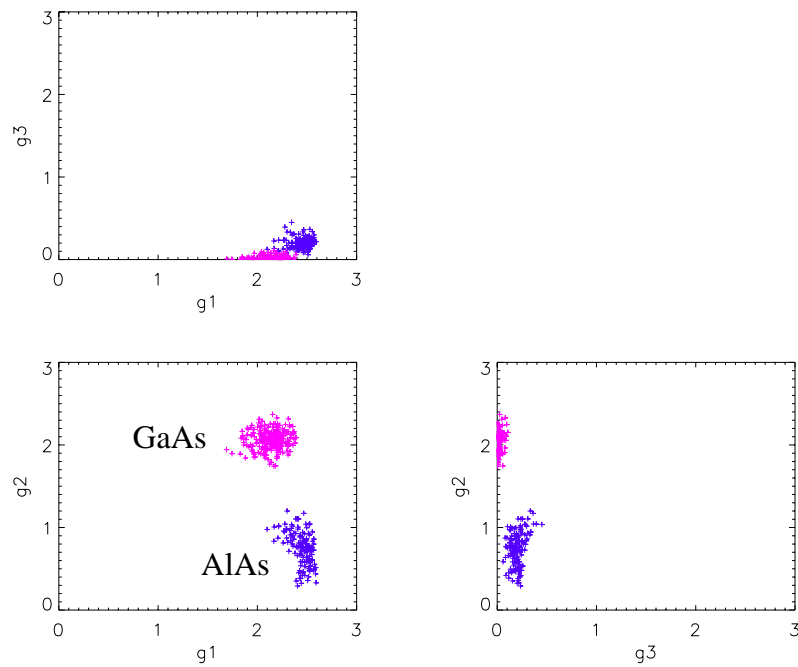


Abbildung 6.2: Experimentelles Beispiel: get3-Vektoren aus den Templatebereichen. Die drei Diagramme zeigen jeweils eine Projektion senkrecht zu einer Komponente der get3-Vektoren.

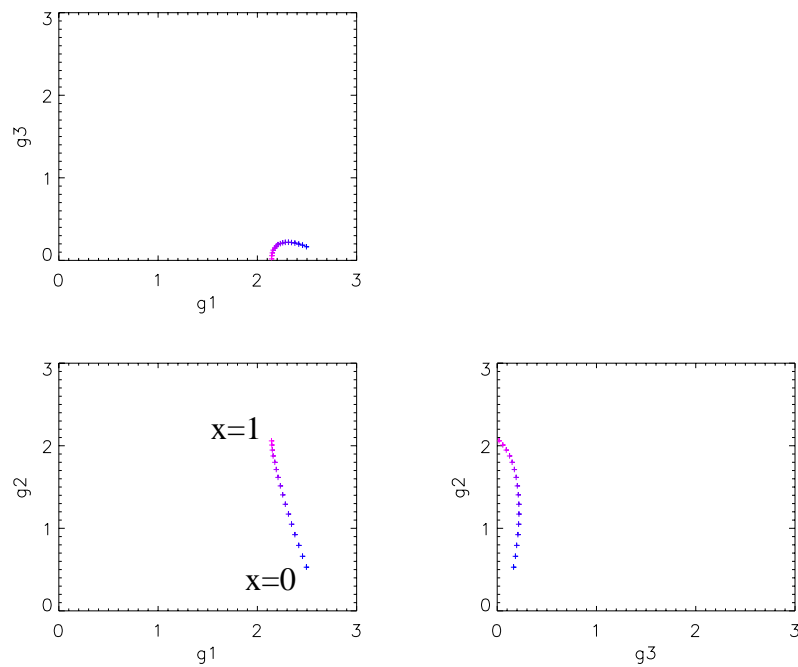


Abbildung 6.3: An die experimentellen Daten in 6.2 angepasste Simulation. Proben-
dicke 21 Ez., Defokussierung 54nm, Komposition x in $\text{Ga}_x\text{Al}_{1-x}\text{As}$ variiert in 16 Schrit-
ten von $x=0$ bis $x=1$.

Grund für die Variation der $get3$ -Vektoren eines Templatebereichs ist einerseits das Rauschen in der HREM-Abbildung. Andererseits aber auch geringfügige Kompositions- und Dickenschwankungen innerhalb des Templatebereichs. Abbildung 6.3 zeigt eine zum Experiment (Abb. 6.2) angepasste Simulation, bei der die Komposition variiert ($x = 0 \dots 1$). Wie man unschwer erkennt, kann eine Kompositionsvariation innerhalb der Templatebereiche nicht ausgeschlossen werden. Bei Templatebereichen mit ternärem Material treten dazu noch verstärkt Kompositionsvariationen auf, durch sogenannte Random-Alloy-Fluktuationen (statistische Besetzung eines Untergitters mit zwei Atomsorten). Die kristalline Probendicke schwankt innerhalb eines Templatebereichs vor allem durch die lateral variierende Amorphisierungstiefe.

Für die anschließend zu besprechenden Verfahren zur Unterdrückung von Rauschen benötigt man Beispiele (Muster) der zu unterdrückenden Störungen. Verwendet man hierfür die $get3$ -Vektoren aus experimentellen Templatebereichen, so werden auch die darin auftretenden Dicken- und Kompositionsvariationen unterdrückt. Diese Vorgehensweise führt zu einer nicht kontrollierbaren Verfälschung (eine Fehlerbestimmung ist schwierig) der Kompositionsbestimmung bzw. Dickenbestimmung.

Daher werden in den folgenden Verfahren statt der experimentellen Muster künstlich verrauschte simulierte $get3$ -Vektoren verwendet. Für diese ist die Komposition und Dicke exakt bekannt.

Kapitel 7

Unterdrückung von Rauschen

Bei der angestrebten lokalen Kompositionsbestimmung ist es notwendig, Störeinflüsse von amorphen Oberflächenschichten und von Schwankungen in der Probendicke zu unterdrücken. Eine lokale Dickenbestimmung wird gestört von Kompositionsschwankungen und wiederum durch amorphe Oberflächenschichten. Auch bei dieser zweiten Zielstellung sollen die Störungen nach Möglichkeit unterdrückt werden.

In diesem Kapitel werden verschiedene Methoden zur Unterdrückung von (evtl. korrelierten) Störungen beschrieben und anhand eines Testszenarios verglichen.

7.1 Beschreibung des Testszenarios

Als Testszenario werden simulierte HREM Abbildungen von $\text{Al}_{1-x}\text{Ga}_x\text{As}$ Proben verwendet. Es wird jeweils der gesamte Kompositionsbereich $x \in [0, 1]$ in zehn Schritte eingeteilt. Zu jeder Komposition werden verrauschte Bilder mit Probendicken von 9 bis 15 Elementarzellen berechnet. Die amorphen Oberflächenschichten hatten dabei eine Gesamtdicke von 30 Angström. Die HREM-Abbildungen wurden mit einer Defokussierung von 58 nm berechnet.

Die zu testenden Methoden müssen zu jedem Muster des Testsatzes die lokale Probendicke und die Komposition bestimmen. Für beide Zielstellungen werden jeweils die durchschnittlichen Fehler auf dem gesamten Testsatz bestimmt.

Den relativen Fehler der lokalen Probendickenbestimmung errechnet man als Verhältnis des absoluten Fehlers zur mittleren kristallinen Probendicke (12 Elementarzellen).

7.2 Vorgehensweise mit neuronalem Netz

Wir trainieren neuronale Netze mit Mustern aus dem gesamten Kompositionsbereich ($x \in [0, 1]$). Zu jeder Komposition werden Probendicken von 8 bis 19 Elementarzellen im Training berücksichtigt. Die Gesamtdicke der amorphen Oberflächenschichten beträgt 30 Angström. Es werden zwei Mustersätze erzeugt. Ein Dickenmustersatz mit der Probendicke als Ausgabewert und ein Kompositionsmustersatz mit der Komposition als Ausgabewert. Ein Mustersatz wird jeweils im Verhältnis 1:3 in Trainingssatz und Validierungssatz aufgeteilt. Mit dem Trainingssatz werden von einer Architekturgröße jeweils 10 Netze trainiert. Untersucht werden Netzwerke mit Architekturgröße $N_H = 2 \cdots 30$.

Von allen trainierten Netzwerken wird das mit dem kleinsten Fehler auf dem Validierungssatz für die Auswertung verwendet. Die Vorgehensweise entspricht der Auswertung von experimentellen HREM Aufnahmen in Abschnitt 8.

Die Auswertung des Validierungsfehlers ergab für die Kompositionsbestimmung eine optimale Netzgröße von 17 Zwischenschichtneuronen und für die Probendickenbestimmung $N_H = 15$.

Die Auswertung des Testfehlers ergab für beide Auswertungen:

Methode: neuronales Netz		
	absoluter Fehler	relativer Fehler
Kompositionsbestimmung	0.067	6.7%
Dickenbestimmung	0.300 Ez.	2.5%

7.3 Minimalabstand zu simulierten Mustern

Eine naheliegende Methode ist es, aus simulierten get3-Vektoren denjenigen auszusuchen, welcher die kleinste quadratische Abweichung zu einem experimentellen get3-Vektor hat. Die Parameter der zugrundeliegenden Simulation sind dann die gesuchten Werte für die lokale Komposition und Probendicke im Experiment.

Für die Suche nach der Simulation mit der kleinsten quadratischen Abweichung zum Experiment wird analog der Vorgehensweise in 5.3 eine Bibliothek mit simulierten get3-Vektoren gerechnet. Die Bibliothek unterteilt den gesamten Kompositionsbereich in 100 Schritte. Bei einer Dicke von 12 Elementarzellen sind zwar die möglichen Kompositionswerte wesentlich gröber quantisiert. Bei der Simulation lassen sich aber über

die VCA-Näherung (siehe Anhang B) auch Zwischenwerte der Komposition berechnen. Damit können die in der Realität auftretenden Kompositionen optimal interpoliert werden und eine Trennung der Kompositionsbestimmung von der Bestimmung der Probendicke ist möglich. Die Bibliothek umfasst den Dickenbereich 8 bis 16 Elementarzellen, und wurde mit einer Defokussierung von 58 nm berechnet. Der Dickenbereich wurde beim Vergleich nur eingeschränkt von 9 bis 15 Elementarzellen betrachtet und ist daher schon optimal voreingeschränkt (Die zusätzlichen Dickenwerte wurden für eine spätere Interpolation berechnet). Die hier durchgeführte Auswertung (Testfehler) stellt also eine untere Abschätzung des bei einer experimentellen Anwendung möglichen Fehlers dar:

Methode: Minimalabstand Simulation		
Kompositionsbestimmung	0.162	16.2%
Dickenbestimmung	1.76 Ez.	14.7%

Eine weitere Ausarbeitung der Methode interpoliert die Bibliothek. Dazu werden für die Kompositionsbestimmung zu der vom Experiment am wenigsten abweichenden Simulation die beiden Simulationen mit gleicher Dicke und den beiden am nächsten liegenden Kompositionen betrachtet.

Für die drei Werte berechnet man jeweils die quadratische Abweichung zum Experiment. Durch diese drei Punkte berechnet man die Parabel. Deren Scheitel markiert die gesuchte Komposition. Kompositionswerte ausserhalb des physikalisch sinnvollen Intervalls $[0, 1]$ werden auf die jeweils näherliegende Intervallgrenze abgebildet.

Bei der Probendickenbestimmung geht man bei dieser Methode analog vor. Für die Interpolation am Rand des in Frage kommenden Dickenintervalls (9-15 Ez.) werden im Gegensatz zur Kompositionsbestimmung die benachbarten Dickenwerte ausserhalb des Intervalls (8 bzw. 16 Ez.) hinzugezogen.

Methode: Minimalabstand Simulation mit Parabelinterpolation		
Kompositionsbestimmung	0.232	23.2%
Dickenbestimmung	1.89 Ez.	15.8%

7.4 Minimalabstand mit Kontrast- und Helligkeitsanpassung

Wie in Abbildung 3.8 schon dargestellt wurde, nimmt der Gesamtkontrast mit steigender Dicke der amorphen Oberflächenschichten ab. Daher wird man beim Vergleich von unverrauschten simulierten Bildern und verrauschten simulierten Bildern keine optimale Übereinstimmung erwarten. Um dieser Abweichung entgegenzuwirken, wird bei

den unverrauschten Simulationen eine Kontrast und Helligkeitsanpassung analog zur Auswertung von experimentellen Mustern durchgeführt. Es werden verrauschte Templatebilder von GaAs und $\text{Al}_{0.4}\text{Ga}_{0.6}\text{As}$ berechnet (Probendicke 12 Ez., Defokus 58 nm). Die get3 Vektoren aus den beiden Templatebereichen werden gemittelt und die Anpassungsparameter zur unverrauschten Simulation berechnet. Wendet man die Methoden aus Abschnitt 7.3 mit einer kontrast- und helligkeitsangepassten Simulation an, so erhält man folgendes Ergebnis:

Methode: Minimalabstand Simulation mit linearer Anpassung		
Kompositionsbestimmung	0.142	14.2%
Dickenbestimmung	3.04 Ez.	25.3%

Methode: Minimalabstand Simulation mit linearer Anpassung und Interpolation		
Kompositionsbestimmung	0.186	18.6%
Dickenbestimmung	3.41 Ez.	28.4%

7.5 Projektion orthogonal zur Hauptkomponente des Rauschens

Die Rauffrequenzfilterung der Abbildung führt zu Korrelation im Rauschen. Geht man von einer idealen Korrelation aus, dann ist die optimale Diskriminierung des Rauschens eine Projektion in die Ebene senkrecht zur Richtung des Rauschens (bei idealer Korrelation wird so das Rauschen vollständig beseitigt).

Die hier durchgeführte Beschreibung von Daten durch wenige Hauptkomponenten bezeichnet man als Karhunen-Loeve Transformation oder *principal component analysis* [25].

Zur näheren Untersuchung des Rauschens wird ein Templatebereich GaAs (Probendicke 12 Ez., Defokus 58 nm, 30Å amorphe Gesamtdicke) berechnet. Von den get3-Vektoren des Bereichs berechnen wir die Korrelationsmatrix. Die Korrelationsmatrix hat folgende Eigenwerte und Eigenvektoren:

$$\begin{aligned}
 w_1 &= 0.00851 & \vec{v}_1 &= (-0.3817, 0.9229, 0.0509) \\
 w_2 &= 0.00300 & \vec{v}_2 &= (0.9124, 0.3674, 0.1806) \\
 w_3 &= 0.00045 & \vec{v}_3 &= (-0.148, -0.1154, 0.9822)
 \end{aligned}$$

Am dominanten Eigenwert w_1 erkennt man, dass eine Vorzugsrichtung im Rauschen auftritt. Der zugehörige Eigenvektor hat nicht verschwindende Komponenten in G_1 und G_2 Richtung. Daher ist die Ursache der Vorzugsrichtung nicht die erhöhte Varianz einer Gruppe sondern vor allem eine Antikorrelation zwischen den Gruppen G_1 und G_2 .

Die Hauptkomponente des Rauschens verläuft in Richtung des Vektors v_1 . Als Projektionsebene für unser Verfahren verwenden wir die Ebene senkrecht zu dieser Richtung. Um mit einem get3-Vektor des Testsatzes eine Kompositions- und Dickenbestimmung durchzuführen, projizieren wir diesen in die Projektionsebene. Jetzt werden alle Simulationen der in 7.3 berechneten get3-Bibliothek in die Ebene projiziert und die Simulation mit geringstem Abstand zum Testmuster bestimmt.

Wir erhalten folgende Ergebnisse:

Methode: Projektion senkrecht zur ersten Hauptkomponente des Rauschens		
Kompositionsbestimmung	0.417	41.7%
Dickenbestimmung	1.81 Ez.	15.1%

Methode: Projektion senkrecht zur ersten Hauptkomponente des Rauschens mit linearer Anpassung		
Kompositionsbestimmung	0.490	49.0%
Dickenbestimmung	3.05 Ez.	25.4%

Das Rauschen ist nicht ideal korreliert. Der zweite Eigenwert w_2 entspricht der maximalen Varianz senkrecht zur ersten Hauptkomponente des Rauschens. Die Varianz senkrecht zur Hauptkomponente kann durch die Projektion nicht reduziert werden. Dagegen wird offensichtlich die gewünschte Kompositions- und Dickeninformation durch die Projektion reduziert. Es folgt ein geringeres Signal-Rausch-Verhältnis und damit verbunden ein drastischer Anstieg in den Fehlern.

7.6 Optimale Projektionsebene

Bei diesem Verfahren vergleichen wir wieder die Projektionen von Testmuster und Simulationen in einer Ebene. Die Projektionsebene wird jetzt aber numerisch optimiert. Die optimale Projektionsebene minimiert nicht das Rauschen wie in Abschnitt 7.5 dargestellt, sondern erreicht ein optimales Signal-Rausch-Verhältnis. In der hier vorgestellten Methode wird die Orientierung der Projektionsebene auf einem gesonderten Mustersatz optimiert. Als Mustersatz zur Optimierung dienen die selben Muster, die

bei der Methode mit neuronalem Netz als Mustersatz (Trainings- und Validierungssatz) verwendet wurden. Es werden für die lokale Probendickenbestimmung und für die lokale Kompositionsbestimmung jeweils gesonderte Projektionsflächen optimiert.

Um eine von der Qualität von numerischen Optimierungsalgorithmen unabhängige Aussage zu treffen, wurden alle Orientierungen der Projektionsebene mit einer Winkelgenauigkeit von 1 Grad getestet. Die optimale Orientierung wurde dann im Bereich ± 1 Grad mit einer Winkelgenauigkeit von 0.01 Grad verfeinert. Diese Vorgehensweise ist sehr zeitaufwendig, hat aber dafür keine Probleme mit lokalen Minima (Methode ist fehleroptimiert nicht rechenzeitoptimiert). Daher sind die hier bestimmten Fehler untere (optimistische) Abschätzungen der Fehler, die mit komplizierteren (schnelleren) numerischen Optimierungsverfahren zu erreichen sind.

Methode: optimale Projektionsebene		
Kompositionsbestimmung	0.091	9.1%
Dickenbestimmung	0.689 Ez.	5.7%

Methode: optimale Projektionsebene mit linearer Anpassung		
Kompositionsbestimmung	0.091	9.1%
Dickenbestimmung	1.22 Ez.	10.2%

7.7 Vergleich der beschriebenen Verfahren

Abbildung 7.1 gibt eine Übersicht über die Ergebnisse (Testfehler im Testszenario, Beschreibung siehe Kapitel 7.1) für die neun untersuchten Verfahren.

Es ergab sich ein Vorteil der Methode mit neuronalem Netz gegenüber dem besten klassischen Verfahren (Abschnitt 7.6) von 2.4 Prozent bei der Kompositionsbestimmung und von 3.2 Prozent bei der Dickenbestimmung. Bei der Kompositionsbestimmung ist der Fehler nur dreiviertelt, bei der Dickenbestimmung nur halb so groß, wie beim besten klassischen Verfahren.

Womit begründet sich der deutliche Vorteil der neuronalen Netze?

Das neuronale Netz könnte je nach Architekturwahl beliebig komplizierte Projektionsflächen leisten. Zusätzlich beschränkt sich die Rauschunterdrückung nicht auf die Projektion auf eine Fläche sondern es kann jedem Raumpunkt des (get3-) Eingaberaumes ein beliebiger Kompositionswert zugeteilt werden. Neuronale Netze vom hier verwendeten Typ sind bei geeigneter Architekturwahl in der Lage, jede beliebig hochdimensionale stetige Funktion beliebig genau zu approximieren [17].

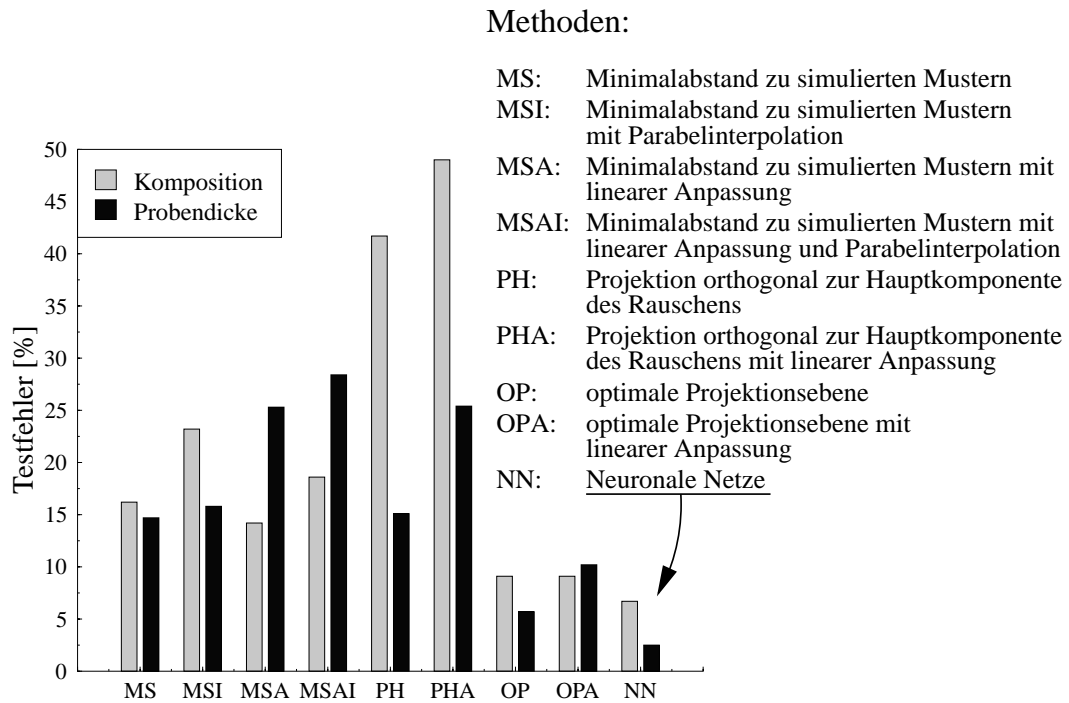


Abbildung 7.1: Fehler der vorgestellten Verfahren zur lokalen Probendicken- und Kompositionsbestimmung

Allerdings ist die große Komplexität der neuronalen Netze nicht alleine das entscheidende Argument. Das hier vorliegende Problem ist die Rekonstruktion von Information aus verrauschten Daten. Dabei ist eine Methode, welche in der Lage ist, die den un-verrauschten Daten zugrundeliegende Funktion genauer zu approximieren, nicht unbedingt im Vorteil. Ein zu komplexes Approximationsmodell folgt zu stark dem Rauschen und ist daher von Nachteil [16]. Diesen Effekt erkennt man auch, wenn man die Methoden dieses Kapitels betrachtet, welche Dreipunkt-Parabelinterpolation verwenden. Die Datenauswertung erfolgt dabei im Vergleich zum reinen Absuchen von Bibliotheken wesentlich feiner. Dies führt aber zu einer größeren Verletzlichkeit gegenüber dem Rauschen in den Daten und damit letztendlich zu höheren Fehlern als bei den nicht-interpolierenden Verfahren.

Bei neuronalen Netzen ist die Modellkomplexität durch die Architekturwahl einstellbar. Erst die Architekturanpassung führt zu den hinreichend bekannten Vorteilen von neuronalen Netzen bei der Rekonstruktion von verrauschter Information [11].

Bei der Vorgehensweise mit neuronalen Netzen können alle Störeinflüsse, welche in der Simulation erfassbar sind, berücksichtigt werden. Insbesondere können auch Störungen berücksichtigt werden, die vom Signal selbst abhängen. Ein Beispiel hierfür ist die Unterdrückung des Einfluss von lokalen Dickenschwankungen bei der Kompositionsbe-

stimmung. Die Berücksichtigung beliebiger Störungen ist ohne vereinfachende Annahmen über den funktionalen Zusammenhang mit dem get3-Vektor möglich. Sämtliche einfließende Näherungen sind Teil der Simulation. Eine Verbesserung der Simulation oder eine Berücksichtigung neuer Störungen führen zu **keinem** neuen Programmieraufwand bei der Auswertung.

Im Gegensatz zur neuronalen Netzen führt bei klassischen Verfahren die Modifikation oder Erweiterung der Eingabedaten zu einer kompletten Reprogrammierung des Verfahrens. Ein Beispiel für eine sinnvolle Ergänzung der get3-Vektoren (Kontrastinformation) wäre das Hinzunehmen von Information über Gitterverspannungen. Eine solche Erweiterung würde es ermöglichen, den Einfluss von Spannungscontrast bei der Kompositionsbestimmung in verspannten Systemen (z.B. InGaAs) zu unterdrücken.

Der Zeitaufwand der Methode mit neuronalen Netzen ist hier unbedeutend. Für die Simulation der Mustersätze benötigt man typischerweise einen Tag. Demgegenüber ist der Aufwand für Training und Architekturwahl vernachlässigbar (PII 400MHz: 1 Stunde). Für zeitkritischere Anwendungen kann durch Verwendung von ausgefeilteren Architektur Anpassungsverfahren [26] und durch die Verkleinerung der verwendeten Mustersätze der numerische Aufwand erheblich reduziert werden.

Kapitel 8

Experimentelle Anwendung

8.1 AlGaAs Grenzschicht

Als Anwendung der vorgestellten Verfahren, wird in diesem Kapitel eine AlGaAs – GaAs Grenzschicht untersucht. Die Probe wurde am Ioffe Institut in St. Petersburg durch Molekularstrahlepitaxie hergestellt. Bei der Probenpräparation erfolgte das Ionenabdünnen mit Ar^+ Ionen der Energie 4 keV (s. Kap. 3.1).

Abbildung 8.1 zeigt die untersuchte HREM-Aufnahme, die von P. Werner (MPI Mikrostrukturphysik Halle) angefertigt wurde. Die Defokussierung wurde bei der Aufnahme als etwas über Scherzerfokus (JEM-4000EX: 50nm) abgeschätzt.

Die Herstellerangaben für die Komposition in den beiden Templatebereichen der $\text{Ga}_x\text{Al}_{1-x}\text{As}$ Probe lauteten:

linke Seite	rechte Seite
$x=0.6$	$x=1.0$

Der Wert für das binäre Halbleitermaterial GaAs kann als exakt angenommen werden. Der Fehler im Wert für das ternäre Material AlGaAs liegt nach Herstellerangabe zwischen 5 und 10 Prozent.

8.1.1 Mittlere Template Parameter

Für die dargestellte HREM-Abbildung (Abb. 8.1) wurde die Zelldetektion durchgeführt (s. Abschnitt 4.2). Abbildung 8.2 zeigt die Positionen zur Helligkeitsauswertung für die get3-Bildvorverarbeitung (s. Abschnitt 4.3) und Abbildung 8.3 die ausgewählten Templatebereiche (s. Abschnitt 5.1).

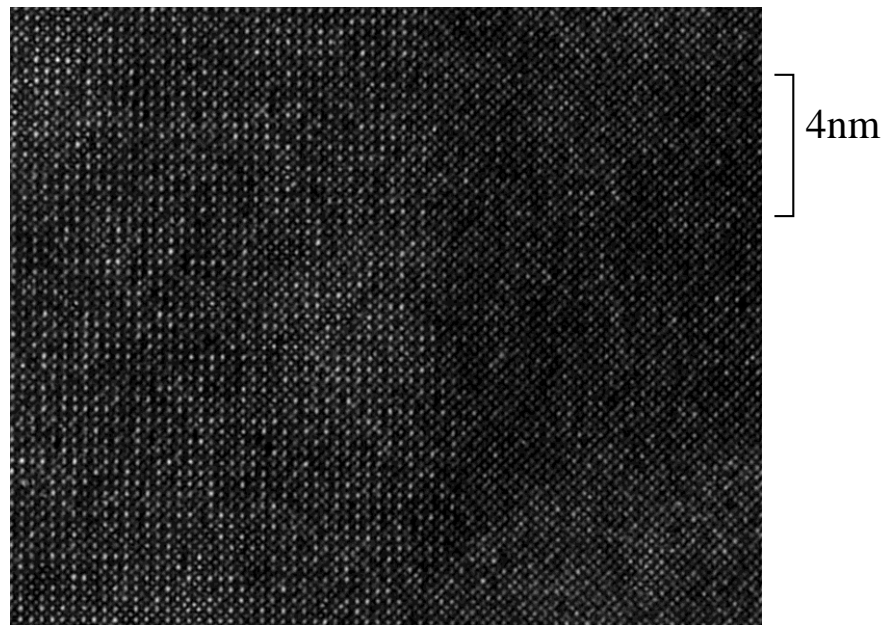


Abbildung 8.1: HREM-Aufnahme einer AlGaAs-GaAs-Grenzschicht.

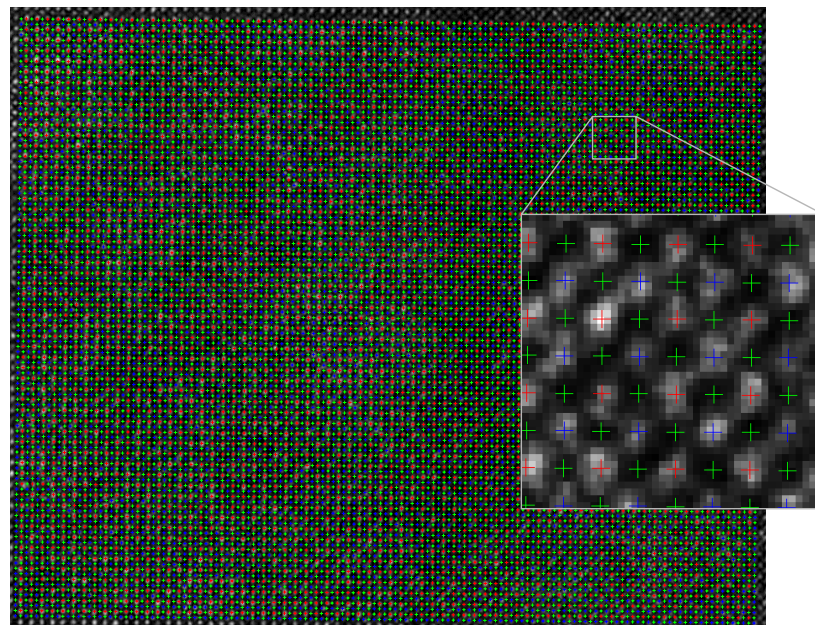


Abbildung 8.2: Ausgewerteter Bereich mit Positionen für Helligkeitsbestimmung (siehe Abschnitt 4.3). Farbe der Positionsmarkierungen wie in Abb. 4.1 und Abb. 4.2.

Die Parametersuche wurde in einem Dickenbereich von 5 bis 40 Elementarzellen (Sampling: 1 Ez.), das entspricht 28.3 bis 226.4Å Schichtdicke, durchgeführt. Schichtdickenunterschiede zwischen den beiden Templatebereichen von bis zu acht Elementarzellen (45.3Å) wurden in Betracht gezogen. Den in Frage kommenden Defokussierungsbereich (10–100nm) untersucht man in Schritten von 1 nm. Der gesamte Kompositionsbereich

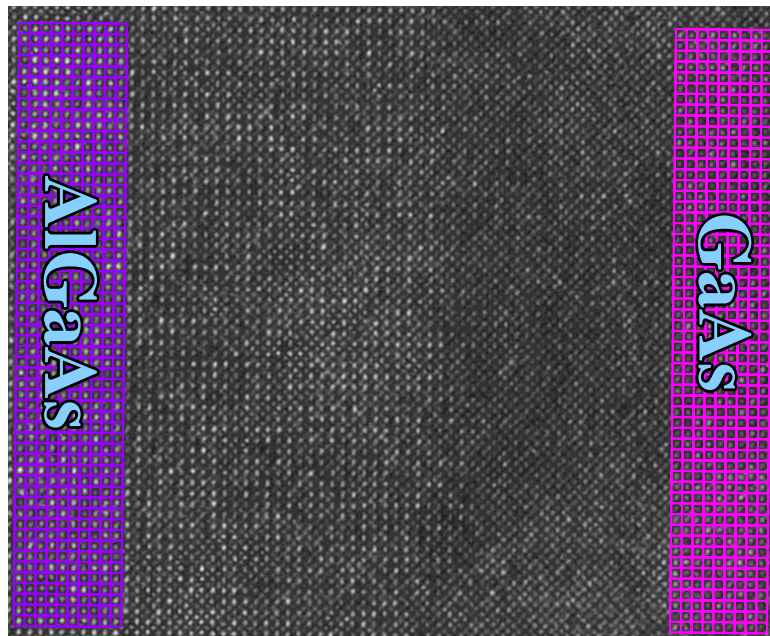


Abbildung 8.3: Für die Parametersuche verwendete Templatebereiche.

($x=0$ bis $x=1.0$) wurde in 32 Schritte geteilt. Als Vorabwissen über die Bildentstehung wurden die Kompositionen beider Templatebereiche auf 10 Prozent bekannt angenommen. Dabei wurde von der Fehlerangabe des Probenherstellers die Obergrenze des Fehlers gewählt. Trotz des auf diese Weise wenig eingeschränkten Parameterraumes war das Ergebnis der Parametersuche eindeutig. Es resultierten folgende mittlere Templateparameter:

	linker Templatebereich	rechter Templatebereich
Komposition $\text{Al}_{1-x}\text{Ga}_x\text{As}$	$x_a = 0.65$	$x_b = 1.0$
kristalline Probendicke	$t_a = 15$ Ez.	$t_b = 12$ Ez.
Defokussierung	$Z = 58\text{nm}$	

Die Defokussierung liegt in dem aus den Experimentatorangaben erwarteten Bereich. Durch die großen verwendeten Templatebereiche ist die Ungenauigkeit (5.21) in den mittleren get3-Vektoren sehr klein. Daher ergab sich ein eindeutiges Minimum bei den oben dargestellten Parametern.

Mit den mittleren Parametern für die Templatebereiche führen wir die in Abschnitt 3.3 beschriebene amorphe Schichtdickenbestimmung durch. Für die experimentelle Auswertung wurde ein quadratischer Ausschnitt aus dem GaAs Bereich gewählt. Dieser lieferte den Wert

$$\frac{I_N}{I_R} = 0.37$$

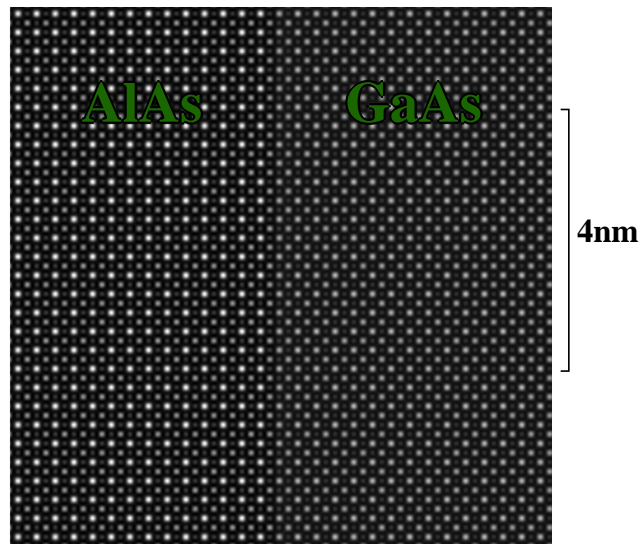


Abbildung 8.4: Mit den mittleren Parametern für die Templatebereiche simulierte Grenzfläche.

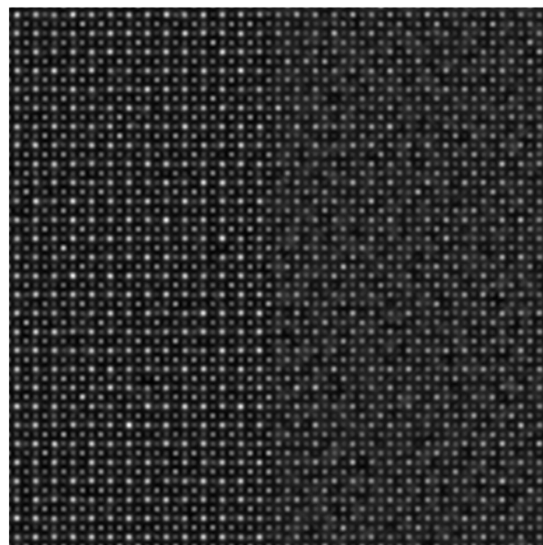


Abbildung 8.5: Simulierte Grenzfläche mit 30 Å Gesamtdicke der amorphen Oberflächenschichten.

Im Vergleich zu simulierten Proben (siehe Abschnitt 3.3) entspricht das einer Gesamtdicke der amorphen Oberflächenschichten von 32 Å.

Abbildung 8.4 zeigt eine simulierte Grenzfläche mit den Parametern des untersuchten experimentellen Beispiels. Simuliert man zusätzlich zu dem kristallinen Bereich amorphe Oberflächenschichten der Gesamtdicke 32 Å so erhält man die in Abb. 8.5 dargestellte simulierte verrauschte HREM-Aufnahme.

8.1.2 Kompositions- und Dickenbestimmung

Für die lateral aufgelöste Kompositions- und Dickenbestimmung wurde ein Mustersatz im vollen Kompositionsbereich ($x \in [0, 1]$, 32 Schritte) und im Dickenbereich 9 bis 17 Elementarzellen berechnet. Aus jeder Parameterkombination (x, t) wurden vier verrauschte Beispiele berechnet. Das Rauschen wurde durch Oberflächenschichten der Gesamtdicke 32 \AA erzeugt. Das Material der Oberflächenschichten wurde als Objekt mit zufälligen Atompositionen (Abschnitt 3.2.2) und der Zusammensetzung entsprechend dem darunterliegenden kristallinen Material berechnet. Dreiviertel der Beispiele im Mustersatz wurden als Trainingssatz der Rest der Muster als Validierungssatz (s. Kap. 6.4) verwendet. Das Netztraining wurde mit den Architekturgrößen $N_h = 2$ bis $N_h = 30$ jeweils zehnmal durchgeführt. Das Netz mit dem kleinsten Fehler auf dem Testsatz hatte für die Kompositionsbestimmung eine Architekturgröße von 17 Zwischenschichtneuronen. Bei der Dickenbestimmung ergab sich $N_h = 15$ als Optimum.

Zur Bestimmung der Fehler des Verfahrens wurde ein Testsatz mit größerem Dickenbereich (6–26 Ez.) und nach Vorabwissen eingeschränktem Kompositionsbereich ($x \in [0.125, 1]$ in $\text{Al}_{1-x}\text{Ga}_x\text{As}$) gerechnet. Bei der später durchgeführten Dickenbestimmung ergaben sich Dicken von 9 bis 19 Ez. Dieser Bereich ist für die Fehlerbestimmung ausreichend, da die Grenzen dieses Bereichs nur sehr selten erreicht wurden ($t < 9.5$: 3 Werte und $t > 18.5$: 6 Werte von insgesamt 3864 Werten).

Der mittlere Fehler für die lokale Kompositionsbestimmung war (im Dickenbereich 9 bis 19 Ez.) 5.9 Prozent. Der größte Fehler (8.96 Prozent) trat bei einer kristallinen Probendicke von 9 Elementarzellen auf. Abbildung 8.6 stellt den Fehler der lokalen Kompositionsbestimmung in Abhängigkeit von der Dicke des kristallinen Teils der Probe dar.

Die lateral aufgelöste Bestimmung der kristallinen Probendicke erreichte auf dem gesamten Kompositionsbereich einen mittleren Fehler von 4.3 Prozent der mittleren Dicke (14 Ez.). Das entspricht einem absoluten Fehler von 0.6 Elementarzellen. Der größte Fehler der Probendickenbestimmung (6.8 Prozent) trat bei der Komposition $x = 0.125$ auf. Abbildung 8.7 ermöglicht die kompositionsabhängige Einschätzung der Fehler.

In Abbildung 8.8 ist das Ergebnis der durchgeführten Kompositions- und Dickenbestimmung als Farbbild dargestellt. Die Auflösung der Darstellung entspricht einem Viertel der Fläche der Elementarzelle ($2.8 \times 2.8 \text{ \AA}^2$). Zusätzlich zur Farbdarstellung der Komposition ist bei dieser Abbildung die kristalline Dicke als Höhe in einem Säulendiagramm wiedergegeben.

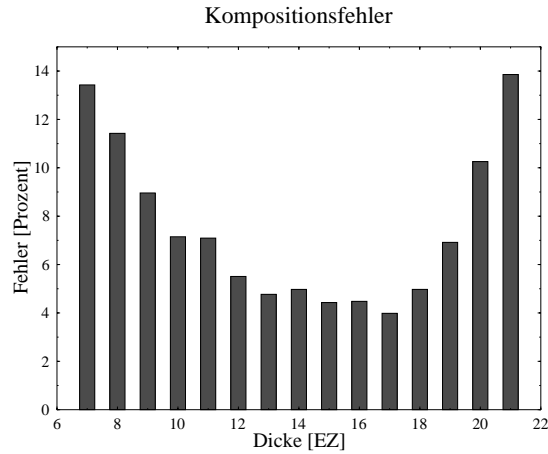


Abbildung 8.6: Fehler der lokalen Kompositionsbestimmung in Abhängigkeit von der kristallinen Probendicke.

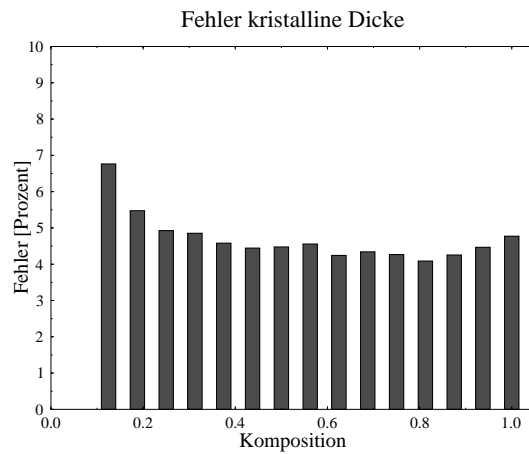


Abbildung 8.7: Kompositionsabhängigkeit des Fehlers der ortsaufgelösten Bestimmung der kristallinen Probendicke.

8.1.3 Diskussion der Kompositionsverteilung

Random-Alloy-Fluktuationen

Es fällt auf, dass die Komposition im AlGaAs Bereich signifikanten Schwankungen unterliegt. Die Mittelwerte x_m und Standardabweichungen σ_x der Komposition in den beiden Templatebereichen sind:

$$x_m(\text{AlGaAs}) = 0.656 \quad \sigma_x(\text{AlGaAs}) = 0.0794 \quad (8.1)$$

$$x_m(\text{GaAs}) = 0.964 \quad \sigma_x(\text{GaAs}) = 0.0262 \quad (8.2)$$

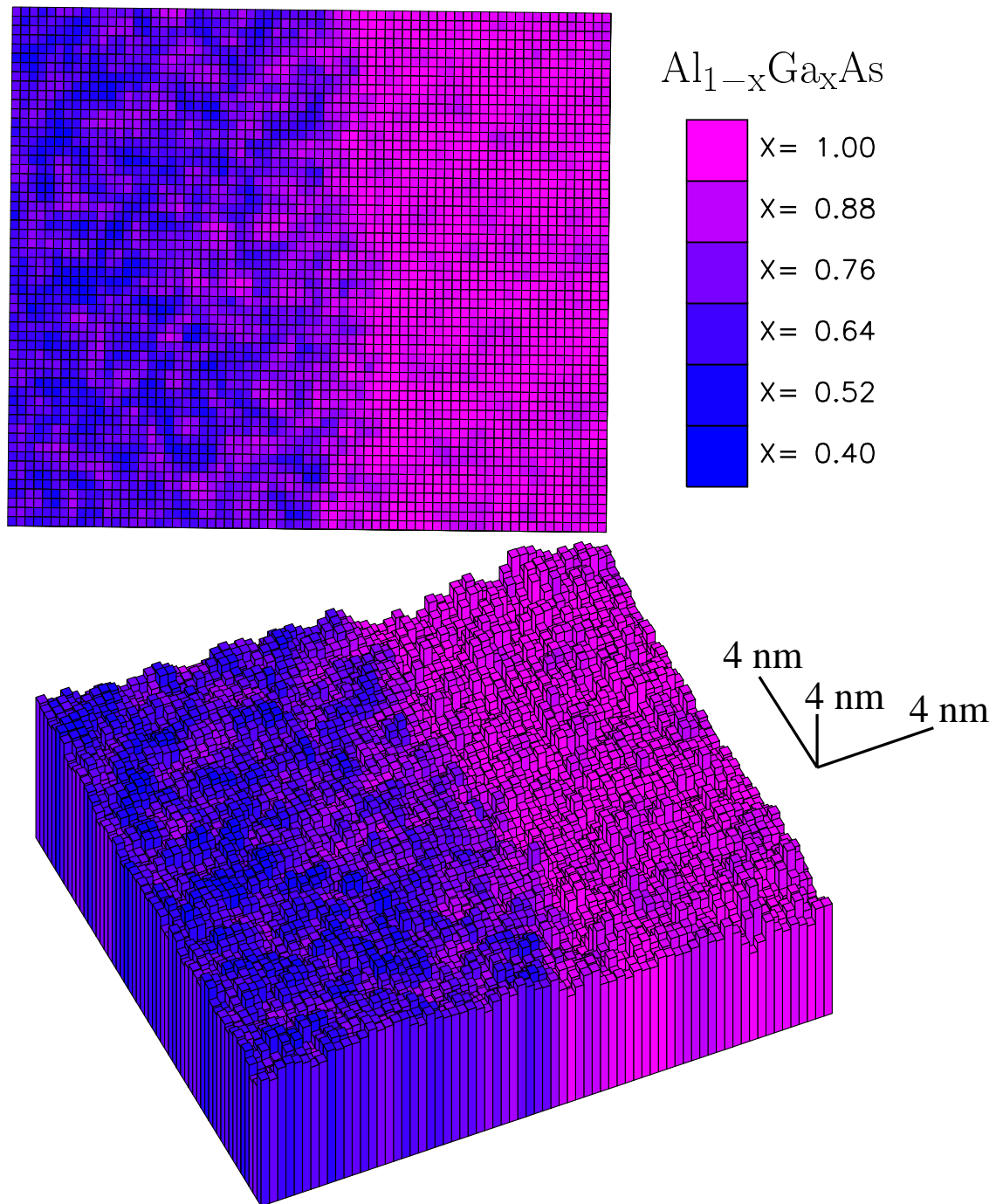


Abbildung 8.8: Ergebnis der Kompositions- und Dickenbestimmung.

Im ternären Material schwankt die Komposition in Folge von Random-Alloy-Fluktuationen. Beim Wachstum des ternären Materials wird das Al/Ga Untergitter der Sphaleritstruktur mit beiden Atomsorten statistisch besetzt. An einem Modell soll das näher untersucht werden:

Wir betrachten die Atomsäulen des Al/Ga Untergitters. Liegt ein Material der mittleren Komposition $\text{Al}_{1-x}\text{Ga}_x\text{As}$ vor, so ist die Wahrscheinlichkeit dafür, dass eine Gitterposition mit Aluminium besetzt wird

$$p = 1 - x \quad (8.3)$$

Die in (8.1) und (8.2) angegebenen statistischen Parameter der Komposition wurden ohne Überlappung der get3-Auswertebereiche bestimmt, die Flächenauflösung entspricht daher der Elementarzelle. Eine Elementarzelle enthält vier Atome des Al/Ga Untergitters. Betrachten wir eine Probendicke von t Elementarzellen. Der Bildbereich einer Elementarzelle beinhaltet $4t$ Atome¹. Die Besetzung der Positionen des Untergitters stellt einen Prozess mit Binomialverteilung dar. Die Komposition des ausgewerteten Bereichs (Flächenauflösung 1 Ez.) hat damit eine Standardabweichung von

$$\sigma_x = \frac{1}{2} \sqrt{\frac{p(1-p)}{t}} \quad (8.4)$$

Eine Berechnung mit den Parametern für die vorliegende Probe ergab eine erwartete Standardabweichung im AlGaAs Bereich ($x = 0.656$, $t = 16$ Ez.) von

$$\sigma_x = 0.0594 \quad (8.5)$$

Bei der MBE-Kristallzüchtung ist das Wachsen eines reinen GaAs Bereichs unproblematisch. Dazu wird einfach die Aluminiumzufuhr während des Wachstumsprozesses vollständig unterbrochen. Eine Aluminiumverunreinigung des GaAs ist sehr unwahrscheinlich [47]. Die Schwankungen der Komposition im GaAs Templatebereich sind ausschliesslich fehlerbedingt.

Die Kompositionsschwankungen im AlGaAs Templatebereich ergeben sich durch Überlagerung der Random-Alloy-Fluktuation mit den fehlerbedingten Schwankungen. Nehmen wir beide Einflüsse als statistisch unabhängig und additiv überlagert an, so addieren sich die Standardabweichungen. Als Abschätzung für den Fehlereinfluss im AlGaAs Bereich nehmen wir die Standardabweichung der Komposition im GaAs Templatebereich (8.2). Es ergibt sich eine Standardabweichung von:

$$0.0262 + 0.0594 = 0.0856 \quad (8.6)$$

Dies ist in ausgezeichneter Übereinstimmung mit dem experimentell bestimmten Wert von 0.0794.

¹Bildet das Al/Ga Untergitter die Oberfläche der Probe, so sind in einer Atomsäule eigentlich $4(t+1)$ Atome. Der Unterschied ist jedoch für die hier betrachteten Schichtdicken vernachlässigbar.

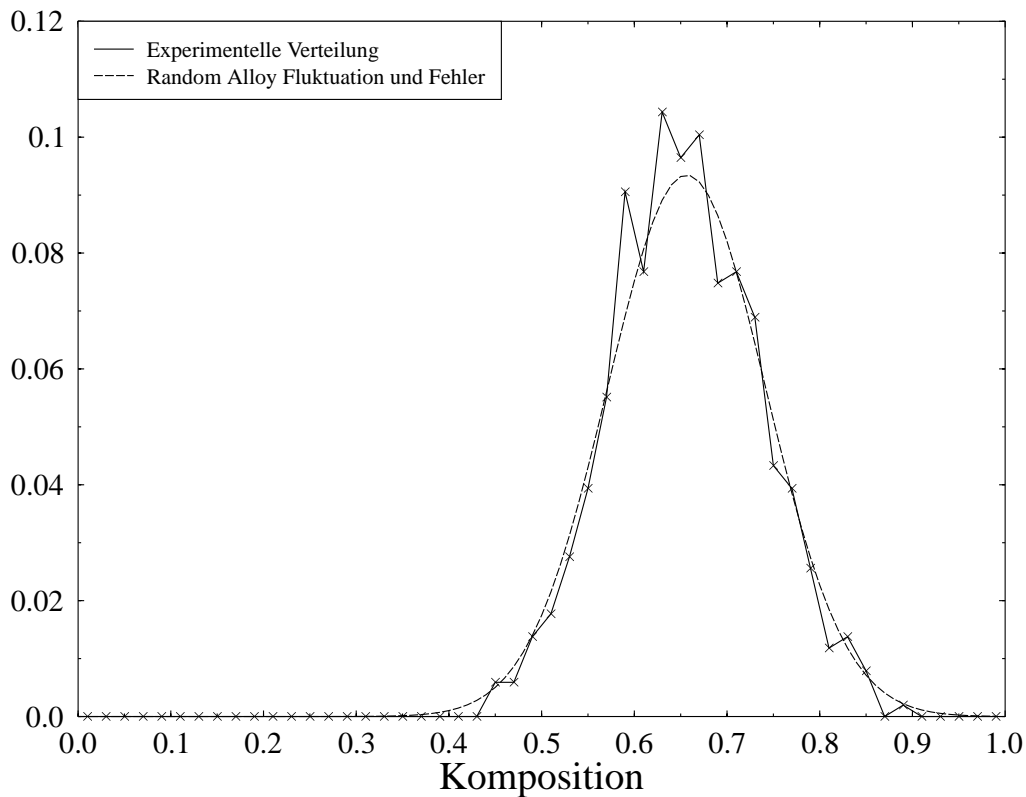


Abbildung 8.9: Verteilung der Komposition im AlGaAs Bereich

Abbildung 8.9 vergleicht die Verteilung der Komposition im AlGaAs Templatebereich des Experiments (durchgezogene Linie) mit der Verteilung in unserem Modell (gestrichelte Linie, Die Binomialverteilung ist als Normalverteilung genähert). Das grobe Modell für die Random-Alloy-Fluktuation zeigt eine hervorragende Übereinstimmung zum Experiment.

Mittelung parallel zur Grenzfläche

Abbildung 8.10 zeigt das Ergebnis einer Mittelung der Komposition parallel zur Grenzfläche. Man erkennt deutlich die erhöhte Kompositionsschwankung im AlGaAs Bereich. Innerhalb des AlGaAs Bereichs tritt in Richtung des Interface ein langsames (signifikantes) Ansteigen der Ga Konzentration auf. Das Profil der Grenzfläche ist deutlich zu erkennen. Der GaAs Bereich hat eine weitgehend konstante Komposition von 0.964. Dies liegt im Rahmen des Fehlers (4.5 % im Dickenbereich des GaAs Templatebereichs) bei dem erwarteten Wert von 1.0.

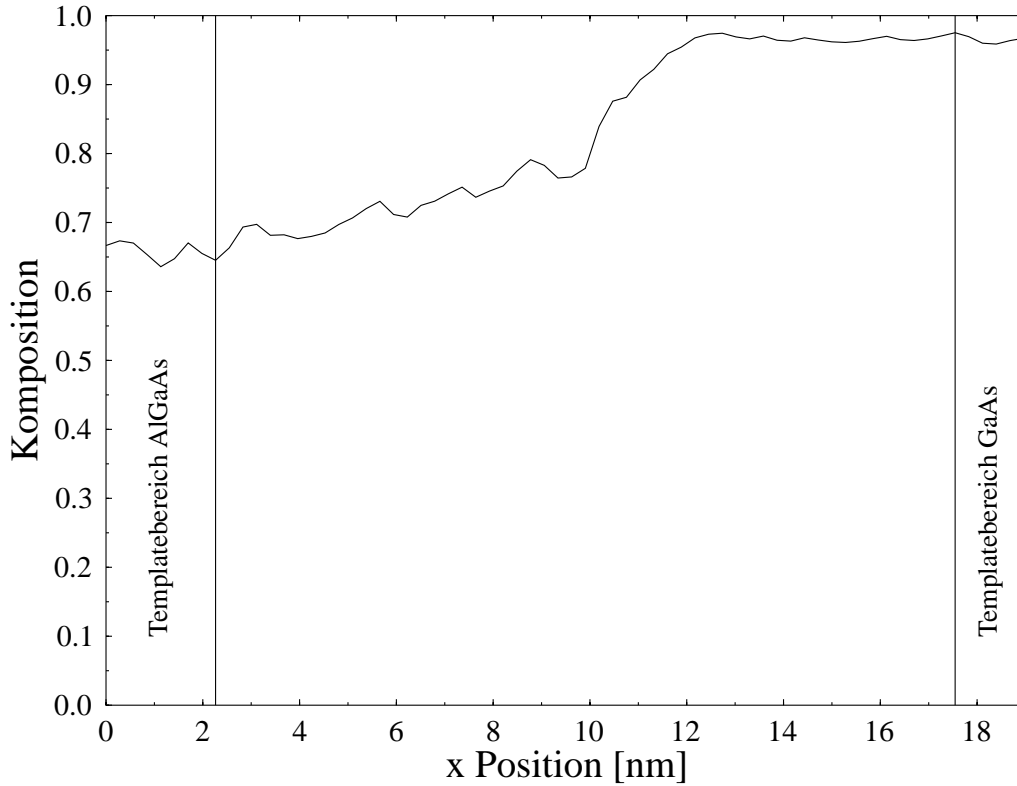


Abbildung 8.10: Komposition parallel zum Interface gemittelt

8.1.4 Diskussion des Dickenverlaufs

Bei der orts aufgelösten Bestimmung der Dicke des kristallinen Teils der Probe ergaben sich in den Templatebereichen folgende Mittelwerte t_m und Standardabweichungen σ_t :

$$\begin{aligned}
 t_m(\text{AlGaAs}) &= 16.3 \text{ Ez.} & \sigma_t(\text{AlGaAs}) &= 0.95 \text{ Ez.} \\
 t_m(\text{GaAs}) &= 12.7 \text{ Ez.} & \sigma_t(\text{GaAs}) &= 1.38 \text{ Ez.}
 \end{aligned}$$

Im Gegensatz zur Kompositionsverteilung, ergeben sich hier deutlich mehr Schwankungen im GaAs als im AlGaAs Bereich. Der kristalline Teil der Probe ist von amorphen Oberflächenschichten umgeben. Diese sind auf das Ionenabdünnen der Probe (HREM-Probenpräparation) zurückzuführen (s. Kap. 3.1). Die Tiefe der Amorphisierung schwankt lateral auf beiden Seiten der Probe (symmetrische Verteilung der Oberflächenschichten s. Kap. 3.3.2). Als Überlagerung ergeben sich laterale Schwankungen der amorphen Gesamtdicke und damit auch bei gleichbleibender Gesamtdicke der Probe laterale Schwankungen in der Dicke des kristallinen Teils der Probe. Es ist hinlänglich bekannt [9], dass AlAs durch den Ionenbeschuss wesentlich weniger geschädigt wird als GaAs. Dies erklärt die wesentlich höheren Schwankungen der Dicke des kristallinen Teils der Probe im GaAs Bereich.

8.2 AlGaAs-Bragg-Reflektor

Abbildung 8.11 zeigt einen Ausschnitt einer HREM-Aufnahme eines AlGaAs-Bragg-Reflektors. Bei der Produktion dieser Heterostruktur mit Hilfe von Molekularstrahl-epitaxie (Ioffe Institut, St. Petersburg) wurden reine AlAs und GaAs Schichtbereiche angestrebt.

Bei der Probenpräparation wurden wiederum 4 keV Ar^+ Ionen zum Absputtern verwendet. Die untersuchte Hochauflösungsaufnahme wurde von P. Werner vom MPI Mikrostrukturphysik Halle angefertigt.

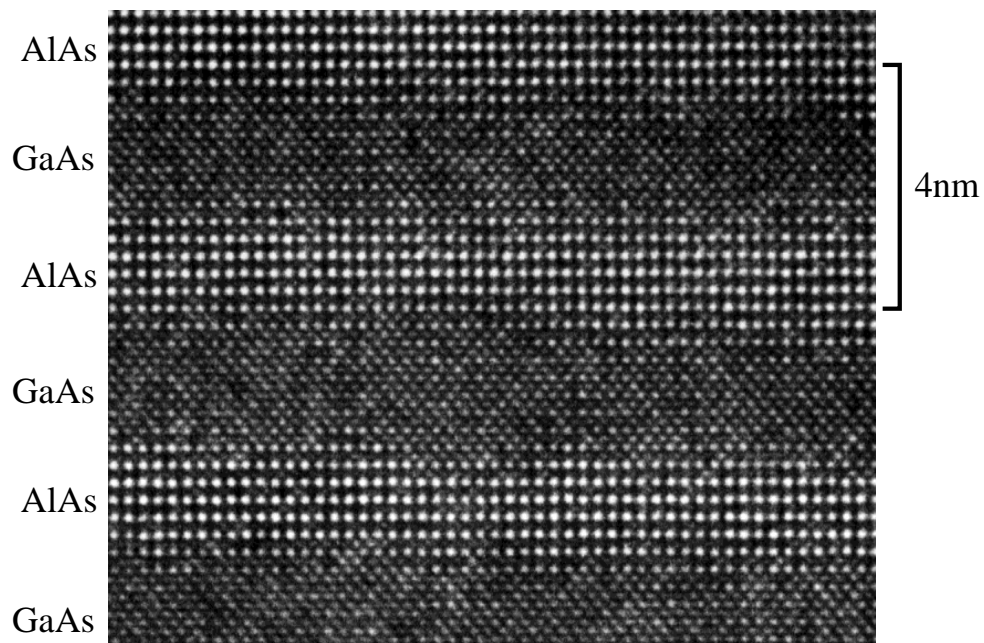


Abbildung 8.11: HREM-Aufnahme eines AlGaAs-Bragg-Reflektors

Bei der Parameterbestimmung in den Templatebereichen ergab sich eine Defokussierung von 57 nm. Die Bestimmung der amorphen Oberflächenschichtdicke wurde an einem der Struktur nahe liegenden GaAs Bulk-Bereich durchgeführt. Die Gesamtdicke der amorphen Oberflächenschichten war 23 Å.

Die Kompositions- und Dickenbestimmung ergab das in Abbildung 8.12 dargestellte Ergebnis. Der Fehler bei der Kompositionsbestimmung lag bei 4.8 Prozent, der Fehler bei der Dickenbestimmung bei 0.78 Elementarzellen entsprechend 5.2 Prozent der mittleren Dicke (15 Ez.).

Innerhalb der einzelnen Schichten kommt es im Vergleich zu der in Abschnitt 8.1 ausgewerteten Aufnahme zu weitaus weniger Schwankungen der Komposition. In den

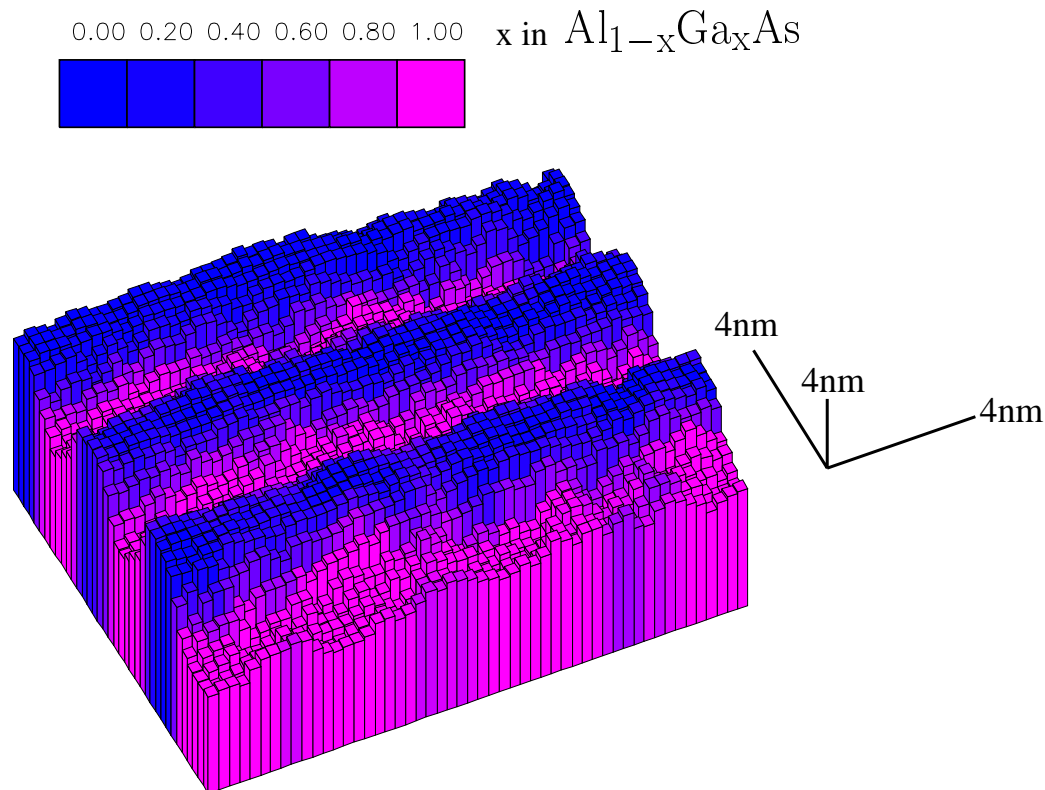


Abbildung 8.12: Kompositions- und Dickenbestimmung

(annähernd) reinen binären Materialien der Schicht kommt es nicht zu ausgeprägten Random-Alloy-Fluktuationen.

Auffällig ist die starke Überhöhung (größere Dicke des kristallinen Teils der Probe) der AlAs Bereiche im Vergleich zu den GaAs Bereichen. Durch das Ionenabdünnen während der Probenpräparation resultiert eine Amorphisierung der Probenoberfläche. Da AlAs gegenüber dem Ionenbeschuss weitaus stabiler ist als GaAs [9], erzeugt das Abdünnen die dargestellten Schwankungen in der kristallinen Probendicke. Im Gegensatz zu der in Abschnitt 8.1 dargestellten Probe liegen hier im zentralen Bereich der jeweiligen Schicht die reinen binären Halbleitermaterialien vor. Daher ergibt sich auch eine wesentlich deutlichere Überhöhung der AlAs Bereiche.

Um Dicken- und Kompositionsverlauf quantitativ auswerten zu können, führen wir eine (in Abb. 8.11 horizontale) Mittelung parallel zu den Schichten durch.

Abbildung 8.13 zeigt die horizontal gemittelte Komposition. Die (in Abb. 8.11 vertikal) lateral dickeren GaAs Schichten haben zentral die von der Probenpräparation vorhergesagte Komposition. Bei den lateral dünneren AlAs Schichten kommt es zu einer signifikanten Abweichung (8–15%) von den Herstellerangaben. Dies kommt einerseits durch eine Mittelung über die Grenzflächenrauigkeit zu Stande. Andererseits

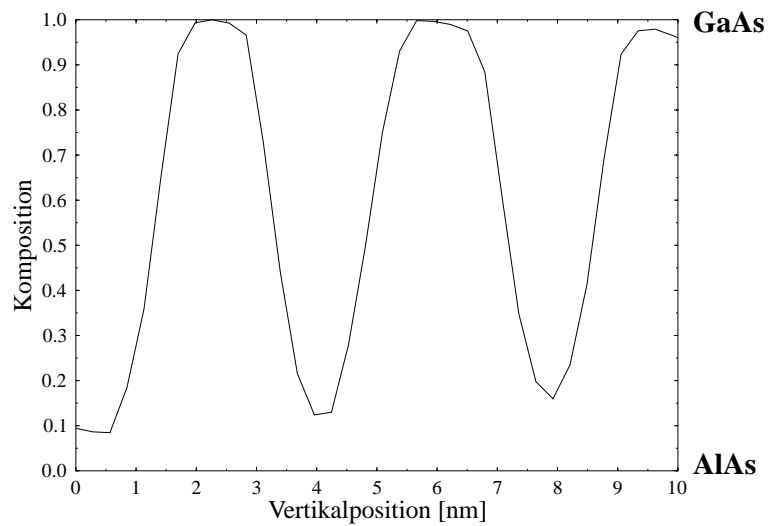


Abbildung 8.13: Horizontalmittel der Komposition

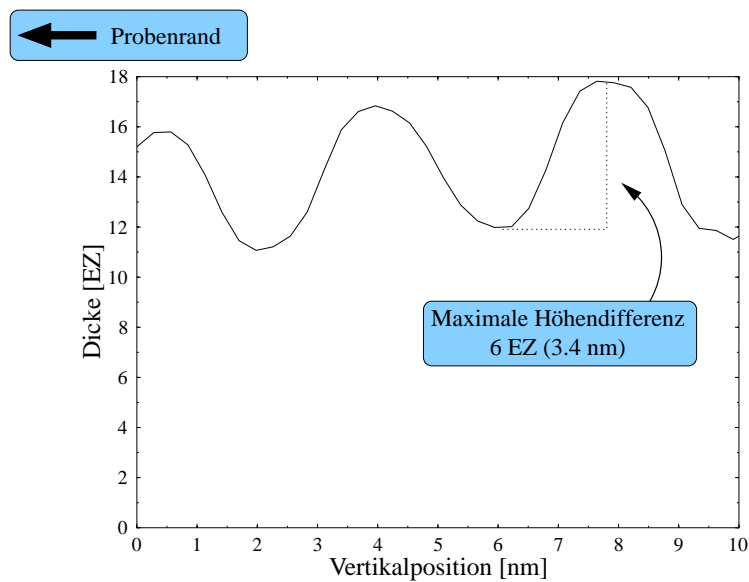


Abbildung 8.14: Horizontalmittel der Dicke des kristallinen Teils der Probe

liegt die Dicke der reinen AlAs Bereiche schon an der Grenze der lateralen Auflösung des Verfahrens (5.9 \AA siehe unten).

In Abbildung 8.14 ist die horizontal gemittelte Dicke dargestellt. Fünf Nanometer vom rechten Rand des ausgewerteten Bereichs liegt der Rand der Probe. Am gemittelten Dickenprofil erkennt man klar den Abfall der Dicke zum Probenrand hin. Aus der Probenpräparation resultiert die größere Dicke der AlAs Schichten im Vergleich zu

den GaAs Schichten. Der maximale Unterschied der beiden Schichtdicken ist im Diagramm markiert. Der GaAs Bereich ist dabei 12 Elementarzellen, der AlAs Bereich ist 18 Elementarzellen dick. Es liegt also ein Unterschied in der Schichtdicke von 6 Elementarzellen und damit von 33.96 Å vor.

Eine naheliegende Frage bei diesem Ergebnis ist: “Würde man diese Veränderung in der Probendicke nicht durch Fresnelstreifen schon in der HREM-Abbildung erkennen?”.

Es wurde in einem *worst-case*-Szenario eine Dickenstufe simuliert, bei der sich die Probendicke abrupt von 12 auf 18 Elementarzellen und die Komposition abrupt zwischen den beiden Extremwerten (AlAs/GaAs) ändert. Die Kompositions- und Dickenänderung erfolgt bei der simulierten HREM-Aufnahme im Gegensatz zur untersuchten Probe (vgl. Abb. 8.14 und Abb. 8.13) abrupt. Durch die abrupte Veränderung wird die Auswirkung von Beugungseffekten (Fresnelstreifen) nach oben abgeschätzt (*worst-case*-Szenario). Die Simulation wurde ohne amorphe Oberflächenschichten berechnet, um eine ungestörte Untersuchung von Beugungseffekten und Delokalisierung zu ermöglichen.



Abbildung 8.15: Simulierte Grenzfläche GaAs/AlAs mit Höhenunterschied 6 Ez.

Abbildung 8.15 zeigt die simulierte HREM-Abbildung. Offensichtlich besteht ohne die quantitative Auswertung keinerlei Hinweis auf den vorliegenden Dickenunterschied. Führt man mit der simulierten Aufnahme eine Kompositions- und Dickenbestimmung durch, so ergibt sich das in Abbildung 8.16 dargestellte Ergebnis.

Zur quantitativen Auswertung betrachten wir das horizontale Profil der Werte aus der Bestimmung der Dicke des kristallinen Teils der Probe (Abb. 8.17). Die abrupte

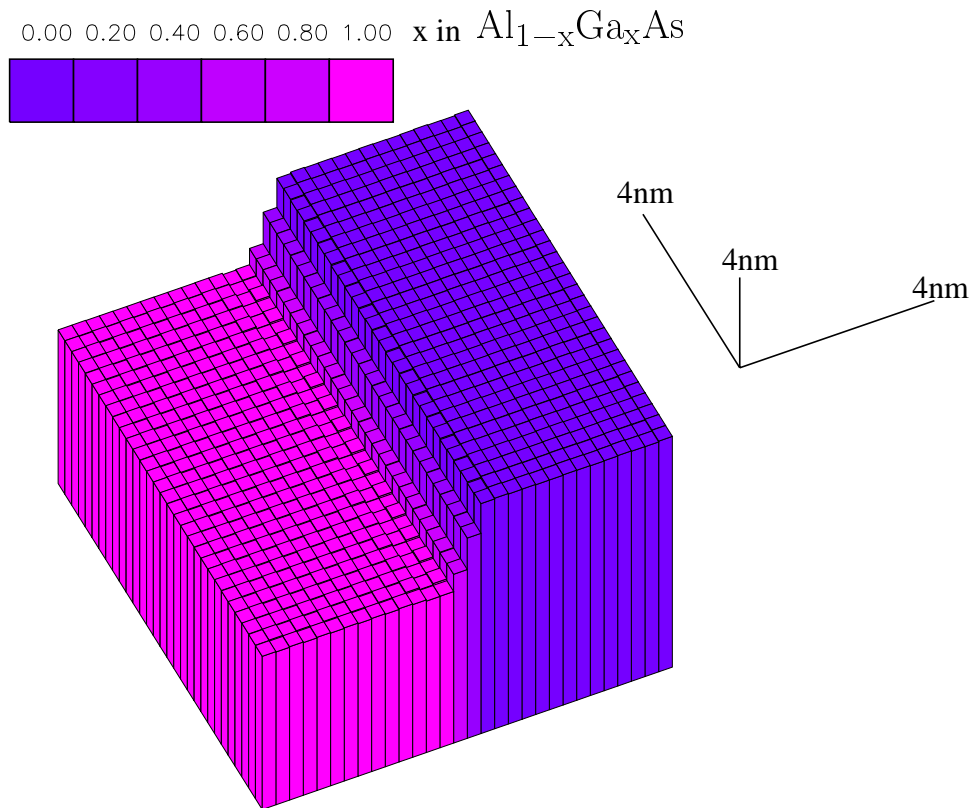


Abbildung 8.16: Kompositions- und Dickenbestimmung der simulierten Grenzfläche.

Dickenänderung an der Grenzfläche der Simulation wird bei der lokalen Dickenbestimmung auf einen Bereich verteilt wiedergegeben. In Abbildung 8.17 eingezeichnet, sind die Grenzen für fünf Prozent Abweichung von der real vorliegenden Dicke. Innerhalb dieser Grenzen beeinflusst die simulierte Grenzfläche die von dem Verfahren bestimmten Werte. Daher ist die Größe des dargestellten Bereichs ein Maß für die laterale Auflösung der besprochenen Methoden. Bei den hier vorliegenden Abbildungsbedingungen ergibt sich eine laterale Auflösung von 5.9\AA .

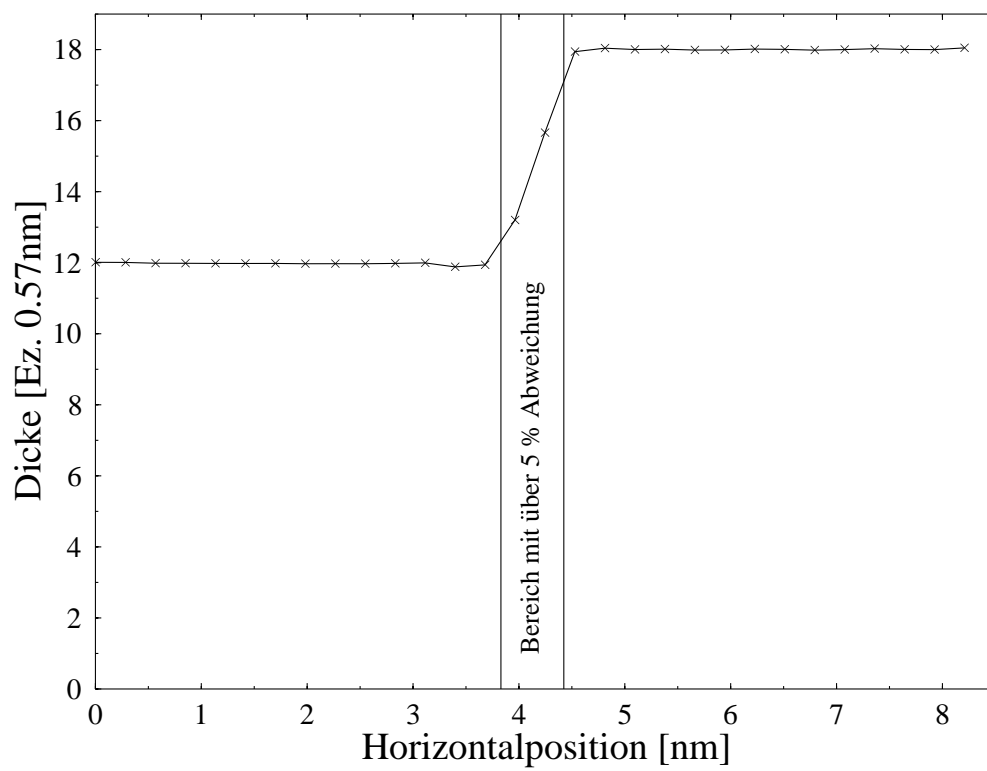


Abbildung 8.17: Horizontalprofil der Werte aus der Dickenbestimmung an der simulierten Grenzfläche. Eingezeichnet ist der Bereich mit mehr als 5% Abweichung von der lokal gegebenen Probendicke.

Kapitel 9

Zusammenfassung

In dieser Arbeit wurde eine neue Methode der orts aufgelösten Kompositionsbestimmung für Materialien mit Sphaleritstruktur in (001)-Projektion entwickelt. Die Methode basiert auf der quantitativen Auswertung von HREM-Aufnahmen und leistet (mit dem JEM-4000EX) eine laterale Auflösung die etwa der Größe der Elementarzelle entspricht (Kap. 8.2). Zusätzlich zur Kompositionsinformation wird die lokale Dicke des kristallinen Teils der Probe bestimmt. Eine notwendige Trennung zwischen Kompositions- und Dickeninformation wird durch den Einsatz neuronaler Netze fehleroptimiert gewährleistet.

Spezielle Aufmerksamkeit wurde den Auswirkungen von Amorphisierung der Probenoberfläche gewidmet. Solche amorphe Oberflächenschichten werden vor allem durch die HREM-Probenpräparation verursacht. Abschnitt 3.2.2 beschreibt die Simulation von amorphen Schichten als Objekte mit gleichverteilten Atomkoordinaten. Wie in Kapitel 3.2.1 dargestellt, ist dies für die vorliegenden Abbildungsbedingungen eine gute Näherung. Es wurde gezeigt, dass sich die Phasengitterfunktion der amorphen Schicht unter Verwendung der atomaren Streuamplituden nach Doyle und Turner teilanalytisch berechnen lässt. Dies führt im Vergleich zur direkten Berechnung mit dem Programmpaket EMS zu einem um den Faktor 12 geringeren Zeitbedarf.

Durch Vergleich mit simulierten Abbildungen amorpher Schichten kann über eine neu entwickelte Diffraktogrammanalyse (Kap. 3.3) die Amorphisierungstiefe von HREM-Proben bestimmt werden. Dies erlaubt eine Beurteilung von HREM-Probenpräparationsverfahren. Bei den zu dieser Arbeit untersuchten, ionenabgedünnten Proben streuten die Amorphisierungstiefen zwischen 20 und 40 Å (Gesamtdicke beider Oberflächenschichten).

In Abschnitt 3.4 wurde für experimentell auftretende Amorphisierungstiefen eine Kontrastabschwächung von 10–20 Prozent errechnet. Der Bildkontrast wird dabei zu 25–40

Prozent vom Rauschen bestimmt (Kap. 3.4). Daher ist für eine quantitative Bildauswertung an HREM-Aufnahmen die Berücksichtigung von amorphen Oberflächenschichten von entscheidender Wichtigkeit.

Quantitative Bildauswertung an HREM-Aufnahmen setzt eine Verwendung von Vorwissen über die Bildentstehung voraus. Die Information über die Bildentstehung wird im Rahmen von sog. Bildvorverarbeitungsverfahren berücksichtigt.

Die verschiedenen Möglichkeiten der Bildvorverarbeitung lassen sich aufteilen in realspacebasierte Verfahren und Fouriermethoden. Aus jeder der beiden Kategorien lässt sich durch prinzipielle Argumente jeweils ein optimiertes Verfahren konstruieren (Kap. 4). Die beiden Verfahren (lokale Fouriertransformierte und get3-Vorverarbeitung) wurden hinsichtlich Kompositions- und Probendickenbestimmung getestet (Kap. 5.4). Es zeigte sich, dass die lateral aperiodische Störung durch amorphe Oberflächenschichten die lokale Fouriertransformierte wesentlich stärker stört. Daher wurde bei der experimentellen Anwendung der Kompositions- und Dickenbestimmung (Kap. 8) das realspacebasierte Verfahren (get3-Vorverarbeitung) verwendet.

In experimentellen HREM-Aufnahmen lassen sich Bereiche finden, in denen eine Mittelung über die orts aufgelösten Werte aus der Bildvorverarbeitung sinnvoll ist. Das Vorwissen über mittlere Werte von experimentellen Parametern in diesen Bereichen ist oft unvollständig und stark fehlerbehaftet. Abschnitt 5 beschreibt ein neues Verfahren um solche mittleren experimentellen Parameter zu präzisieren.

Mit den so gefundenen experimentellen Parametern können HREM-Bildsimulationen berechnet werden, die (kontrast- und helligkeitsangepasst) direkt mit der experimentellen Aufnahme vergleichbar sind. Insbesondere lässt sich so der Einfluss von amorphen Oberflächenschichten simulieren. Die Störung durch amorphe Oberflächenschichten verursacht korreliertes Rauschen in der Aufnahme. Die Elimination dieses Typs von Rauschen wurde in Kapitel 7 untersucht. Zur Lösung dieser Problemstellung wurde neben einer Reihe von klassischen Verfahren auch eine Methode unter Verwendung neuronaler Netze untersucht. Es stellte sich heraus, dass neuronale Netze im Vergleich zu klassischen Verfahren wesentlich geringere Fehler erreichen. Für die lokale Probendicken- und Kompositionsbestimmung trainiert man neuronale Netze mit simulierten HREM-Aufnahmen oberflächenamorphisierter Proben. Die trainierten neuronalen Netze bestimmen orts aufgelöst Komposition und Dicke des kristallinen Teils der Probe.

In Kapitel 8 wurden die entwickelten Methoden auf experimentelle HREM-Aufnahmen angewendet. Im Rahmen des Fehlers der Kompositionsbestimmung konnte dabei im ternären Halbleiter die statistische Schwankung der Probenzusammensetzung (Random-Alloy-Fluktuation) nachgewiesen werden. Die Verteilung der Komposition stimmt da-

bei sehr gut mit einem theoretischen Modell überein. Bei den untersuchten HREM-Proben traten Schwankungen der Probendicke auf, welche durch das unterschiedliche Abtragverhalten von AlAs und GaAs beim Ionenstrahlabdünnen (*ion-milling*) hervorgerufen werden.

Die Messung solcher Schwankungen in der Probendicke ist nur durch die Trennung von Dicken- und Kompositionsinformation im Verfahren möglich. Insbesondere kann diese Information durch intuitive Bildinterpretation nicht erhalten werden. Quantitative Bildauswertung extrahiert bis dahin unzugängliche Information aus den Bilddaten.

Ausblick

Bei der Probenpräparation kann das Ergebnis durch einen nach dem Ionenabdünnen durchgeführten nasschemischen Ätzprozess wesentlich verbessert werden [35]. Dies führt zu einer Verringerung der Amorphisierungstiefe und damit sowohl zu einer Verkleinerung der statistischen Fehler (Verringerung des Rauschens) wie auch der systematischen Fehler (kleinere Dickenschwankungen). Wird auf diese Art der Fehler der Verfahren verkleinert, so ist eine vergrößerte Ortsauflösung möglich (Kap. 4.3.3).

Die Mikrostrukturierung von III-V-Halbleitern kann durch Selbstorganisation z.B. beim Stranski-Krastanov Wachstum erfolgen [30]. Die hierfür notwendigen Gitterverspannungen erzeugen in der HREM-Aufnahme Spannungskontrast. Der Spannungskontrast stört die Kompositions- und Dickenbestimmung, weil sich der Bildkontrast durch die Verspannung auch ohne Änderung der Probendicke oder Komposition stark verändern kann. Abhilfe kann hier eine zusätzliche Auswertung von Verspannungsinformation (Positionen der Helligkeitsmaxima) leisten. Diese Zusatzinformation lässt sich bei den beschriebenen Methoden in den Trainingsprozess des neuronalen Netzes integrieren. Mit dieser Erweiterung wird es möglich sein, auch an verspannten Proben ortsaufgelöst Dicken- und Kompositionsbestimmung durchzuführen.

Anhang A

Atomare Streuamplituden nach Doyle und Turner

Für die analytischen Rechnungen wurden die von P.A. Doyle und P.S. Turner [13] eingeführten atomaren Streuamplituden verwendet:

$$f(\vec{s}) = \sum_{j=1}^4 \alpha_j \exp(-s^2 \beta_j) \quad (\text{A.1})$$

Hierin sind α_j und β_j elementspezifische Konstanten. Konkrete Werte für einige Elemente von Verbindungshalbleitern sind¹:

Element	α_1 [Å]	β_1 [Å ²]	α_2 [Å]	β_2 [Å ²]	α_3 [Å]	β_3 [Å ²]	α_4 [Å]	β_4 [Å ²]
Al	2.276	72.322	2.428	19.773	0.858	3.080	0.317	0.408
As	2.399	45.718	2.790	12.817	1.529	2.280	0.594	0.328
Ga	2.321	65.602	2.486	15.458	1.688	2.581	0.599	0.351
In	3.153	66.649	3.557	14.449	2.818	2.976	0.884	0.335
P	1.888	44.876	2.469	13.538	0.805	2.642	0.320	0.361

In der hier vorliegenden Arbeit wird der reziproke Raumvektor \vec{k} durch die Fouriertransformation

$$FT[f(\vec{r})] = \int_{\vec{r} \in \mathbb{R}^3} f(\vec{r}) \exp(2\pi i \vec{r} \vec{k}) d\vec{r} \quad (\text{A.2})$$

definiert.

¹Werte zu weiteren Elementen erhält man aus [13].

In der Originalarbeit von Doyle und Turner wird die Größe $s = \frac{k}{2}$ als Argument der atomaren Streuamplitude verwendet. Rechnet man die Streuamplitude von Ångström in Volt um, so erhält man:

$$f(\vec{k}) = \frac{h^2}{2\pi m_0 e} \sum_{j=1}^4 \alpha_j \exp\left(-\frac{k^2 \beta_j}{4}\right) \quad (\text{A.3})$$

Bei dem in (A.3) auftretenden Vorfaktor muss die Elektronenmasse **nicht** relativistisch korrigiert werden. Die relativistische Korrektur ist schon in der korrigierten Wellenlänge der Elektronen enthalten [43] [7].

Anhang B

Symmetrische Gitterbesetzung bei ternären Halbleitern

In dieser Arbeit werden ternäre Halbleitermaterialsysteme untersucht, bei denen die Komposition eines Untergitters variiert. Die Symmetrie des Streuvermögens der Struktur bestimmt die Auslöschungsregeln für die an der Abbildung beteiligten Reflexe. Auf dem gesamten Kompositionsbereich eines Materialsystems, gibt es immer eine Komposition, bei der sich die Streuvermögen der beiden Untergitter minimal unterscheiden und die 200-Reflexintensität im Beugungsbild des Kristalls minimal wird. Diese Komposition soll hier für verschiedene Materialsysteme berechnet werden.

Die Rechnung wird am Beispiel $\text{Al}_{1-x}\text{Ga}_x\text{As}$ vorgestellt. Die atomaren Streuamplituden approximieren wir durch einen numerischen Fit wie in der Arbeit von Doyle und Turner (siehe Anhang A). Für das konstant besetzte Untergitter (As) ergibt sich:

$$f_c(\vec{k}) = \frac{h^2}{2\pi m_0 e} \sum_{j=1}^4 \alpha_{cj} \exp\left(-\frac{k^2 \beta_{cj}}{4}\right) \quad (\text{B.1})$$

mit den Konstanten $\vec{\alpha}_c$ und $\vec{\beta}_c$ (in unserem Beispiel für Arsen)¹.

Die atomare Streuamplitude des binär besetzten Untergitters berechnen wir durch eine mit der Komposition x gewichtete Mittelung (VCA-Näherung):

$$f(\vec{k}) = \frac{h^2}{2\pi m_0 e} \sum_{j=1}^4 x \alpha_{1j} \exp\left(-\frac{k^2 \beta_{1j}}{4}\right) + (1-x) \alpha_{0j} \exp\left(-\frac{k^2 \beta_{0j}}{4}\right) \quad (\text{B.2})$$

¹für Werte siehe Anhang A und [13]

mit den Doyle-Turner Konstanten $\vec{\alpha}_1, \vec{\beta}_1$ für Gallium und $\vec{\alpha}_0, \vec{\beta}_0$ für Aluminium.

Wir berechnen einen Kennwert D für die Abweichung der Streuamplituden der beiden Untergitter. Dazu mitteln wir den quadratischen Abstand der beiden Streuamplituden über den gesamten k -Raum (Die Mittelung ist zweidimensional durchzuführen, genaueres siehe Abschnitt 3.2.2).

$$D = \frac{h^4}{4\pi^2 m_0^2 e^2} \int_0^\infty \int_0^{2\pi} \left(\sum_{j=1}^4 x \alpha_{1j} \exp\left(-\frac{k^2 \beta_{1j}}{4}\right) + (1-x) \alpha_{0j} \exp\left(-\frac{k^2 \beta_{0j}}{4}\right) - \alpha_{cj} \exp\left(-\frac{k^2 \beta_{cj}}{4}\right) \right)^2 K d\phi dK \quad (\text{B.3})$$

Nach Integration ergibt sich:

$$D = \frac{h^4}{\pi m_0^2 e^2} (x^2 S_1 + (1-x^2) S_2 + 2x(1-x) S_3 - 2x S_4 - 2(1-x) S_5 + S_6) \quad (\text{B.4})$$

Mit den Summen:

$$S_1 := \sum_{i=1}^4 \sum_{j=1}^4 \frac{\alpha_{1i} \alpha_{1j}}{\beta_{1i} + \beta_{1j}} \quad (\text{B.5})$$

$$S_2 := \sum_{i=1}^4 \sum_{j=1}^4 \frac{\alpha_{0i} \alpha_{0j}}{\beta_{0i} + \beta_{0j}} \quad (\text{B.6})$$

$$S_3 := \sum_{i=1}^4 \sum_{j=1}^4 \frac{\alpha_{1i} \alpha_{0j}}{\beta_{1i} + \beta_{0j}} \quad (\text{B.7})$$

$$S_4 := \sum_{i=1}^4 \sum_{j=1}^4 \frac{\alpha_{1i} \alpha_{cj}}{\beta_{1i} + \beta_{cj}} \quad (\text{B.8})$$

$$S_5 := \sum_{i=1}^4 \sum_{j=1}^4 \frac{\alpha_{0i} \alpha_{cj}}{\beta_{0i} + \beta_{cj}} \quad (\text{B.9})$$

$$S_6 := \sum_{i=1}^4 \sum_{j=1}^4 \frac{\alpha_{ci} \alpha_{cj}}{\beta_{ci} + \beta_{cj}} \quad (\text{B.10})$$

Abbildung B.1 stellt die Funktion D in Abhängigkeit von der Komposition x dar.

Die Komposition x mit minimaler Abweichung D berechnen wir als Nullstelle der ersten Ableitung:

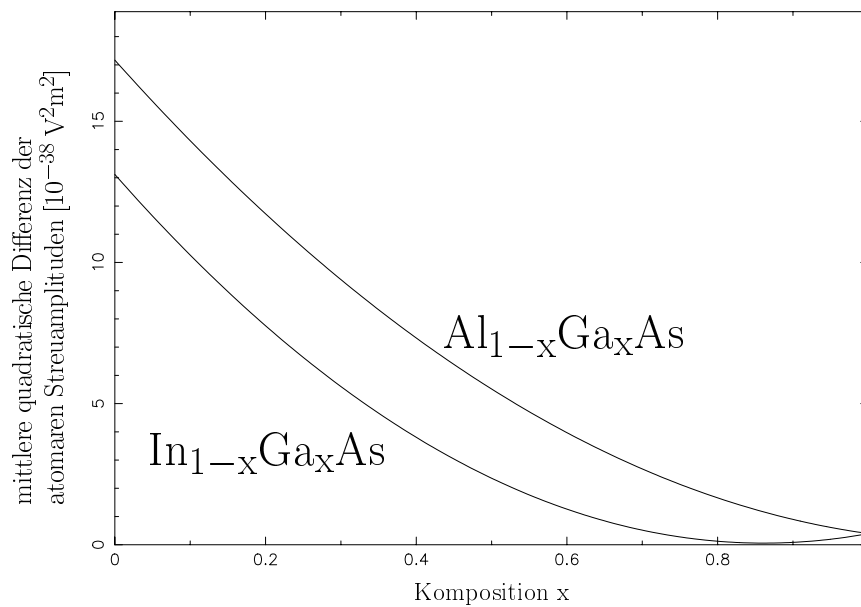


Abbildung B.1: Mittlere quadratische Differenz der atomaren Streuamplituden beider Untergitter in Abhängigkeit von der Komposition x . Dargestellt sind Graphen zweier Materialsysteme AlGaAs und InGaAs.

$$\frac{\partial D}{\partial x} \stackrel{!}{=} 0 \quad (\text{B.11})$$

$$\Rightarrow x_s = \frac{S_2 - S_3 + S_4 - S_5}{S_1 + S_2 - 2S_3} \quad (\text{B.12})$$

Für $\text{Al}_{1-x}\text{Ga}_x\text{As}$ ergibt sich daraus ein Wert

$$x_s = 1.14 \quad (\text{B.13})$$

Im Materialsystem $\text{Al}_{1-x}\text{Ga}_x\text{As}$ ist also die Komposition mit maximaler Symmetrie $x=1$. Dies entspricht dem reinen GaAs.

Für das Materialsystem $\text{In}_{1-x}\text{Ga}_x\text{As}$ errechnet sich

$$x_s = 0.862 \quad (\text{B.14})$$

Die maximale Symmetrie wird also bei einem Indiumgehalt von etwa 14 Prozent erreicht.

Betrachten wir ein Verfahren zur Kompositionsbestimmung welches nur die Intensität der 200-Reflexe auswertet. Bei $\text{In}_{0.14}\text{Ga}_{0.86}\text{As}$ haben beide Untergitter der Sphaleritstruktur das gleiche Streuvermögen. Daher präsentiert sich die Struktur bezüglich

ihrer Streueigenschaft für Elektronen mit der Symmetrie der Diamantstruktur. Die Auslöschungsregeln der Diamantstruktur verbieten die 200-Reflexe in (001)-Projektion. Die 200-Reflexintensität hat bei dieser Komposition ein Minimum. Die angestrebte Kompositionsbestimmung stellt eine Umkehrung des Zusammenhangs 200-Intensität und Komposition dar. Durch das Minimum bei $x=0.86$ ist der Zusammenhang aber nicht auf dem gesamten Kompositionsintervall, sondern nur auf $x \in [0, 0.86]$ umkehrbar. Kompositionswerten die größer als die Intervallobergrenze sind, kann bestenfalls der Wert 0.86 zugewiesen werden. Einer GaAs Probe ($x=1$) wird also ein Indiumgehalt von 14 Prozent zugeordnet. Bei Verfahren zur Kompositionsbestimmung, die ausschliesslich die 200-Reflexintensität auswerten, treten für das Materialsystem InGaAs auch ohne Störeinflüsse (Rauschen, Variation der Probendicke) bereits Fehler bis zu 14 Prozent auf.

Anhang C

Fouriertransformierte des atomaren Potentials

Als Nebenrechnung für Abschnitt 3.2.2 berechnen wir das nach x und y fouriertransformierte atomare Potential an einer Stelle z :

$$f(\vec{K}, z) = \int_{R_x \in \mathbb{R}} \int_{R_y \in \mathbb{R}} V_a \begin{pmatrix} R_x \\ R_y \\ z \end{pmatrix} \exp(2\pi i \vec{R} \vec{K}) dR_y dR_x \quad (\text{C.1})$$

Die atomare Streuamplitude ist die dreidimensionale Fouriertransformierte des atomaren Potentials:

$$f(\vec{k}) = \int_{\vec{r} \in \mathbb{R}^3} V_a(\vec{r}) \exp(2\pi i \vec{r} \vec{k}) d\vec{r} \quad (\text{C.2})$$

Wir verwenden hierfür die atomare Streuamplitude nach Doyle und Turner (siehe Anhang A):

$$f(\vec{k}) = \frac{h^2}{2\pi m_0 e} \sum_{j=1}^4 \alpha_j \exp\left(-\frac{k^2 \beta_j}{4}\right) \quad (\text{C.3})$$

Um hieraus (C.1) zu berechnen, führen wir zunächst mit (C.2) eine dreidimensionale inverse Fouriertransformation durch:

$$V_a(x, y, z) = \int_{\vec{k} \in \mathbb{R}^3} f(k) \exp(-2\pi i \vec{r} \vec{k}) d\vec{k} \quad (\text{C.4})$$

Da $f(k)$ in der Doyle-Turner-Näherung winkelunabhängig ist, kann man die Winkelintegration ausführen:

$$V_a(x, y, z) = \frac{2}{r} \int_{k=0}^{\infty} k f(k) \sin(2\pi r k) dk \quad (\text{C.5})$$

Setzt man nun die atomare Streuamplitude (C.3) ein, ergibt sich:

$$\begin{aligned} V_a(x, y, z) &= \frac{h^2}{2\pi m_0 e r} \sum_{j=1}^4 \alpha_j \int_{k=0}^{\infty} k \exp\left(-\frac{k^2 \beta_j}{4}\right) \sin(2\pi r k) dk \\ &= \frac{4h^2 \sqrt{\pi}}{m_0 e} \sum_{j=1}^4 \frac{\alpha_j}{\beta_j^{\frac{3}{2}}} \exp\left(-\frac{4\pi^2 r^2}{\beta_j}\right) \end{aligned} \quad (\text{C.6})$$

Setzt man (C.6) in (C.1) ein, so folgt

$$f(\vec{K}, z) = \frac{4h^2 \sqrt{\pi}}{m_0 e} \sum_{j=1}^4 \frac{\alpha_j}{\beta_j^{\frac{3}{2}}} \underbrace{\int_{R_x \in \mathbb{R}} \int_{R_y \in \mathbb{R}} \exp\left(-\frac{4\pi^2 (R^2 + z^2)}{\beta_j}\right) \exp(2\pi i \vec{R} \vec{K}) dR_y dR_x}_{:=H_1} \quad (\text{C.7})$$

Den Wert des Doppelintegrals in (C.7) berechnen wir als

$$H_1 = \exp\left(-\frac{4\pi^2 z^2}{\beta_j}\right) \underbrace{\int_{R_x \in \mathbb{R}} \int_{R_y \in \mathbb{R}} \exp\left(-\frac{4\pi^2 R^2}{\beta_j}\right) \exp(2\pi i \vec{R} \vec{K}) dR_y dR_x}_{H_2} \quad (\text{C.8})$$

In zweidimensionalen Polarkoordinaten ergibt sich:

$$\begin{aligned} H_2 &= \int_{R=0}^{\infty} R \exp\left(-\frac{4\pi^2 R^2}{\beta_j}\right) \int_{\phi=0}^{2\pi} \exp(2\pi i R K \cos(\phi)) d\phi dR \\ &= 2\pi \int_{R=0}^{\infty} R \exp\left(-\frac{4\pi^2 R^2}{\beta_j}\right) I_0(2\pi i K R) dR \end{aligned} \quad (\text{C.9})$$

$$= \frac{1}{4\pi} \beta_j \exp\left(-\frac{K^2 \beta_j}{4}\right) \quad (\text{C.10})$$

(in (C.9) tritt $I_0(z)$, die modifizierte Besselfunktion erster Gattung nullter Ordnung auf).

Wir erhalten das Ergebnis:

$$\begin{aligned}
 f(\vec{K}, z) &= \frac{4h^2\sqrt{\pi}}{m_0e} \sum_{j=1}^4 \frac{\alpha_j}{\beta_j^{\frac{3}{2}}} \frac{1}{4\pi} \beta_j \exp\left(-\frac{K^2\beta_j}{4}\right) \exp\left(-\frac{4\pi^2z^2}{\beta_j}\right) \\
 &= \frac{h^2}{m_0e\sqrt{\pi}} \sum_{j=1}^4 \frac{\alpha_j}{\sqrt{\beta_j}} \exp\left(-\frac{K^2\beta_j}{4}\right) \exp\left(-\frac{4\pi^2z^2}{\beta_j}\right)
 \end{aligned} \tag{C.11}$$

Literaturverzeichnis

- [1] H. Bethe. Theorie der Beugung von Elektronen an Kristallen. *Ann. Physik*, 87:55–128, 1928.
- [2] D. Bimberg, F. Heinrichsdorff, R.K. Bauer, D. Gerthsen, D. Stenkamp, D.E. Mars, and J.N. Miller. *J. Vac. Sci. Technol. B*, 10:1793ff, 1992.
- [3] C.M. Bishop. *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford, 1995.
- [4] M. Bode and A. Ourmazd. *J. Vac. Sci. Technol. B*, 10:1787ff, 1992.
- [5] E.O. Brigham. *FFT Schnelle Fourier-Transformation*. Oldenbourg Verlag, München, 4. edition, 1989.
- [6] D. Bruno and J.S. Hurley. *Superlattices Microstructures*, 11:23ff, 1992.
- [7] J.M. Cowley. *Diffraction Physics*. North-Holland Publishing Company, second edition, 1981.
- [8] J.M. Cowley and A.F. Moodie. The scattering of electrons by atoms and crystals. *Acta Cryst.*, 1:609–619, 1957.
- [9] A.G. Cullis, N.G. Chew, and C.R. Whitehouse. Material-dependent amorphization and epitaxial crystallization in ion-implanted AlAs/GaAs layer structures. *Appl. Phys. Lett.*, 55:1211–1213, 1989.
- [10] Y. Le Cun, J.S. Denker, and Sara A. Solla. Optimal brain damage. *Advances in Neural Information Proccession Systems I*, 1989.
- [11] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2:303–314, 1989.
- [12] E. Domany, J.L van Hemmen, and K. Schulten, editors. *Models of Neural Networks II*. Springer-Verlag, New York, 1994.

- [13] P.A. Doyle and P.S. Turner. Relativistic hartree-fock x-ray and electron scattering factors. *Acta Cryst.*, A24:390ff, 1968.
- [14] A. Förster, J. Lange, D. Gerthsen, Ch. Dieker, and H. Lüth. *J. Vac. Sci. Technol. B*, 11:23ff, 1993.
- [15] J. Frank. The envelope of electron microscopic transferfunctions for partially coherent illumination. *Optik*, 38:519–536, 1973.
- [16] S. Geman, E. Bienenstock, and R. Doursat. Neural networks and the bias/variance dilemma. *Neural Computation*, 1992.
- [17] F. Girosi and T. Poggio. Networks and the best approximation property. *Biological Cybernetics*, 63:169–176, 1990.
- [18] P. Goodman and A.F. Moodie. Numerical evaluation of n-beam wave functions in electron scattering by the multislice method. *Acta Cryst.*, A30:280–290, 1974.
- [19] H. Hashimoto, A. Howie, and M.J. Whelan. *Proc. R. Soc. London Ser. A*, 269:80–103, 1962.
- [20] J. Hertz, A. Krogh, and R. Palmer. *Introduction to the theory of Neural Computation*. Addison Wesley, Redwood City, 1991.
- [21] R. Hillebrand. Beiträge zur Computersimulation Elektronenmikroskopischer Hochauflösungsabbildungen. Dissertation, Martin-Luther-Universität Halle-Wittenberg, 1987.
- [22] R. Hillebrand, K. Scheerschmidt, W. Neumann, P. Werner, and A. Pippel. *Bildinterpretation in der Hochauflösungs-Elektronenmikroskopie*. Akademie-Verlag, Berlin, 1984.
- [23] R. Hillebrand, P. Werner, H. Hofmeister, and U. Gösele. A fuzzy logic approach to image analysis of HREM micrographs of III-V compounds. *Ultramicroscopy*, 66:73–88, 1996.
- [24] Electron Microscope JEM-4000EX. Manual, JEOL LTD. Tokyo.
- [25] I.T. Jolliffe. *Principal Component Analysis*. Springer-Verlag, New York, 1986.
- [26] H. Kirschner. Architekturabhängiges Lern- und Anpassungsverhalten bei Neuronalen Mehrschichtnetzen. Diplomarbeit, Institut für angewandte Physik der Universität Regensburg, 1997.

- [27] C. Kisielowski, P. Schwander, F.H. Baumann, M. Seibt, Y. Kim, and A. Ourmazd. An approach to quantitative high resolution transmission electron microscopy of crystalline materials. *Ultramicroscopy*, 58:131–155, 1995.
- [28] C. Kittel. *Einführung in die Festkörperphysik*. Oldenbourg Verlag, München, 10. edition, 1993.
- [29] W. Kleber, H. Bautsch, and J. Bohm. *Einführung in die Kristallographie*. Verlag Technik, Berlin, 1990.
- [30] A.O. Kosogov, P. Werner, and U. Gösele. Structural and optical properties of InAs–GaAs quantum dots subjected to high temperature annealing. *Appl. Phys. Lett.*, 69:3072–3074, 1996.
- [31] T. Kwok and D. Yeung. Constructive feedforward neural networks for regression problems. September 1995.
- [32] J. Larsen, L.K. Hansen, C. Svarer, and M. Ohlsson. Design and regularization of neural networks: The optimal use of a validation set. *Proceedings of the IEEE Workshop on Neural Networks for signal Processing VI*, 1996.
- [33] K.L. Merkle, R. Csencsits, K.L. Rynes, J.P. Withrow, and P.A. Stadelmann. The effect of three-fold astigmatism on measurements of grain boundary volume expansion by high-resolution electron microscopy. *Journal of Microscopy*, 190:204–213, 1998.
- [34] G. Möbus, R. Schweinfest, T. Gemming, T. Wagner, and M. Rühle. Iterative structure retrieval techniques in HREM: a comparative study and a modular program package. *Journal of Microscopy*, 190:109–130, 1998.
- [35] A. Ourmazd, F.H. Baumann, M. Bode, and Y. Kim. Quantitative chemical lattice imaging: theory and practice. *Ultramicroscopy*, 34:237–255, 1990.
- [36] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B. Flannery. *Numerical Recipes in C*. Cambridge University Press, Cambridge, 1992.
- [37] L. Reimer. *Transmission Electron Microscopy*. Springer, second edition, 1989.
- [38] M. Riedmiller and H. Braun. The RPROP algorithm. *Proceedings of the IEEE International Conference on Neural Networks*, 1993.
- [39] A. Rosenauer, S. Kaiser, T. Reisinger, J. Zweck, and W. Gebhart. Digital analysis of high resolution transmission electron microscopy lattice images. *OPTIK*, 102:63–69, 1996.

- [40] A. Rosenauer, T. Remmele, D. Gerthsen, K. Tillmann, and A. Förster. Atomic scale strain measurements by the digital analysis of transmission electron microscopic lattice images. *OPTIK*, 105:99–107, 1997.
- [41] K. Scheerschmidt, S. Ruvimov, P. Werner, A. Höpner, and J. Heydenreich. HREM structure characterization of interfaces in semiconducting multi-layers using molecular-dynamics supported image interpretation. *Journal of Microscopy*, 179:214–228, 1995.
- [42] P.A. Stadelmann. EMS - a software package for electron diffraction analysis and HREM image simulation in materials science. *Ultramicroscopy*, 21:131–146, 1987.
- [43] P.A. Stadelmann. Image calculation techniques. Technical report, EPFL Lausanne, 1995.
- [44] D. Stenkamp. Detection and quantitative assessment of image aberrations from single HRTEM lattice images. *Journal of Microscopy*, 190:194–203, 1998.
- [45] D. Stenkamp and W. Jäger. Compositional and structural characterization of Si_xGe_{1-x} alloys and heterostructures by high-resolution transmission electron microscopy. *Ultramicroscopy*, 50:321–354, 1993.
- [46] D. Stenkamp and H.P. Strunk. Quantitative determination of defocus, thickness and composition from high-resolution transmission electron microscopy lattice images. *Appl Phys. A*, 62:369–372, 1996.
- [47] V.M. Ustinov. *private communication*. Ioffe PTI, June 1999.
- [48] T. Walther. Untersuchung struktureller Eigenschaften von AlAs/GaAs-Quantentopfstrukturen mittels hochauflösender Transmissions-Elektronenmikroskopie. Diplomarbeit, Lehrstuhl IV e für Experimentalphysik der Technischen Hochschule Aachen, 1993.
- [49] T. Walther and D. Gerthsen. Quantitative characterization of AlAs/GaAs interfaces by HRTEM along the (100) and the (110) projection. *Appl. Phys. A*, 57:393–400, 1993.
- [50] T. Walther, C.J. Humphreys, M.P. Grimshaw, and A.C. Churchill. Detection of random alloy fluctuations in high-resolution transmission electron micrographs of AlGaAs. *Phil. Mag. A*, 72:1015–1030, 1995.
- [51] A. Warwick, W.Y. Jan, A. Ourmazd, and T.D. Harris. Does luminescence show semiconductor interfaces to be atomically smooth? *Appl. Phys. Lett.*, 56:2666ff, 1990.

- [52] A. Weickenmeier and H. Kohl. Computation of absorptive form factors for high-energy electron diffraction. *Acta Cryst.*, 47:590–597, 1991.
- [53] N. Wirth. *Algorithmen und Datenstrukturen*. Teubner Verlag, Stuttgart, 3. edition, 1983.

Ein herzliches Dankeschön ...

an Prof. Dr. U. Gösele für die Möglichkeit in seiner Abteilung dieses sehr interessante Thema bearbeiten zu können,

an Dr. R. Hillebrand, der die Arbeit betreuend begleitete, für zahlreiche Diskussionen und Anregungen zu TEM und Bildverarbeitung, und für die Unterstützung auf den Projekttreffen der Volkswagen Stiftung,

an Dr. P. Werner, ohne dessen HREM-Aufnahmen die Arbeit reine Theorie geblieben wäre,

an Dr. K. Scheerschmidt für einige wichtige Hinweise zur TEM Theorie und zur Simulation von HREM-Abbildungen, und für's Korrekturlesen,

an Dr. A. Feldhoff für Tips zu Layout und \LaTeX , und für's Korrekturlesen,

an meine Eltern, die stets verständnisvoll und geduldig waren auch wenn ich mal nicht so toll drauf war, und

an die Volkswagen Stiftung, die diese Arbeit unter Az. I/71108 finanzierte.

Lebenslauf

04.01.1969	geboren in Freising
1975–1980	St. Severin Volksschule in Garching bei München
1980–1983	Werner-Heisenberg-Gymnasium in Garching
1983–1989	Burkhart-Gymnasium in Mallersdorf-Pfaffenberg
1989	Abitur
1989–1990	Zivildienst im Kreiskrankenhaus Landshut
1990–1997	Physikstudium an der Universität Regensburg
1.1.1997–1.1.1998	Diplomarbeit: Architekturabhängiges Lern- und Anpassungsverhalten bei neuronalen Mehrschichtnetzen
1.1.1997–1.1.2000	Doktorand am Max-Planck-Institut für Mikrostrukturphysik in Halle (Saale)

Erklärung

Hiermit erkläre ich, dass ich diese Dissertation selbstständig und ohne fremde Hilfe verfasst habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt. Die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen habe ich als solche kenntlich gemacht.

Halle (Saale), den 19. Januar 2000

(Holger Kirschner)