

2 Grundlagen

In Kapitel 2 werden die Grundlagen zum Verständnis der orts aufgelösten Charakterisierung von $\text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y}$ -Halbleiter-Heteroschichten im RTEM dargelegt. Zunächst werden die für die Raster-Transmissionselektronenmikroskopie relevanten physikalischen und messtechnischen Grundlagen erläutert. Dabei wird auf die wichtigsten Streuprozesse und Beugungseffekte eingegangen, danach der apparative Aufbau des RTEM und die verwendeten Charakterisierungstechniken vorgestellt. Anschließend werden die für die Entstehung der Mischungslücke verantwortlichen thermodynamischen Voraussetzungen diskutiert und abschließend die relevanten Materialeigenschaften der zu untersuchenden Heteroschichten präsentiert.

2.1 *Elektronenmikroskopische Grundlagen*

2.1.1 Wechselwirkungen zwischen Elektronen und Festkörper

Wird eine Probe mit einer fein fokussierten, hochenergetischen Elektronensonde bestrahlt, entstehen aufgrund der Interaktion zwischen Primärelektronen und Probenatomen eine Vielzahl von Wechselwirkungsprodukten, die spezifische Informationen über die Probe bzw. deren Eigenschaften liefern. Im Raster-Transmissionselektronenmikroskop werden die transmittierten Elektronen als Informationsträger detektiert. Zu den transmittierten Elektronen gehören Elektronen des Primärelektronenstrahls, die keine Wechselwirkung erfahren haben oder an den Probenatomen elastisch und inelastisch gestreut werden sowie gebeugte Elektronen. Im Folgenden wird auf die Physik der oben genannten relevanten Wechselwirkungsprodukte eingegangen.

2.1.1.1 *Elastische Streuung an einzelnen Atomen*

In der Elektronenmikroskopie nennt man alle Streuprozesse elastisch, bei denen die Elektronen innerhalb einer Probe keine oder nahezu keine Energieverluste erleiden und nur eine Änderung ihrer Bewegungsrichtung erfahren. Der elastische Streuprozess beruht hauptsächlich auf der Coulomb-Kraft, die zwischen den negativ geladenen Primärelektronen

und den positiv geladenen Atomkernen der Probe herrscht. Bild 2.1 zeigt die schematische Darstellung der Ablenkung der Elektronen im Coulomb-Feld des Kerns nach Reimer [13].

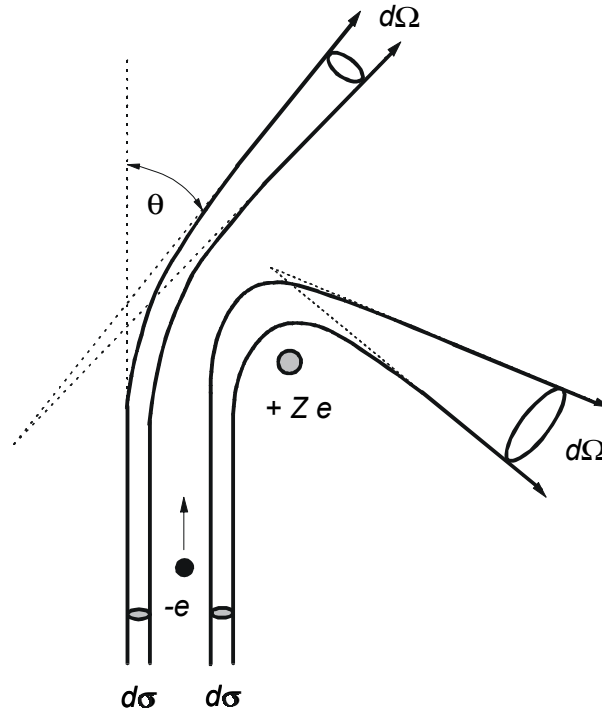


Bild 2.1 : Schematische Darstellung des elastischen Streuprozesses

Aufgrund des großen Massenunterschieds zwischen den einfallenden Elektronen und den im Kristallgitter gebundenen Atomkernen können die eindringenden Elektronen nahezu keinen Impuls auf das streuende Probenatom übertragen, wodurch die Änderung ihrer kinetischen Energie nur sehr gering ist. Die bei der Coulomb-Streuung auftretenden Gitterschwingungen (Phononenanregung) oder Intrabandübergänge können energetisch nicht aufgelöst werden, da der Energieverlust unterhalb des Auflösungsvermögens des RTE (0,3 eV) liegt.

Die Streuprozesse können mit Hilfe des sogenannten differentiellen Streuquerschnitts $d\sigma/d\Omega$ oder der Streuamplitude $f(\theta)$ quantitativ beschrieben werden. Der Quotient $d\sigma/d\Omega$ gibt die Wahrscheinlichkeit an, mit der ein Elektron aus einem Elektronenbündel vom Querschnitt $d\sigma$ unter einem bestimmten Streuwinkel θ in einem Raumwinkelement $d\Omega = 2\pi \sin(\theta)d\theta$ auftritt. Der Zusammenhang zwischen dem differentiellen Streuquerschnitt und der Streuamplitude ergibt sich zu

$$\frac{d\sigma}{d\Omega} = |f(\theta)|^2 \quad (2.1)$$

Ein einfaches Modell zur Beschreibung der elastischen Streuung geht davon aus, dass die im weiten Winkel gestreuten Elektronen den Atomkern so nahe passieren, dass sie von der

abschirmenden Wirkung der Elektronenhülle des Atoms nicht beeinflusst werden. Man spricht hierbei von der Rutherford-Streuung. Für den Rutherford-Streuquerschnitt $d\sigma_R/d\Omega$ gilt:

$$\frac{d\sigma_R}{d\Omega} = \frac{e^4 Z^2}{(4\pi\epsilon_0)^2 m^2 v^4 \sin^4(\theta/2)} \quad (2.2)$$

Hierbei ist e die Elementarladung, Z die Ordnungszahl, ϵ_0 die Permittivität des Vakuums, m die Elektronenmasse und v die Elektronengeschwindigkeit. Eine ausführliche Herleitung der Rutherford'schen Streuformel ist in [14] gegeben.

Für die Weitwinkelstreuung ist der Rutherford-Streuquerschnitt eine gute Näherung und es zeigt sich ein quadratischer Zusammenhang zwischen der Ordnungszahl und dem Streuquerschnitt. Bei kleinen Streuwinkeln hingegen ist der Abstand der Elektronen zum Atomkern größer. Somit nimmt der Einfluss der durch die Elektronenhülle entstehenden Abschirmung der Kernladung zu und der quadratische Zusammenhang geht verloren. Daher muss zur quantitativen Beschreibung der Elektronenstreuung im Elektronenmikroskop die abschirmende Wirkung der Elektronenhülle mit berücksichtigt werden. Dies geschieht durch die quantenmechanische Berechnung der Wellenfunktionen der Elektronen bzw. ihrer Streuamplitude. Je nach Wahl des entsprechenden Atommodells gibt es große Unterschiede für Streuwinkel $\theta < 50$ mrad. Bei größeren Streuwinkeln stimmen die Modelle gut überein [15].

Für die Berechnung der Elektronenstreuamplitude $f_e(s)$ eignet sich für Streuwinkel $\theta > 50$ mrad die Mott-Formel [16]. Sie berücksichtigt den Einfluss relativistischer Effekte, die Abschirmung durch die Elektronenhülle sowie den Elektronenspin.

$$f_e(s) = \frac{8\pi^2 m e^2}{h^2} \frac{Z - f_x(s)}{s^2}, \quad \text{mit } s = \frac{1}{2} |\vec{k}_0 - \vec{k}_s| \quad (2.3)$$

Dabei ist $f_x(s)$ die Röntgenstreuamplitude, h das Planck'sche Wirkungsquantum, \vec{k}_0 der Wellenvektor des einfallenden Elektrons und \vec{k}_s der Wellenvektor des gestreuten Elektrons. Die Werte der Röntgenstreuamplituden sind zum Beispiel in [17] dokumentiert.

2.1.1.2 Inelastische Streuung an einzelnen Atomen

Bei der inelastischen Streuung von Elektronen handelt es sich um Streuprozesse, bei denen die Elektronen innerhalb der Probe teilweise erhebliche Energieverluste erfahren, ihre Bewegungsrichtung jedoch unverändert bzw. nahezu unverändert bleibt ($\theta < 50 \text{ mrad}$). Die inelastische Streuung resultiert aus der Wechselwirkung zwischen den einfallenden schnellen Elektronen und den Elektronen der Atomhülle. Die Höhe des jeweiligen Energieverlustes hängt von der Art des Streuprozesses ab und kann zwischen einigen eV bis zu einigen keV betragen. Zwischen folgenden Streuprozessen kann unterschieden werden [18]:

1. Die Primärelektronen treten mit mehreren äußeren Elektronen (Leitungsband- bzw. Valenzbandelektronen) in Wechselwirkung und regen damit kollektive quantisierte Gitterschwingungen an (Plasmonenanregung). Der Plasmonverlust hängt von der Konzentration der Valenz- und Leitungsbandelektronen ab und wird durch ihre chemische Bindung und das Bandstruktur beeinflusst. Der Energieverlust liegt bei den meisten Festkörpern zwischen 5 bis 30 eV.
2. Anregung eines Elektrons äußerer Schalen. In einem Isolator oder Halbleiter entspricht dies einem Übergang eines Valenzelektrons über den Bandabstand. Bei Metallen wird ein Leitungsbandedelektron innerhalb eines Energiebands in einen höheren energetischen Zustand versetzt. Befindet sich der Endzustand des angeregten Elektrons oberhalb der Vakuumenergie des Festkörpers und ist die Anregungsenergie groß genug, kann das Elektron die Probe als Sekundärelektron verlassen. Der Energieverlust des Elektrons beträgt bis zu einigen 10 eV und der Streuwinkel liegt bei einer Primärenergie von 100 keV zwischen 1 bis 2 mrad.
3. Ionisation eines Elektrons einer inneren Schale. Der Grundzustand eines Elektrons einer inneren Schale liegt typischerweise einige 100 eV bis zu einigen 1000 eV unterhalb des Fermi-Niveaus des Festkörpers. Nur oberhalb des Fermi-Niveaus existieren unbesetzte Zustände. Daher kann ein Übergang eines Elektrons aus den inneren Schalen nur dann erfolgen, wenn es einen Energiebetrag absorbiert, der höher als seine ursprüngliche Bindungsenergie ist. Aufgrund der Energieerhaltung verliert das Elektron einen entsprechenden Energiebetrag und wird in einem Winkel von etwa 10 mrad gestreut ($E_0 = 100 \text{ keV}$). Das Probenatom hingegen befindet sich in einem angeregten ionisierten Zustand. Die „überschüssige“ Ionisierungsenergie wird in der Form freigesetzt, dass Elektronen der äußeren Schalen oder Elektronen innerer Schalen mit geringer Bindungsenergie auf einen unbesetzten Platz in einer inneren Schale übergehen. Die

dabei frei werdende Energie kann in Form von Röntgenstrahlen oder in Form von kinetischer Energie, die auf ein drittes Elektron übertragen wird, abgegeben werden (Auger-Elektronen).

Die bei der Ionisation innerer Schalen auftretenden Elektronenenergieverluste sind charakteristisch für die jeweiligen Probenatome, da ihre Bindungsenergie charakteristisch für die entsprechende Atomsorte ist. Somit können aus den Energieverlusten der inelastisch gestreuten Elektronen Informationen über die in der Probe vorkommenden Atomsorten und damit Informationen über die chemische Zusammensetzung der Probe gewonnen werden.

Zur genauen Beschreibung der inelastischen Streuung von Elektronen an einem Atom muss die Eigenschaft der Elektronen für den nicht angeregten Zustand durch die Wellenfunktion Ψ_0 und für den angeregten Zustand durch die Wellenfunktion Ψ_n formuliert werden. Die Übergangswahrscheinlichkeit, dass ein gebundenes Elektron durch den Impulsübertrag \vec{q} vom Grundzustand Ψ_0 in den angeregten Zustand Ψ_n übergeht, ist der sogenannte inelastische Formfaktor $|\varepsilon_n(\vec{q})|^2$. Er berücksichtigt die elementspezifischen Eigenschaften des ionisierten Atoms und ist unabhängig von der Energie des gestreuten Elektrons. Dies führt zu dem von Inokuti ermittelten Ausdruck für den differentiellen Streuquerschnitt inelastisch gestreuter Elektronen [19]:

$$\frac{d\sigma_n}{d\Omega} = \left(\frac{4\gamma^2}{a_0^2 q^4} \right) \frac{k_1}{k_0} |\varepsilon_n(\vec{q})|^2 \quad (2.4)$$

$$\varepsilon_n = \left\langle \Psi_n \left| \sum_j \exp(i\vec{q} \cdot \vec{r}_j) \right| \Psi_0 \right\rangle \quad (2.5)$$

$k_0 = |\vec{k}_0|$ und $k_1 = |\vec{k}_1|$ sind die Beträge der Wellenvektoren des angeregten Elektrons vor und nach dem Stoß, $\vec{q} = \vec{k}_0 - \vec{k}_1$ der dabei übertragene Impuls, \vec{r}_j der jeweilige Ortsvektor des Elektrons, $a_0 = 0,529 \cdot 10^{-10} \text{ m}$ der Bohr'sche Atomradius und $\gamma = (1 - v^2/c^2)^{-1/2}$ ein relativistischer Korrekturfaktor.

2.1.1.3 Elektronenbeugung an idealen Kristallen

2.1.1.3.1 Einführung in die Kristallographie

Nach de Broglie weisen die als Teilchen interpretierten Elektronen auch Welleneigenschaften auf. Die Wellenlänge λ der Elektronen hängt dabei mit deren Impuls p bzw. deren kinetischen Energie E zusammen (m_0c^2 =Ruheenergie):

$$\lambda = \frac{h}{p} = \frac{h}{mv} = \frac{hc}{\sqrt{2m_0c^2E + E^2}} \quad (2.6)$$

Gleichung 2.6 berücksichtigt dabei den relativistischen Massezuwachs des Elektrons [20].

Trifft eine Elektronenwelle auf ein Kristallgitter, so kommt es aufgrund der periodischen Anordnung der Atome zu Beugungserscheinungen, sofern die Elektronenwellenlänge im Bereich der Periodizität der Atome bzw. der Netzebenen liegt. Jedes Kristallgitter besteht aus der dreidimensionalen Aneinanderreihung sogenannter Einheitszellen, welche durch drei nicht koplanare Basisvektoren \vec{a}_1 , \vec{a}_2 und \vec{a}_3 definiert werden. Der Ursprung der Einheitszelle wird im realen Kristall durch den Ortsvektor

$$\vec{r}_g = m\vec{a}_1 + n\vec{a}_2 + o\vec{a}_3 \quad (m, n, o \in \mathbb{Z}) \quad (2.7)$$

beschrieben. Die Einheitszelle enthält im Allgemeinen mehrere Atome k ($k = 1, \dots, n$) an den Positionen

$$\vec{r}_k = u_k\vec{a}_1 + v_k\vec{a}_2 + w_k\vec{a}_3 \quad (0 \leq u_k, v_k, w_k \leq 1) \quad (2.8)$$

Der Ort eines einzelnen Atoms im Kristall ist durch die Vektorsumme $\vec{r}_g + \vec{r}_k$ festgelegt.

Das Auftreten von Beugung an einem periodischen Objekt kann auf unterschiedliche Weise formuliert werden. Bezieht man sich auf das Kristallgitter im realen Raum, gibt das Bragg-Gesetz

$$n\lambda = 2d_{hkl} \sin \vartheta_B \quad \text{mit } n \in \mathbb{Z} \quad (2.9)$$

an, unter welchen Bedingungen eine einfallende Elektronenwelle gebeugt wird, d.h. konstruktive Interferenz auftritt. Hier ist d_{hkl} der Gitterebenenabstand (Netzebenenabstand)

und ϑ_B der Bragg'sche Reflexionswinkel (vgl. Bild 2.2). Der Netzebenenabstand d_{hkl} kann für kubische Werkstoffe durch

$$d_{hkl} = \frac{a}{\sqrt{h^2 + k^2 + l^2}} \quad \text{mit } h, k, l \in \mathbb{Z} \quad (2.10)$$

berechnet werden, wobei h , k und l die Miller'schen Indizes sind und a die Gitterkonstante ist.

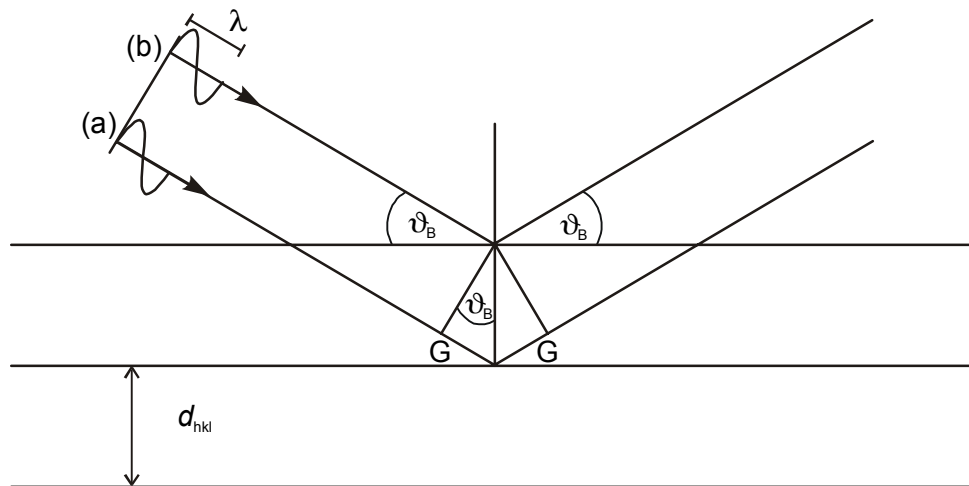


Bild 2.2 : Veranschaulichung der Bragg-Beziehung

Die Bragg-Bedingung besagt, dass nur dann ein Interferenzmaximum (Bragg-Reflex) auftreten kann, wenn die an einer tiefer liegenden Netzebene gebeugte Elektronenwelle (a) einen Gangunterschied $G = 2d_{hkl} \sin \vartheta$, welcher ein ganzes Vielfaches der de Broglie Wellenlänge λ ist, zur Welle (b) besitzt. Dieser Zusammenhang wird in Bild 2.2 graphisch verdeutlicht.

Die Beugungsaufnahme eines Kristallgitters bzw. die Abbildung der Bragg-Reflexe steht in direkter Beziehung zu der geometrischen Anordnung der Atome im Kristallgitter und ist - mathematisch betrachtet - die dreidimensionale Fouriertransformation des Kristallgitters im realen Raum (Ortsraum). Man führt daher das sogenannte reziproke Gitter (im reziproken Raum) ein, welches von den Bragg-Reflexen gebildet wird. Ortsraum und reziproker Raum liefern gleiche physikalische Aussagen, jedoch steht das reziproke Gitter in direktem Zusammenhang mit den Beugungsaufnahmen, weshalb häufig eine Darstellung im reziproken Raum gewählt wird.

Ein unendlich ausgedehntes Kristallgitter im Ortsraum liefert im reziproken Raum ein Punktgitter. Alle Punkte des reziproken Gitters können durch die reziproken Basisvektoren \vec{a}_1^* , \vec{a}_2^* und \vec{a}_3^* beschrieben werden:

$$\vec{a}_1^* = \frac{\vec{a}_2 \times \vec{a}_3}{V_{EZ}}, \quad \vec{a}_2^* = \frac{\vec{a}_3 \times \vec{a}_1}{V_{EZ}}, \quad \vec{a}_3^* = \frac{\vec{a}_1 \times \vec{a}_2}{V_{EZ}} \quad (2.11)$$

mit $V_{EZ} = \vec{a}_1(\vec{a}_2 \times \vec{a}_3)$ Volumen der Einheitszelle

$$\text{und } \vec{a}_i^* \cdot \vec{a}_j = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases} \quad i, j = 1, 2, 3$$

In Analogie zum Ortsvektor \vec{r}_r wird der reziproke Gittervektor \vec{g}

$$\vec{g} = h\vec{a}_1^* + k\vec{a}_2^* + l\vec{a}_3^* \quad (2.12)$$

definiert. Dieser steht senkrecht zu den Netzebenen mit den Miller'schen Indizes hkl . Sein Betrag ist gleich dem reziproken Netzebenenabstand.

$$|\vec{g}| = \frac{1}{d_{hkl}} \quad (2.13)$$

Das Auftreten von Bragg-Reflexen im reziproken Raum wird durch die Laue-Bedingung beschrieben. Bei einem einfallenden Strahl der Richtung \vec{k}_0 treten in allen Richtungen \vec{k} Reflexe auf, wenn gilt:

$$\vec{k} - \vec{k}_0 = \vec{g} \quad (2.14)$$

Gleichung 2.14 ist das Bragg-Gesetz in vektorieller Schreibweise und ist die Grundlage zur graphischen Bestimmung der Lage der Intensitätsmaxima mittels der Ewald-Konstruktion. Bild 2.3 veranschaulicht diesen Sachverhalt.

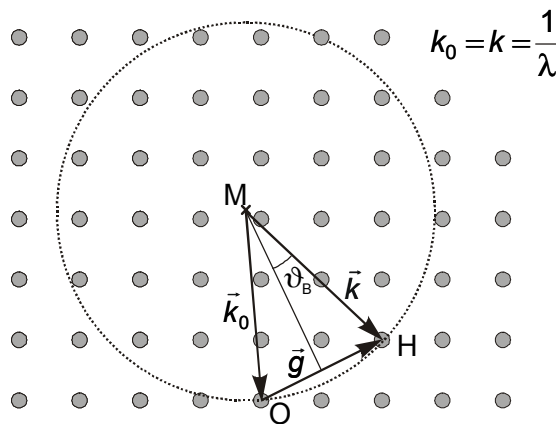


Bild 2.3 : Ewald-Konstruktion nach Reimer [20]

Der einfallende Wellenvektor endet im Ursprung O des reziproken Gitters. Man zeichnet nun eine sogenannte Ewald-Kugel mit dem Radius $|\vec{k}_0| = 1/\lambda$ um den Fußpunkt M von \vec{k}_0 . Schneidet die Kugel einen beliebigen Punkt des reziproken Gitters, so entsteht ein gebeugter Strahl \vec{k} . Der reziproke Gittervektor \vec{g} verbindet den einfallenden Wellenvektor mit dem gebeugten Wellenvektor.

2.1.1.3.2 Kinematische Beugungstheorie

Die kinematische Beugungstheorie ist der einfachste Ansatz zur Beschreibung der Beugung von Elektronen und ermöglicht die Ermittlung der Lage der Intensitätsmaxima und -minima einer Beugungsaufnahme. Eine in einen idealen Kristall einfallende Elektronenwelle Ψ_0 mit dem Wellenvektor \vec{k}_0 wird einmal an einem Atom um den Winkel θ mit der Streuamplitude $f(\theta)$ gebeugt. Da vorausgesetzt wird, dass die Elektronen nur elastisch gestreut werden, d.h. keine Verluste durch Stoßprozesse erfahren haben, hat der Wellenvektor \vec{k} der gebeugten Welle denselben Betrag wie \vec{k}_0 . Die resultierende Wellenfunktion Ψ ergibt sich aus der Summe der in einer Richtung gestreuten Elektronenwellen:

$$\Psi = \Psi_0 \sum_{m=1}^{M_1} \sum_{n=1}^{M_2} \sum_{o=1}^{M_3} \sum_{k=1}^N f_k(\theta) \exp[-2\pi i(\vec{k} - \vec{k}_0) \cdot (\vec{r}_g + \vec{r}_k)] \quad (2.15)$$

Durch Umformung kann Gleichung 2.15 in zwei Faktoren zerlegt werden:

$$\Psi = \underbrace{\Psi_0 \sum_{k=1}^N f_k(\theta) \exp[-2\pi i(\vec{k} - \vec{k}_0) \cdot \vec{r}_k]}_F \cdot \underbrace{\sum_{m=1}^{M_1} \sum_{n=1}^{M_2} \sum_{o=1}^{M_3} \exp[-2\pi i(\vec{k} - \vec{k}_0) \cdot \vec{r}_g]}_G \quad (2.16)$$

Der erste Faktor F wird als Strukturamplitude bezeichnet und ist nur vom Atomort und -typ innerhalb der Einheitszelle abhängig. Der zweite Faktor G wird Gitteramplitude genannt und berücksichtigt die Art des Kristallgitters. Setzt man Gleichung 2.12 und 2.8 in den Term der Strukturamplitude F aus Gleichung 2.16 ein, wird deutlich, dass F für bestimmte Werte von hkl Null wird. Aufgrund der „Auslöschung“ bestimmter Netzebenen und der Größe von F kann zwischen unterschiedlichen Kristallgittertypen unterschieden werden. Im für diese Arbeit relevantem Zinkblendegitter treten bei der kinematischen Elektronenbeugung nur dann Bragg-Reflexe auf ($F \neq 0$), wenn hkl alle ungerade oder hkl alle gerade sind und $h+k+l$ durch vier teilbar ist. Weitere Werte der Strukturamplituden der gebräuchlichsten Kristalltypen sind z.B. in [14] aufgelistet.

In der Transmissionselektronenmikroskopie werden sehr dünne Kristalle verwendet. Das heißt, der Kristall kann als eine dünne Scheibe angesehen werden, deren Kantenlängen sehr viel größer sind als ihre Dicke. Im reziproken Gitter hat dies die Konsequenz, dass sich die Punkte in dünne Striche bzw. Nadeln aufweiten, wobei die Längsrichtung dieser Nadeln im Ortsraum der Durchstrahlrichtung entspricht. Beim idealen Kristall treten wie bereits erwähnt nur dort Bragg-Reflexe auf, wo die Ewaldkugel Gitterpunkte exakt schneidet. Durch die Aufweitung der Gitterpunkte zu Strichen in einem Kristall mit endlicher Dicke entstehen nun auch dort Intensitätsmaxima, wo die Ewaldkugel einen Strich schneidet. Die Abweichung von der exakten Bragg-Bedingung bezeichnet man als Anregungsfehler \vec{s} . Gleichung 2.12 erweitert sich damit zu

$$\vec{k} - \vec{k}_0 = \vec{g} + \vec{s} \quad (2.17)$$

Das angeregte Intensitätsmaximum wird um so schwächer, je größer der Anregungsfehler ist und ist nach Hirsch [21] durch

$$I_g = I_e \cdot \frac{\pi^2}{\xi_g^2} \cdot \frac{\sin^2(\pi t s_z)}{(\pi s_z)^2} \quad (2.18)$$

definiert. I_e ist die Intensität der einfallenden Elektronenwelle, t die Probendicke, s_z die Komponente des Anregungsfehlers in Durchstrahlrichtung und ξ_g die Extinktionslänge. Die Extinktionslänge ist von der Strukturamplitude abhängig und ein Maß dafür, wie schnell die Amplitude eines Bragg-Reflexes anwächst. Ist der Anregungsfehler aber sehr klein oder gleich Null, wird $I_g > 1$ wenn $t > \xi_g / \pi$, das heißt, die Intensität des gebeugten Elektronenstrahls wird größer als die des einfallenden Elektronenstrahls, was physikalisch nicht sinnvoll ist. Gleichung 2.18 zeigt, dass der kinematische Beugungsansatz nur für sehr dünne Kristalle Gültigkeit hat.

2.1.1.3.3 Dynamische Beugungstheorie

Für eine quantitative Berechnung der Intensitäten der Bragg-Reflexe wird die sogenannte dynamische Beugungstheorie angewendet. Im Gegensatz zur kinematischen Theorie berücksichtigt sie die Tatsache, dass die gestreuten Wellen im Kristall wiederum gestreut werden, d.h. die Intensitäten der gestreuten Wellen interagieren dynamisch sowohl untereinander also auch mit der einfallenden Welle. Zusätzlich werden relativistische Effekte in die Rechnung mit einbezogen, also auch inelastische Streuprozesse und thermische Effekte berücksichtigt. Es existieren verschiedene Lösungsansätze, die beispielsweise in

[21], [14], [22], [23] oder [20] ausführlich dokumentiert sind. Der klassische Ansatz ist die Blochwellentheorie, die auf Bethe (1928) zurückzuführen ist. Dieser Ansatz und das sogenannte Multislice-Verfahren, welches von Cowley und Moodie (1957) entwickelt wurde, sollen hier kurz vorgestellt werden.

Blochwellen-Theorie

Ausgangspunkt des Ansatzes nach Bethe zur Beschreibung der Wellenfunktionen schneller Elektronen ist die Schrödingergleichung [24]:

$$\Delta\Psi(\vec{r}) + \frac{8\pi^2 m_0 e}{h^2} [U + \varphi(\vec{r})] \Psi(\vec{r}) = 0 \quad (2.19)$$

Das Kristallpotential $\varphi(\vec{r})$ ergibt sich aus der Überlagerung der Potentiale aller Atome des Kristalls und kann im reziproken Raum durch eine Fourierreihe über alle Gitterpunkte \vec{g} dargestellt werden:

$$\varphi(\vec{r}) = \sum_{\vec{g}} \varphi_{\vec{g}} \exp(2\pi i \vec{g} \cdot \vec{r}) \quad (2.20)$$

Die Fourierkoeffizienten $\varphi_{\vec{g}}$ sind dabei direkt proportional zur Strukturamplitude aus Gleichung 2.16. Eine Lösung der Schrödingergleichung, welche die gleiche Periodizität wie das Kristallgitter aufweisen muss, ist die Überlagerung sogenannter Blochwellen $B_j(\vec{r})$

$$\begin{aligned} \Psi(\vec{r}) &= \sum_j \alpha_j \cdot B_j(\vec{r}) \\ &= \sum_j \alpha_j \cdot \sum_{\vec{g}} C_{\vec{g}} \exp[2\pi i (\vec{k}_0 + \vec{g}) \cdot \vec{r}] \end{aligned} \quad (2.21)$$

Die Koeffizienten α_j beschreiben die Anregungsamplituden der Blochwellen $B_j(\vec{r})$, die wiederum aus Teilwellen mit der Amplitude $C_{\vec{g}}$ und dem Wellenvektor $\vec{k} = \vec{k}_0 + \vec{g}$ bestehen. Bei der Berechnung der Blochwellen mit N betrachteten Einfallsstrahlen entsteht eine zu lösende $(N \times N)$ -Matrix, die N Eigenwerte besitzt. Diese N Eigenwerte beschreiben N Blochwellen, die jeweils aus N ebenen Wellen bestehen, d.h. die Lösung besteht aus N^2 Teilwellen. Nur für unendliche viele Bragg-Reflexe \vec{g} ist die Lösung des Gleichungssystems exakt, in der Praxis jedoch werden die Anzahl der einfallenden Strahlen je nach Genauigkeitsbedarf begrenzt.

In einer ersten Näherung werden im sogenannten Zweistrahlfall nur zwei Blochwellen betrachtet, was nach Humphreys [25] für Elektronen mit einer Primärenergie bis zu 100 keV ausreichend genaue Ergebnisse liefert. Physikalisch betrachtet bedeutet der Zweistrahlfall, dass neben dem ungebeugten Elektronenstrahl nur noch ein gebeugter Elektronenstrahl existiert. Mathematisch betrachtet ergeben sich vier Lösungen, zwei positive, die die vorwärtsgestreuten Blochwellen repräsentieren und zwei negative, die den rückwärtsgestreuten Blochwellen entsprechen. Da transmittierte, hochenergetische Elektronen hauptsächlich vorwärtsgestreut werden, müssen nur die zwei vorwärtsgestreuten Blochwellen berücksichtigt werden.

Für die Intensitäten des einfallenden (I_0) und des gebeugten (I_g) Elektronenstrahls gilt folgender Zusammenhang:

$$\begin{aligned} I_0(t) &= 1 - I_g(t) \\ I_g(t) &= \left(\frac{\pi}{\xi_g} \right)^2 \cdot \left(\frac{\sin(\pi \cdot t \cdot s_{g,\text{eff}})}{\pi \cdot s_{g,\text{eff}}} \right)^2 \\ \text{und} \\ s_{g,\text{eff}} &= \sqrt{s_g^2 + \xi_g^{-2}} \end{aligned} \tag{2.22}$$

Gleichung 2.22 entspricht Gleichung 2.17 aus der kinematischen Theorie bis auf die Tatsache, dass der Anregungsfehler s_g durch den effektiven Anregungsfehler $s_{g,\text{eff}}$ ersetzt wurde. Dies ist ein wichtiger Unterschied, da im Allgemeinen die Kristallorientierung nahe einer Bragg-Bedingung gewählt wird, so dass der Anregungsfehler klein ist und der effektive Anregungsfehler $s_{g,\text{eff}}$ merklich von s_g abweicht.

Die bisher vorgestellte Theorie gilt nur für elastische Streuung in einem idealen, nicht-absorbierenden Kristall. Zur Berücksichtigung inelastischer Streuung und Absorption wird das bisher reelle Kristallpotential $\varphi(\vec{r})$ als komplex angesehen, wobei der Realteil des Kristallpotentials das Kristallpotential selbst und der Imaginärteil die Absorption beschreibt.

Einige Blochwellen werden äußerst stark absorbiert und einige Blochwellen transmittieren außerordentlich stark, was man als anormale Absorption bzw. anormale Transmission oder auch als „Channelling Effekte“ bezeichnet. Der Grund für anormale Absorption ist darin zu sehen, dass der periodische Imaginärteil des Kristallpotentials an dem Ort der Einheitszelle maximal ist, an der die inelastische Streuung am stärksten ist. Dies ist in der Nähe der Atome der Fall. Es ist daher zu erwarten, dass der Imaginärteil des Kristallpotentials Maxima an den Atomorten aufweist. Somit werden die Blochwellen, die ihre maximale Elektronendichte (Amplituden) an den Atompositionen besitzen (Blochwellentyp I), stärker

absorbiert und diejenigen, die die maximale Elektronendichte zwischen den Atomen erreichen (Blochwellentyp II), nur schwach absorbiert (siehe Bild 2.4).

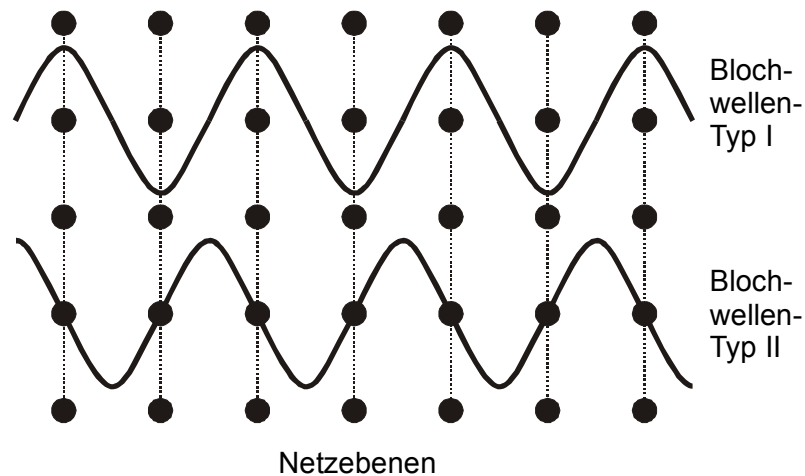


Bild 2.4 : Schematische Darstellung der Blochwellen-Absorption im Zweistrahlfall

Multislice-Theorie

Beim Multislice-Ansatz nach Cowley und Moodie wird ein Kristall der Dicke z in eine Vielzahl dünner Scheiben (engl. Slices) gleicher Dicke Δz geteilt [26]. Der Elektronenstrahl bzw. die Elektronenwelle breitet sich senkrecht zu den Scheiben und parallel zur optischen Achse aus. Unter der Annahme, dass die Scheiben sehr dünn sind, wirken sie auf die einfallende Elektronenwelle wie ein reines Phasenobjekt. Das bedeutet, bei der Transmission durch die Scheibe ändert sich lediglich die Phase der Elektronenwelle, ihre Amplitude hingegen bleibt konstant. Somit wird der Elektronenstrahl von den einzelnen Scheiben rein kinematisch gestreut. Weiter geht man davon aus, dass der Abstand zwischen den Scheiben sehr gering ist, so dass die Ausbreitung der Elektronenwelle zwischen den Scheiben wie im potentialfreien Raum beschrieben werden kann.

Die Transmission der Elektronenwelle durch jede Scheibe kann somit in zwei Schritten beschrieben werden:

Die einfallende Elektronenwelle erfährt innerhalb einer Scheibe eine Phasenänderung $\Delta\phi$. Dies lässt sich durch die Multiplikation der Elektronenwellenfunktion ψ mit der sogenannten Transmissionsfunktion $t(\vec{x})$ beschreiben. Die Transmissionsfunktion gibt die jeweilige lokale Phasenverschiebung wieder, welche vom projizierten Atompotential $v_z(\vec{x})$ abhängt:

$$t(\vec{x}) = \exp[j v_z(\vec{x})] \quad (2.23)$$

Im reziproken Raum entspricht dabei das projizierte Atompotential $v_z(\vec{x})$ gerade der Fouriertransformierten $V_z(\vec{k})$ [27]

$$V_z(\vec{k}) = \frac{\lambda}{F_{\text{Slice}}} \cdot \frac{m}{m_0} \cdot \sum_i f_{ei}(k) \cdot \exp[2\pi j \vec{k} \cdot \vec{x}_i] \quad (2.24)$$

Die Summe beläuft sich über alle Atome innerhalb der Scheibe, wobei die Elektronenstreuamplituden $f_{ei}(k)$ nach Gleichung 2.3 berechnet werden, F_{Slice} ist die Fläche der Scheibe und \vec{x}_i der Ort des i-ten Atoms. Durch die inverse Fouriertransformation von $V_z(\vec{k})$ ergibt sich somit das Atompotential im Ortsraum:

$$v_z(\vec{x}) = \text{FT}^{-1}\{V_z(\vec{k})\} \quad (2.25)$$

Die Ausbreitung der modulierten Welle zwischen den äquidistanten Scheiben entlang der Einstrahlrichtung wird im Ortsraum durch die Faltung der transmittierten Wellenfunktion mit der Ausbreitungsfunktion $p(\vec{x})$ beschrieben.

$$p(\vec{x}, \Delta z) = \exp\left(j \cdot \pi \cdot \frac{x^2 + y^2}{\lambda \cdot \Delta z}\right) \quad (2.26)$$

Die Ausbreitung der Welle durch die gesamte Probe kann nun mit Hilfe der Transmissions- und der Ausbreitungsfunktion rekursiv ermittelt werden und es ergibt sich folgender Ausdruck für die Wellenfunktion nach der Transmission durch die n-te Scheibe (Ortsraum):

$$\psi_n(\vec{x}) = [\psi_{n-1}(\vec{x}) \otimes p(\vec{x}, \Delta z)] \cdot t_n(\vec{x}) \quad (2.27)$$

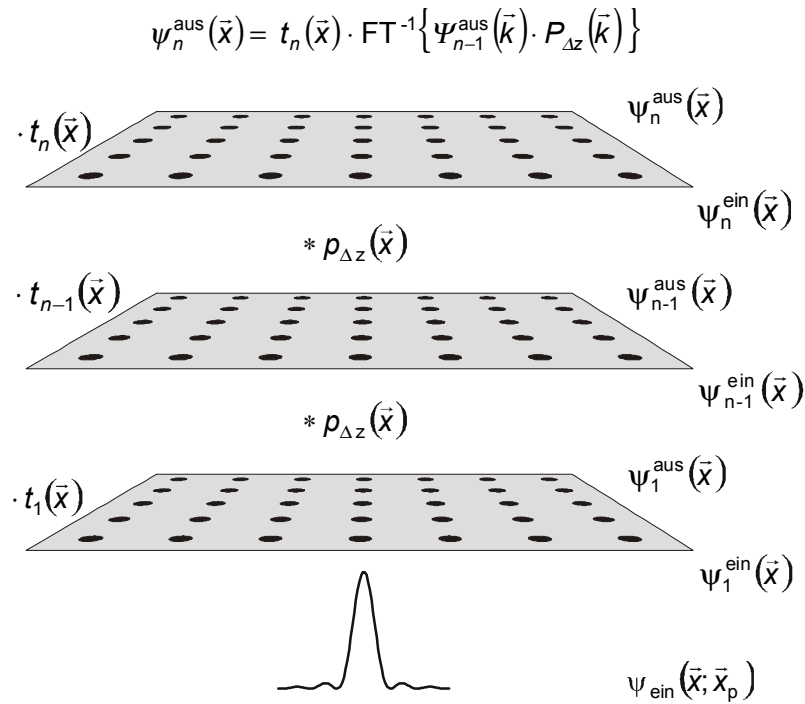
Die Faltung im Ortsbereich entspricht der Multiplikation im reziproken Raum. Da eine Multiplikation mathematisch einfacher zu handhaben ist, wird Gleichung 2.27 in der Regel im reziproken Raum behandelt und ergibt sich zu

$$\psi_n(\vec{x}) = \text{FT}^{-1}[\psi_{n-1}(\vec{k}) \cdot P(\vec{k}, \Delta z)] \cdot t_n(\vec{x}) \quad (2.28)$$

mit

$$P(\vec{k}, \Delta z) = \exp\left[-j \cdot \pi \cdot \Delta z \cdot \lambda \cdot |\vec{k}|^2\right]$$

Bild 2.5 gibt den rekursiven Ablauf nochmals schematisch wieder.

**Bild 2.5 : Die Multislice-Methode**

Beide hier vorgestellten Ansätze werden für die Simulation von elektronenmikroskopischen Abbildungen eingesetzt. Je nach Anwendungsgebiet haben beide Verfahren folgende Vor- bzw. Nachteile: Für die Simulation von Hochauflösungsabbildungen ist der Multislice-Ansatz besser geeignet, insbesondere können Kristalldefekte und Materialübergänge bei der Bildberechnung berücksichtigt werden, jedoch steigt der Rechenaufwand linear mit wachsender Kristalldicke. Der Blochenwellenansatz ist sehr leicht in Simulationsprogramme umzusetzen und eignet sich besonders zur Berechnung von Beugungsmustern eines perfekten Kristallgitters [23]. Die Implementierung, Erweiterung und Anwendung dieser Simulationsansätze werden in den Dissertationen von Bollig [28] und Radefeld [29], die am Fachgebiet „Werkstoffe der Elektrotechnik“ angefertigt wurden, ausführlich erläutert.

2.1.2 Apparativer Aufbau und Funktionsprinzip des Raster-Transmissions-elektronenmikroskops (RTEM)

Bei dem genutzten RTEM handelt es sich um ein im Ultra-Hoch-Vakuum (UHV) arbeitendes, kommerziell erhältliches Mikroskop vom Typ HB 501 der Firma "VG Microscopes". Die elektronenoptische Säule des Mikroskops besteht aus zwei Bereichen, welche durch eine Druckstufe voneinander getrennt sind. Im Bereich der Feldemissionsquelle herrscht durch die Kombination einer Ionengetterpumpe und einer Titan-Sublimationspumpe ein Vakuum von $3 \cdot 10^{-11}$ mbar. Dieser Wert ist Voraussetzung für den stabilen Betrieb der Feldemissionsquelle. Im übrigen Bereich der Säule wird durch den Einsatz einer Diffusionspumpe und einer weiteren Titan-Sublimationspumpe ein Vakuum im Bereich von $5 \cdot 10^{-10}$ mbar erzeugt. Die genannten Vakuumwerte gelten für den Fall, dass das gesamte System vorher ausgeheizt wurde. Mittels einer externen Probenschleuse, die an die Probenkammer gekoppelt ist, kann ein rascher Probenwechsel erfolgen, ohne dass die Probenkammer belüftet werden muss.

Der Elektronenstrahl wird mittels einer kalten Feldemissionsquelle erzeugt. Bei der genutzten Elektronenquelle handelt es sich um eine dünne, in $\{3\ 1\ 0\}$ -Richtung orientierte, einkristalline Wolframspitze. Das Prinzip der kalten Feldemission beruht darauf, dass bei Raumtemperatur zwischen der feinen Spitze und der Extraktionsanode eine Extraktionsspannung, die zwischen 3 kV und 4,5 kV liegt, angelegt wird. Aufgrund des hohen elektrischen Feldes, welches in der Größenordnung von 10^9 V/m liegt, sind die Elektronen in der Lage, die Potentialbarriere zwischen Spitze und Vakuum zu durchtunneln. Anschließend werden die Elektronen durch die nachfolgende Beschleunigungsanode auf eine Primärenergie von 100 keV beschleunigt.

Der so erzeugte Elektronenstrahl weist einen Richtstrahlwert von 10^9 Acm⁻²sr⁻¹ auf, der etwa um den Faktor 10^3 größer ist als der einer herkömmlichen thermischen Elektronenquelle [30]. Somit liefert die Feldemissionsquelle bei kleinen Sondendurchmessern 1000fach höhere Strahlströme, was für ein gutes Signal-Rausch-Verhältnis (SRV) bei den RTEM-Untersuchungen von großer Bedeutung ist. Ein weiterer Vorteil der kalten Feldemissionsquelle gegenüber thermischen Quellen ist der mit $\Delta E = 0,2$ eV Energiebreite nahezu monochromatische Charakter der Elektronensonde. Dies erleichtert die Erzeugung von feinsten Sonden, da bestimmte Abbildungsfehler bei kleinen Energiebreiten minimal werden.

Der Primärelektronenstrahl wird mittels einer Hilfslinse durch eine kleine Apertur, die die beiden Vakuumbereiche voneinander trennt (Druckstufe), hindurchfokussiert. Zur weiteren

Fokussierung und Verkleinerung des Sondendurchmessers dienen magnetische Linsen, und zwar zwei Kondensorlinsen sowie die Objektivlinse. Die sphärische Aberrationskonstante des verwendeten Hochauflösungs-Polschuhs der Objektivlinse wurde zu $C_s = 1,3 \text{ mm}$ bestimmt [28]. Zur Kompensation des Astigmatismus (Unsymmetrien des Magnetfeldes) von Kondensor- und Objektivlinse werden Korrekturlinsen (sog. Stigmatoren) eingesetzt.

Zwischen der zweiten Kondensorlinse und der Objektivlinse befinden sich zwei senkrecht zueinander angeordnete Rasterspulenpaare, die zur Positionierung des Elektronenstrahls während der Abrasterung der Probe dienen. Alternativ wird bei bestimmten Beugungstechniken mit Hilfe der Rasterspulen der Einfallswinkel der Sonde bei annähernd konstantem Aufpunkt auf der Probe moduliert. Der Konvergenzwinkel der Elektronensonde wird durch die Objektivblende bzw. deren Durchmesser bestimmt. Wahlweise können vier verschiedene Blenden in den Strahlengang gebracht werden.

Die Probe ist in einen Probenhalter eingebaut, der mittels Schrittmotoren lateral verschoben werden kann. Die Probe kann außerdem in zwei zueinander senkrechten Richtungen verkippt werden.

In der Probe werden die Elektronen des transmittierenden Strahls elastisch oder inelastisch gestreut (vgl. Kapitel 2.1.1). Zum Nachweis der entstehenden Wechselwirkungsprodukte sind im Strahlengang des RTEM hinter der Probe unterschiedliche Detektoren positioniert. Zur Erfassung der in große Winkel gestreuten Elektronen dient der ringförmige Dunkelfelddetektor (High Angle Annular Dark Field, HAADF). Er besteht aus einer Szintillator-Photomultiplier-Kombination und dient der Aufnahme von Ordnungszahlkontrast-Abbildungen (vgl. Kapitel 2.1.3.2). Als Szintillatormaterial wird ein YAG-Kristall (Yttrium Aluminium Garnet, YAG) verwendet.

Zwischen HAADF-Detektor und Probe kann ein Phosphorschirm (Diffraction Pattern Observation Screen, DPOS) in den Strahlengang eingebracht werden. Mit Hilfe eines Umlenkspiegels wird der Leuchtschirm auf eine CCD-Kamera (Charged Coupled Device, CCD) abgebildet. Der Leuchtschirm dient der Detektion von Beugungsaufnahmen mit konvergentem Elektronenstrahl (Convergent Beam Electron Diffraction, CBED).

Im Rahmen eines Sonderforschungsprojektes der Deutschen Forschungsgesellschaft wurde das Duisburger PEELS-System (Parallel Electron Energy Loss Spectroscopy, PEELS) in Form der sogenannten „Psi-Tube“ aufgebaut, welches von McMullan [31] entwickelt wurde. Die „Psi-Tube“ schließt sich an das magnetische Sektorfeld (Spektrometer) an und ermöglicht auf einfache Weise das schnelle Umschalten des Elektronenstrahls auf drei unterschiedliche Detektionssysteme. Zunächst passieren die inelastisch oder nicht gestreuten Elektronen das Spektrometer. Die Kollektorblende bestimmt den

Kollektionswinkel des Spektrometers. Insgesamt stehen sieben verschiedene Blenden zur Verfügung.

Bei der Nutzung des Hellfelddetektors wird der Einfluss des Spektrometers nicht wirksam. Der Elektronenstrahl passiert das Spektrometer und wird im Anschluss von den Ablenkspulen der „Psi-Tube“ auf den Hellfelddetektor positioniert, der ebenfalls aus einer Szintillator-Photomultiplier-Kombination besteht. Hier besteht der Szintillator aus einem YAP-Kristall (Yttrium Aluminium Perovskit, YAP). Neben der Hellfeldabbildung erlaubt er die Aufnahme von Dunkelfeldabbildungen. Dabei werden die Grigson-Spulen, die direkt hinter der Probe positioniert sind, genutzt. Mittels dieser Spulen können die durch die Probe transmittierten Elektronen gezielt abgelenkt werden, so dass nur ausgewählte Bragg-Reflexe detektiert werden können und somit Beugungs-kontrastabbildungen ermöglichen (vgl. Kapitel 2.1.3.1). Außerdem dient der Hellfelddetektor zur Aufzeichnung von SAD-Beugungs-aufnahmen (Selected Area Diffraction, SAD).

Bei EELS-Messungen dient das Spektrometer zusammen mit den Quadrupollinsen zur Erzeugung der Energiedispersion und der Vergrößerung und Fokussierung der energiedispersiven Ebene. Innerhalb des Sektorfeldes werden die Elektronen durch ein elektrisch isoliertes Metallrohr (Drift-Tube) geleitet. Zur Nachbeschleunigung der Elektronen können an die Drift-Tube Spannungen zwischen 0 V und 2000 V angelegt werden. Dadurch ist es möglich, EEL-Spektren bei unterschiedlichen Energieverlusten und gleicher Dispersion zu erfassen. Die Elektronen treffen hinter dem Sektorfeld und der „Psi-Tube“ auf einen YAG-Szintillator und erzeugen am Auftreffpunkt Licht. Außerhalb des Vakuums wird der YAG-Szintillator mittels einer Lichtoptik, bestehend aus zwei Rodenstocklinsen, auf eine Slow-Scan CCD-Kamera abgebildet. Des Weiteren steht eine Fast-Scan CCD-Kamera, die mittels eines Glasfaserkabels an einen YAG-Szintillator gekoppelt ist, als Detektionseinheit zur Verfügung. Sie dient einerseits zur Justage des PEELS-Systems, andererseits kann sie zur Aufnahme von Beugungsbildern eingesetzt werden. Durch Einbringen des Spaltes ist es außerdem möglich, energiegefilterte Abbildungen zu erzeugen. In Bild 2.6 wird der Aufbau des verwendeten RTEM schematisch dargestellt.

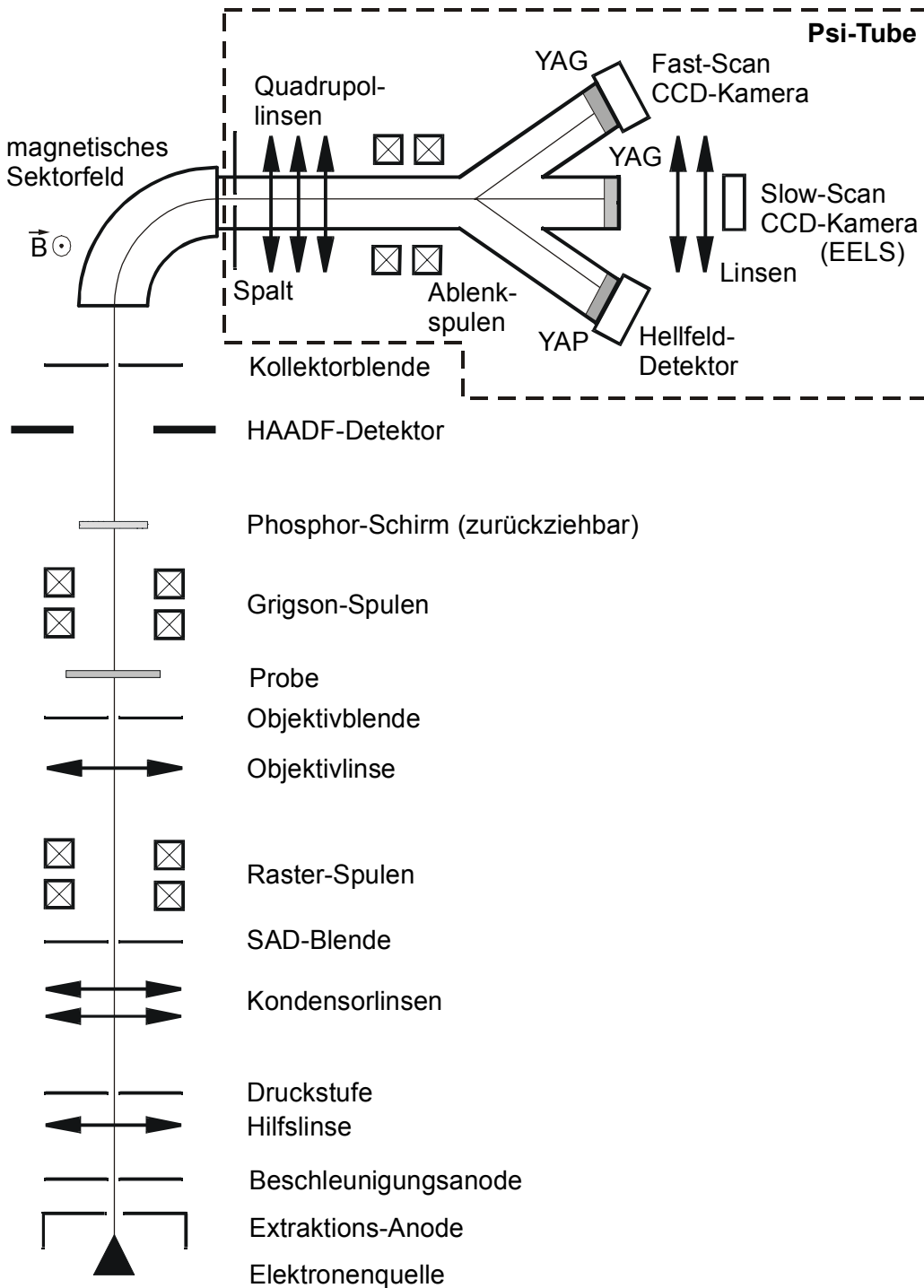


Bild 2.6 :Schematischer Aufbau des Raster-Transmissionselektronenmikroskops

2.1.3 Messtechnische Grundlagen

Im folgenden Kapitel wird die Messtechnik der zur Verfügung stehenden Charakterisierungswerkzeuge beschrieben. Dabei kann generell zwischen abbildenden und analytischen Untersuchungsmethoden unterschieden werden. Bis auf eine Ausnahme (siehe Kapitel 2.1.3.4) gehen sowohl die abbildenden als auch die analytischen Messtechniken von derselben Startvoraussetzung aus. Dies wird in Bild 2.7 verdeutlicht.

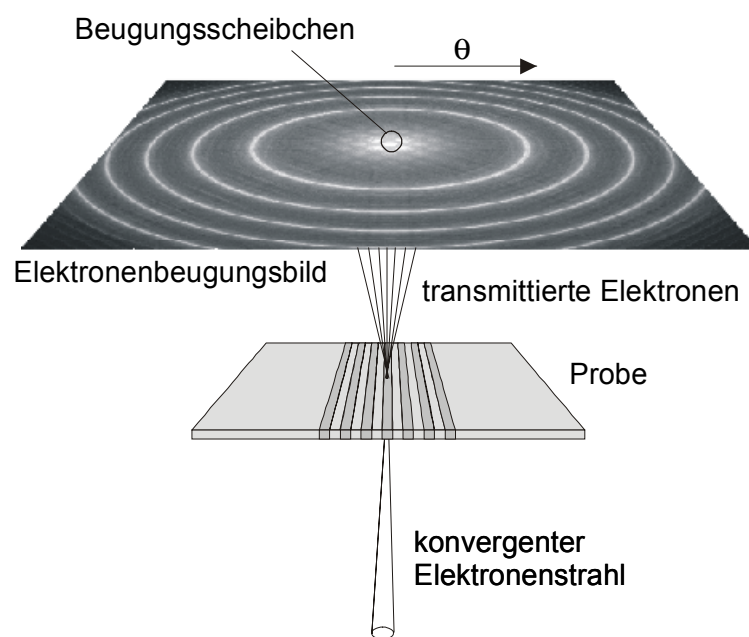


Bild 2.7 : Ausgangspunkt der Detektion der angewandten Messtechniken

Die zu untersuchende Probe (Objekt) wird von einem fein fokussierten, konvergenten Elektronenstrahl durchstrahlt und es entsteht ein Elektronenbeugungsbild (Fouriertransformierte des Objekts), welches die Winkelverteilung der transmittierten Elektronen wiedergibt und dessen Muster abhängig vom Öffnungswinkel der Sonde ist (vgl. Kapitel 2.1.3.4). Entsprechend der Messtechnik werden bestimmte Winkelbereiche bzw. elastisch sowie inelastisch gestreute und/oder gebeugte Elektronen detektiert. Bei den analytischen Messverfahren wird mit einer stationären Sonde gearbeitet. Im Gegensatz dazu wird bei den Abbildungsverfahren der Elektronenstrahl zeilenweise über einen bestimmten Probenbereich gerastert. Dabei sind die Ablenkspulen des RTEM mit den Ablenkspulen des Bildschirms synchronisiert, so dass wie beim konventionellen Raster-Elektronenmikroskop eine

Punkt-zu-Punkt-Abbildung des abgerasterten Bereiches erfolgt. Im Folgenden wird auf die einzelnen Messtechniken detaillierter eingegangen.

2.1.3.1 Hellfeld- und Dunkelfeldabbildung

Bei der Hellfeldabbildung werden die Elektronen, die nicht oder in einem sehr kleinen Winkel ($\theta < 50 \text{ mrad}$) um die optische Achse gestreut worden sind, vom Hellfelddetektor detektiert, der aus einer Szintillator-Photomultiplier-Kombination besteht. Die Objektivblende legt den Öffnungswinkel α der Elektronensonde fest und die Kollektorblende bestimmt die Größe des Raumwinkelbereiches, welcher detektiert wird (vgl. Kapitel 2.1.2).

Je nach Größe des Öffnungswinkels der Elektronensonde wird zwischen der konventionellen Hellfeldabbildung und der Hochauflösungs-Hellfeldabbildung unterschieden. Bei der konventionellen Hellfeldabbildung ist der Öffnungswinkel der Sonde kleiner als der kleinstmögliche doppelte Bragg-Winkel der Probe, so dass sich die Beugungsscheibchen der angeregten Bragg-Reflexe nicht überlappen. Der Kollektorwinkel β_0 des Detektors ist etwa so groß wie der Öffnungswinkel der Sonde, was bedeutet, dass nur der ungebeugte transmittierte Primärstrahl (Nullstrahl) auf den Detektor trifft. Bei idealen Kristallen ist in diesem Fall keine Kontrastentstehung zu erwarten, da die Bragg-Bedingung in der gesamten Probe gleichermaßen erfüllt ist und die gebeugten Wellen nicht miteinander interferieren. Ist jedoch beispielsweise durch Kristalldefekte die Bragg-Bedingung für irgendeinen Bragg-Reflex lokal erfüllt, so wird an dieser Probenstelle ein beträchtlicher Teil der einfallenden Elektronen in diesen Bragg-Reflex gestreut werden. Die Intensität des Nullstrahles wird dadurch am Ort des Kristalldefektes geschwächt und die Probenstelle erscheint im Hellfeldbild dunkel. Bei dem so entstehenden Kontrast spricht man auch von Beugungskontrast. Wird nun anstatt des Nullstrahles die Intensität des lokal angeregten Bragg-Reflexes detektiert, erhält man eine komplementäre Abbildung, in der der Kristalldefekt hell erscheint. Diese Abbildung wird als Dunkelfeldabbildung bezeichnet und ist ebenfalls auf Beugungskontrast zurückzuführen. Technisch wird die Dunkelfeldabbildung im RTEM folgendermaßen realisiert: Durch Anregung der Grigson-Spulen (vgl. Bild 2.6) wird ein ausgewählter Bragg-Reflex auf den Hellfelddetektor umgelenkt, während die Probe mit der Elektronensonde abgerastert wird.

Bei der Hochauflösungs-Hellfeldabbildung muss der Öffnungswinkel der Elektronensonde größer als der doppelte Bragg-Winkel ($\alpha > 2\theta_B$) der entsprechenden Gitterebene, welche abgebildet werden soll, sein. Dadurch kommt es zur Überlagerung der Beugungsscheibchen der einzelnen Bragg-Reflexe und die Reflexe interferieren innerhalb des

Überlappungsbereiches miteinander. Durch die kohärente Überlagerung der Reflexe kommt es zur einer Phasenänderung der Elektronenwelle. Die Phase der transmittierten Elektronenwelle entspricht dem Netzebenenabstand des Kristalls und ermöglicht somit die Abbildung des Kristallgitters mit atomarer Ortsauflösung. Die Hochauflösungs-Hellfeldabbildung wird entsprechend auch kohärente Abbildung genannt und zeigt Phasenkontrast. In Bild 2.8 wird zur Verdeutlichung der Detektionsbereich für die Hellfeld- und Dunkelfeldabbildung gezeigt. Die hier vorgestellten Abbildungsverfahren werden zur Lokalisierung und Identifizierung von strukturellen Defekten in Kristallen eingesetzt.

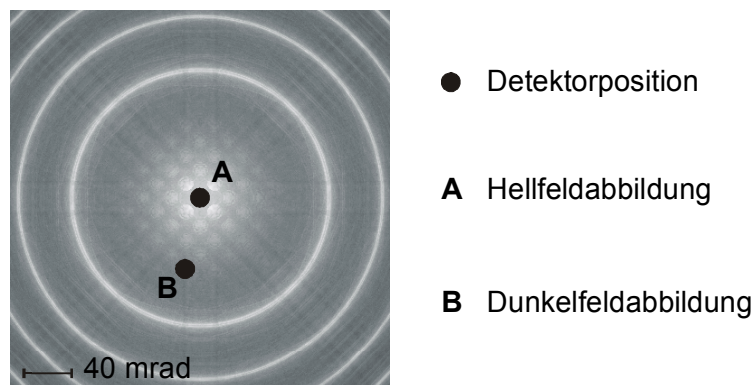


Bild 2.8 : Detektionsbereich für die Hell- und Dunkelfeldabbildung im RTEM

Das erreichbare Auflösungsvermögen hängt nicht nur vom gewählten Abbildungsmodus (d.h. überlappende bzw. nicht überlappende Beugungsscheibchen), sondern auch von den durch das Mikroskop vorgegebenen Betriebs- bzw. Abbildungsbedingungen ab. Der Einfluss der Abbildungsbedingungen lässt sich durch die Kontrastübertragungsfunktion (engl.: Contrast Transfer Funktion, CTF) beschreiben. Die CTF beschreibt die Übertragung der Objektintensität in die Bildintensität und hängt im Wesentlichen von den Linsenfehlern, der Elektronenwellenlänge und dem Defokus (Abweichung von der Lage des exakten Fokus) ab. Aufgrund der CTF können in der Hochauflösungs-Hellfeldabbildung je nach Defokus die Atomsäulen entweder als helle oder dunkle Kontraste erscheinen, d.h. es sind Kontrastumkehrungen möglich. In der Praxis sind möglichst breite Bereiche, in denen der Wert der CTF konstant bleibt, von Bedeutung. Hierbei spricht man vom sogenannten Scherzer-Fokus. Die für den Scherzer-Fokus theoretisch erzielbare Auflösungsgrenze ist nach Scherzer [32]

$$d_{\min, \text{HF}} = 0,66 \cdot \sqrt[4]{C_s \cdot \lambda^3} \quad (2.29)$$

Für das in der Arbeit verwendete RTEM ergibt sich mit $C_s = 1,3 \text{ mm}$ und $\lambda = 3,7 \text{ pm}$ eine theoretisch erzielbare Auflösungsgrenze von $d_{\text{min,HF}} = 0,334 \text{ nm}$.

2.1.3.2 Ordnungszahlkontrastabbildung

Die Ordnungszahlkontrastabbildung (kurz: Z-Kontrastabbildung) beruht auf der Detektion von transmittierten Elektronen, die in einem weiten Winkelbereich ($\theta > 50 \text{ mrad}$) gestreut werden. Bei dem von Crewe et al. entwickelten sogenannten Ordnungszahlkontrastdetektor [33], mit dem es gelang, einzelne Atome schwerer Elemente auf einem Kohlenstofffilm abzubilden, handelt es sich um einen ringförmigen Detektor, der radialsymmetrisch auf der optischen Achse angebracht ist (siehe Kapitel 2.1.2). Dieser erfasst alle Elektronen, die in einem Winkel zwischen $\beta_{\text{min}} \approx 50 \text{ mrad}$ und $\beta_{\text{max}} \approx 180 \text{ mrad}$ gestreut werden (siehe Bild 2.9). Dazu gehören in erster Linie elastisch und quasielastisch gestreute Elektronen. Unter der quasielastischen Streuung versteht man die Wechselwirkung der einfallenden Primärelektronen mit thermisch verursachten Gitterschwingungen. Diese Schwingungen entstehen dadurch, dass die Kristallatome aufgrund ihrer thermischen Energie um ihren lokalen Gitterplatz oszillieren. Die Wechselwirkung der Primärelektronen mit den Gitterschwingungen wird durch die Erzeugung oder Absorption eines Phonons beschrieben. Man spricht deshalb auch von der Phononenstreuung oder der thermisch diffusen Streuung (engl.: Thermal Diffuse Scattering, TDS). Die TDS weist eine sehr weite Winkelverteilung auf, was einen diffusen Untergrund in der Intensitätsverteilung der Beugungsmuster bewirkt.

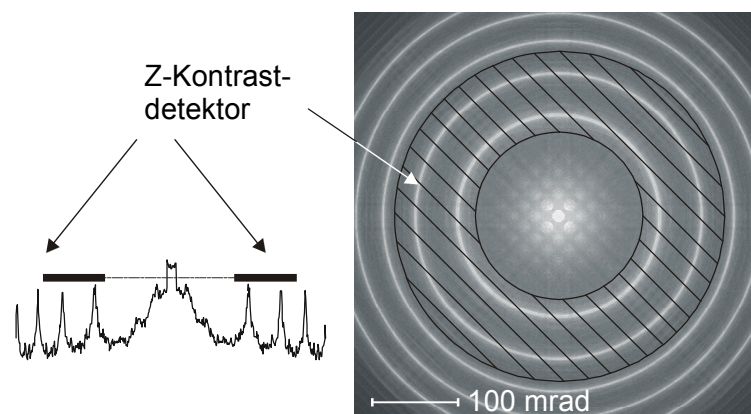


Bild 2.9 : Detektionsbereich für die Ordnungszahlkontrastabbildung

Die vom Ordnungszahlkontrastdetektor detektierten Elektronen können in erster Nähe als Rutherford-gestreut betrachtet werden [34]. Die Rutherford-Streuung ist quadratisch von der

Ordnungszahl des streuenden Elements abhängig (vgl. Gleichung 2.2). Daraus ergibt sich, dass die Intensität eines mit dem Ordnungszahlkontrastdetektor aufgenommenen Bildpunktes näherungsweise proportional dem Quadrat der mittleren Ordnungszahl Z der im durchstrahlten Probenvolumen enthaltenen Elemente ist [35]. Es ergibt sich somit in der Z -Kontrastabbildung: Materialbereiche mit einer höheren mittleren Ordnungszahl erscheinen heller und Materialbereiche mit einer niedrigeren mittleren Ordnungszahl erscheinen dunkler. Somit eignet sich die Z -Kontrastabbildung ausgezeichnet zur qualitativen Bestimmung von chemischen Zusammensetzungen.

Bei der Ordnungszahlkontrastabbildung wird das Intensitätssignal mittels des ringförmigen Detektors über eine Vielzahl sich überlappender Beugungsscheiben inkohärent aufintegriert (In der Literatur spricht man daher auch häufig von der inkohärenten Abbildung). Dadurch ergibt sich die Bildpunktintensität einfach aus der Faltung zwischen der Objektfunktion der Probe, welche die Streukraft der Probenatome repräsentiert, und dem Intensitätsverteilungsprofil der Elektronensonde (engl.: Point Spread Function, PSF). Das Signal ist somit nicht von der Phase der gestreuten Elektronen abhängig. Die Objektfunktion zeigt sehr schmale „Peaks“ an den Positionen der Atomsäulen, welche näherungsweise als Deltafunktionen betrachtet werden können. Daher ist der Kontrast und das erreichbare Auflösungsvermögen vom Intensitätsprofil der Elektronensonde abhängig. Um Atomsäulen innerhalb kristalliner Proben abbilden zu können, muss der Strahldurchmesser kleiner als der aufzulösende Gitterebenenabstand sein. In diesem Fall spricht man von der Hochauflösungs-Ordnungszahlkontrastabbildung. Neben der hinreichend schmalen Sonde müssen für die Hochauflösungsabbildung weitere Bedingungen erfüllt sein. Die Probe muss in einer Zonenachse orientiert sein, um Bedingungen zu erreichen, in denen der Strahl nur entlang einer Atomsäule wandert (Channeling). Beim Channeling wird der Elektronenstrahl durch die Atome, die quasi als Mikrolinsen fungieren, auf die Atomsäule fokussiert [36]. Durch das Verhindern der Strahlaufweitung ist auch bei „dickeren“ Proben - im Gegensatz zur Hellfeld-Hochauflösungsabbildung - Hochauflösung möglich. Des Weiteren ist ein erschütterungsfreier sowie ein stabiler elektrischer Betrieb des Mikroskops erforderlich. Außerdem sollten die Abbildungsfehler des elektrooptischen Linsensystems ausreichend klein gehalten werden (Die letzten beiden Voraussetzungen gelten auch für die Hellfeld-Hochauflösungsabbildung).

In Analogie zur Hellfeld- kann auch für die Hochauflösungs-Ordnungszahlkontrastabbildung die für den Scherzer-Fokus theoretisch erzielbare Auflösungsgrenze angegeben werden:

$$d_{\min,ADF} = 0,43 \cdot \sqrt[4]{C_s \cdot \lambda^3} \quad (2.30)$$

Die theoretisch erzielbare Auflösungsgrenze der Z-Kontrastabbildung des genutzten RTE-M ist somit $d_{\text{min,ADF}} = 0,218 \text{ nm}$. Im Gegensatz zur Kontrastübertragungsfunktion für die Hellfeld-Hochauflösungsabbildung zeigt die CTF für die Ordnungszahlkontrastabbildung keinen Vorzeichenwechsel, sondern einen monoton fallenden Verlauf, das heißt, über einen weiten Defokus-Bereich findet keine Kontrastumkehr statt.

2.1.3.3 Elektronenenergieverlust-Spektroskopie (EELS)

2.1.3.3.1 Funktionsprinzip

Die Elektronenenergieverlust-Spektroskopie untersucht die Energieverteilung von Elektronen, die durch eine sehr dünne Probe transmittiert sind. Bei den zu untersuchenden Elektronen handelt es sich um inelastisch gestreute Elektronen, das heißt um Elektronen, die eine Energieänderung (bzw. Impulsänderung) bei einer sehr geringen Richtungsänderung durch Streuung an der Probe erfahren haben (vgl. Kapitel 2.1.1.2). Die Elektronen werden durch einen energiedispersiven Filter, dem sogenannten magnetischen Sektorfeld, räumlich in ihrer Energieverteilung getrennt. Nachfolgend angeordnete Quadrupollinsen dienen der weiteren Fokussierung des Elektronenstrahls und bestimmen gemeinsam mit dem Sektorfeld die Dispersion des Elektronenspektrometers, das heißt die Breite des Energiefensters, welches bei EELS-Messungen erfasst werden kann. Die Elektronen mit unterschiedlichen Energieverlusten treffen an unterschiedlichen Orten auf einen Szintillator-Kristall (Yttrium Aluminium Granat, YAG) und erzeugen an den Auftreffpunkten Lichtquanten. Die Intensitätsverteilung des YAG, die einem Elektronenenergieverlust-Spektrum entspricht, wird mit einer Lichtoptik auf eine CCD-Kamera abgebildet. Dadurch wird das komplette Energiefenster zeitlich parallel detektiert (siehe Bild 2.6). In der Regel ruht die Elektronensonde während der Aufzeichnung des EEL-Spektrums an einer bestimmten Probenposition. Daher lässt sich die gemessene Energieverteilung der bestrahlten Probenposition zuordnen und ermöglicht eine orts aufgelöste Analyse des Spektrums. Eine detaillierte Beschreibung des gesamten Messsystems für die Elektronenenergieverlust-Spektroskopie ist in [37] zu finden.

2.1.3.3.2 Das Elektronenenergieverlust-Spektrum

Ein typisches, schematisches Energieverlust-Spektrum ist in Bild 2.10 dargestellt. Es wird wie in der Literatur üblich [z.B.18] in drei Bereiche unterteilt,

- den Zero-Loss-Bereich,
- den Low-Loss-Bereich und
- den Core-Loss-Bereich.

Im ersten Bereich (Zero-Loss-Bereich) bei $\Delta E \approx 0 \text{ eV}$ befindet sich ein schmaler intensitätsreicher Peak, der sogenannte "Zero-Loss-Peak". Dieser Peak beinhaltet Beiträge aus drei Gruppen von Elektronen. Die erste Gruppe besteht aus ungestreuten Elektronen, welche die Probe passiert haben, ohne in Wechselwirkung mit der Probe zu treten und folglich keine Bedeutung für die Mikroanalyse haben. Zur zweiten Gruppe zählen die elastisch gestreuten Elektronen, die nur eine Richtungsänderung erfahren. Die dritte Gruppe besteht aus Elektronen, die eine Phononenanregung erzeugen. Die Phononenanregung tritt in allen Materialien auf und ist insbesondere bei kristallinen Materialien von Bedeutung. Die Energieverluste durch die Phononenanregung sind sehr klein und liegen im Bereich kleiner 0,1 eV. Sie erlauben keine chemische Analyse der Probe, führen jedoch zu einer Reduktion der inelastisch gestreuten Elektronen, die auf den Analysator treffen. Trotz seines geringen Informationsgehaltes zur quantitativen Analyse sollte der Zero-Loss-Peak nicht außer Acht gelassen werden. So ist die Halbwertsbreite des Zero-Loss-Peaks ein Maß für das Energieauflösungsvermögen des gesamten Systems. Die Position des Zero-Loss-Peaks selbst dient zur Kalibrierung des Nullpunktes der Spektren.

Der zweite Bereich (Low-Loss-Bereich) deckt den Energieverlustbereich bis etwa 50 eV ab. Typischerweise besitzt die Intensität des Signals in dieser Region nur noch 5 % bis 10 % von der Intensität des Zero-Loss-Peaks. Die Energieverluste im Low-Loss-Bereich setzen sich aus Interbandanregungen (0 eV bis 8 eV) und sogenannten Plasmonenanregungen (15 eV bis 40 eV) zusammen. Der Interbandübergang entspricht bei Halbleitern und Isolatoren dem Energieverlust in der Größe der Bandlücke. Es ist somit möglich, aus den Low-Loss-Spektren Informationen über die lokale elektronische und optoelektronische Struktur der Probe zu erhalten. Mit Hilfe der Kramers-Kronig-Analyse ist es darüber hinaus möglich, Aufschlüsse über die dielektrischen Eigenschaften der Probe zu gewinnen. Eine detaillierte Darstellung der Kramers-Kronig-Analyse an EEL-Spektren ist bei Rafferty [38] zu finden.

Die Plasmonenanregung tritt bevorzugt in Metallen und Legierungen mit einer hohen Anzahl freier Elektronen auf. Diese freien Elektronen verhalten sich wie ein Elektronengas, wobei aufgrund der Coulombabstoßung dieser freien Elektronen ein Gleichgewicht der Elektronendichte erzeugt wird. Ein schnelles Elektron, das in dieses Gas eindringt, stört das Gleichgewicht und das Elektronengas beginnt zu oszillieren. Da im Prinzip alle Valenzelektronen an diesem Vorgang beteiligt sind, spricht man auch von der sogenannten "kollektiven Anregung". Prinzipiell ist eine chemische Analyse möglich, da sich aber die Plasmonenverluste unterschiedlicher Materialien stark überlappen und sich spektral nur geringfügig voneinander unterscheiden, wird von einer chemischen Analyse in der Regel abgesehen. Mit Hilfe der Plasmonenanregung kann die Dicke der zu untersuchenden Probe bestimmt werden.

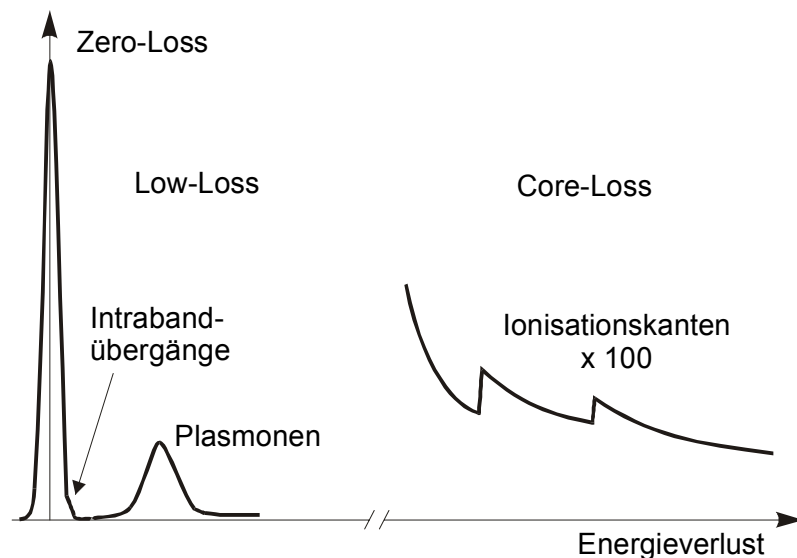


Bild 2.10 : Schematisches Elektronenenergieverlust-Spektrum

Der dritte charakteristische Bereich (Core-Loss-Bereich) des Elektronenenergieverlust-Spektrums liegt bei Energien oberhalb von 50 eV. Dieser Bereich ist für die quantitative Bestimmung chemischer Materialzusammensetzungen von entscheidender Bedeutung. In dieser Region entsteht der Energieverlust der Primärelektronen durch die Ionisation der Elektronen der inneren Schalen der Probenatome. Dabei entstehen elementspezifische charakteristische Ionisationskanten mit einem steilen Anstieg bei der entsprechenden Ionisationsenergie und einem danach exponentiell abklingenden Verlauf. Der abklingende Verlauf stellt die Untergrundintensität dar, die durch die Ionisation von Valenzelektronen erzeugt wird. Die Intensität der Ionisationskanten ist im Vergleich zum Zero-Loss-Peak und dem Low-Loss-Bereich wesentlich geringer, so dass das Energieverlustspektrum aufgrund seiner Intensitätsabnahme zu höheren Energien hin begrenzt ist.

2.1.3.3.3 Quantitative Analyse von EEL-Spektren

Bei der quantitativen Analyse nutzt man die Tatsache aus, dass die Intensität der Ionisationskante ein direktes Maß für die zur Kante gehörenden Elementkonzentration an der untersuchten Probenstelle ist. Da die Kante jedoch von einem Untergrund überlagert ist, ist es notwendig, diesen Untergrund von der jeweiligen Kante zu subtrahieren. Der Untergrund entsteht durch die Ionisation von Valenzelektronen. Zusätzlich wird die Intensität des Untergrunds durch den Dunkelstrom des eingesetzten Detektors beeinflusst. Der Untergrund fällt in erster Näherung mit einem Potenzgesetz gemäß

$$I_B = A \Delta E^{-r} \quad (2.31)$$

ab [18]. Die Parameter A und r der Anpassung des Untergrunds sind im EEL-Spektrum in einem Intervall der Breite ΔE kurz vor der Ionisationskante zu ermitteln. Zur Quantifizierung der Ionisationskante wird nun die angepasste Untergrundfunktion bis unter die Kante extrapoliert und vom gemessenen Spektrum subtrahiert. Nun kann die eigentliche Kantenintensität I_K durch einfache numerische Integration im Integrationsbereich ΔE bestimmt werden. Die Bestimmung der tatsächlichen Elementkonzentration erfolgt über den Zusammenhang

$$N_x = \frac{I_K}{I_0 \sigma_K} \quad (2.32)$$

Hierbei ist N_x die Anzahl der Atome/cm², I_K die Intensität der Kante in Abhängigkeit vom Integrationsbereich ΔE , I_0 die Intensität des einfallenden Primärelektronenstrahls und σ_K der Wirkungsquerschnitt (pro Atom). Um Vielfachstreuungen auszugleichen, muss I_0 durch I_1 ersetzt werden (siehe Bild 2.11). Somit wird Gleichung 2.32 zu

$$N_x = \frac{I_K}{I_1 \sigma_K} \quad (2.33)$$

Die relative Elementkonzentration von zwei Elementen A und B kann einfach durch das Verhältnis

$$\frac{N_A}{N_B} = \frac{\frac{I_{K,A}}{I_1 \cdot \sigma_{K,A}}}{\frac{I_{K,B}}{I_1 \cdot \sigma_{K,B}}} = \frac{I_{K,A} \cdot \sigma_{K,B}}{I_{K,B} \cdot \sigma_{K,A}} \quad (2.34)$$

bestimmt werden. Hierbei wird nicht mehr das Zero-Loss-Spektrum benötigt, sondern es ist möglich, das atomare Verhältnis der Elemente A und B aus einem Spektrum, das beide spezifische Kanten enthält, zu bestimmen. Die Streuquerschnitte können dabei entweder rechnerisch oder bei bekannter Zusammensetzung experimentell bestimmt werden [39].

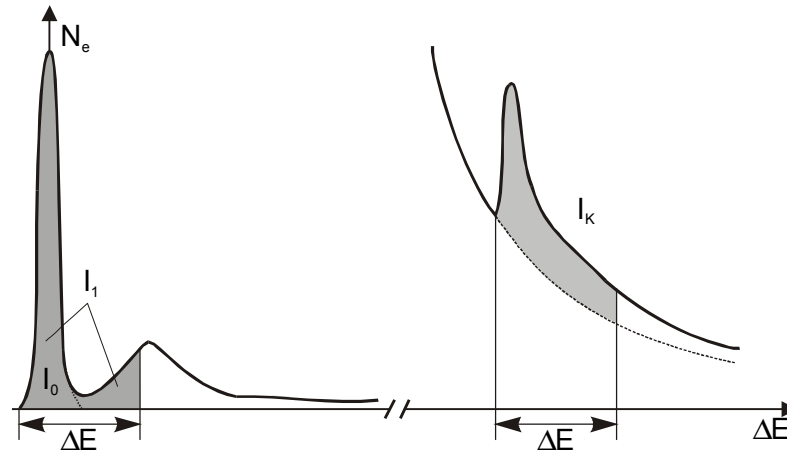


Bild 2.11 : Schematische Darstellung zur Bestimmung der Elementkonzentration

2.1.3.4 Elektronenbeugungstechniken

Zur Abbildung von Elektronenbeugungsmustern in einem RTEM werden im Allgemeinen zwei unterschiedliche Methoden genutzt. Diese sind die Feinbereichsbeugung (engl.: Selected Area Diffraction, SAD) und die Elektronenbeugung im konvergenten Bündel (engl.: Convergent Beam Electron Diffraction, CBED).

Bei der Feinbereichsbeugung wird der zu untersuchende Probenbereich durch die SAD-Blende (vgl. Kapitel 2.1.2) definiert. Die Elektronensonde bleibt fest auf einer Probenposition stehen, während die Einfallsrichtung des Elektronenstrahls durch die Raster-Spulen moduliert wird. Durch diesen technischen Trick wird die Probe von einem nahezu parallelen Elektronenstrahlbündel bestrahlt, d.h. der Konvergenzwinkel und somit der Sondendurchmesser ist relativ groß. Daher beträgt die Ortsauflösung im SAD-Modus nur etwa 200 nm [28]. Auf dem Detektor (DPOS oder Hellfelddetektor) entsteht ein nahezu diskretes Punktmuster, bei dem jeder Punkt einem Bragg-Reflex entspricht. Bild 2.12a) zeigt ein Beispiel eines SAD-Beugungsmuster einer InP-Probe entlang der $\langle 011 \rangle$ -Zonenachse. Der detektierte Winkelbereich wurde hier so groß gewählt, dass der sogenannte FOLZ-Ring

(engl.: First Order Laue Zone, FOLZ) noch erfasst ist. Die SAD-Methode dient u.a. zur exakten Orientierung einer Zonenachse der zu untersuchenden Probe relativ zur optischen Achse des Mikroskops. Darüber hinaus gibt ein SAD-Beugungsmuster Aufschluss über die Kristallstruktur und die Größe der Einheitszelle.

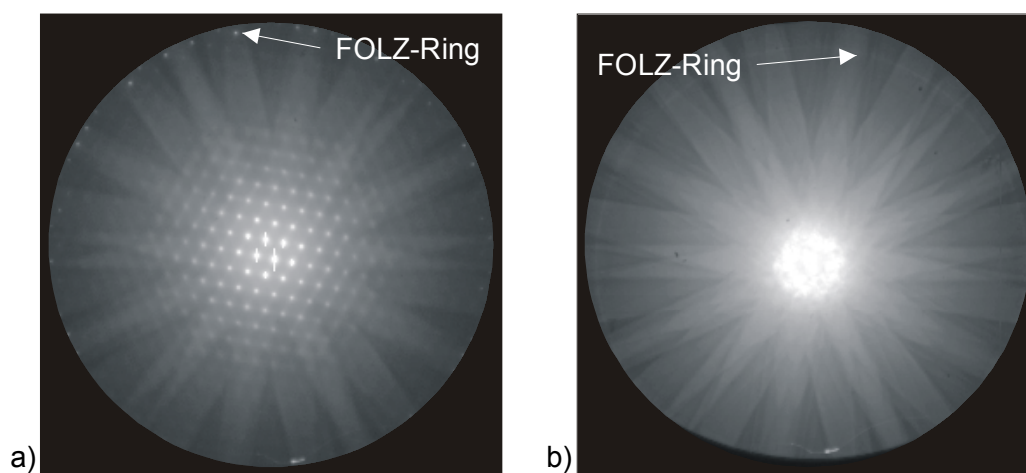


Bild 2.12 : a) SAD-Beugungsmuster einer InP-Probe entlang der $\langle 011 \rangle$ -Zonenachse
b) CBED-Beugungsmuster einer InP-Probe entlang der $\langle 011 \rangle$ -Zonenachse

Bei der CBED-Messtechnik wird die Probe ortsfest mit einem konvergenten Elektronenstrahl bestrahlt. Die erreichbare laterale Ortsauflösung liegt dabei im Bereich von etwa 1 nm [40]. Das so erzeugte Elektronenbeugungsbild besteht aus kreisförmigen Scheiben, den sogenannten Bragg-Scheiben. Diese entsprechen Bragg-Reflexen, die sich zu eben diesen Scheiben aufgeweitet haben. Die Bragg-Scheiben in der Mitte des CBED-Beugungsmusters bilden die Laue-Zone nullter Ordnung, die auch ZOLZ (engl.: Zero Order Laue Zone) genannt wird. Die weiter außen liegenden Bragg-Scheiben entarten zu Linien, die kreisförmig um die Laue-Zone nullter Ordnung angeordnet sind. Sie gehören zu den Laue-Zonen höherer Ordnung, auch HOLZ (engl.: Higher Order Laue Zone) genannt, die einzeln nach ihrer Ordnung unterschieden werden (FOLZ: First Order Laue Zone, SOLZ: Second Order Laue Zone, etc.). Innerhalb der Laue-Zone nullter Ordnung entstehen linienförmige Intensitätsminima, die sogenannten HOLZ-Linien, die gemäß dem Bragg'schen Gesetz Intensitätsmaxima innerhalb der Higher Order Laue Zone entsprechen. Neben den HOLZ-Linien entstehen im Beugungsbild die sogenannten Kikuchilinen bzw. -bänder. Diese entstehen durch die elastische Streuung von zuvor inelastisch gestreuten Elektronen. Wird der Kristall aus der Zonenachse verkippt, bleiben die Bragg-Scheiben an ihrer Position innerhalb des Beugungsbildes. Lediglich ihre Intensität variiert. Kikuchi- und HOLZ-Linien

sind dagegen an den Kristall gebunden und bewegen sich mit der Verkipfung im Beugungsbild. Die Lage der Linien ist charakteristisch für den kristallinen Aufbau der Probe. Die CBED-Methode eignet sich zur Bestimmung der Probenorientierung und wird über die Verschiebung von bestimmten HOLZ-Linien mit dem Grad der Verspannung zur Ermittlung lokaler Gitterverspannungen mit einer Genauigkeit von $\Delta a/a \geq 10^{-3}$ eingesetzt [41], [42]. Bild 2.12b) zeigt ein Beispiel eines CBED-Beugungsmuster einer InP-Probe entlang der $\langle 011 \rangle$ -Zonenachse. Der detektierte Winkelbereich wurde wie bei dem SAD-Beugungsmuster aus Bild 2.12 a) so groß gewählt, dass der FOLZ-Ring noch erfasst ist.

2.1.4 Maximum-Entropie-Analyse von Hochauflösungs-Ordnungszahlkontrastabbildungen

In Kapitel 2.1.3.2 wurde der physikalische Hintergrund und die Messtechnik der Ordnungszahlkontrastabbildung beschrieben. Diese kann wie bereits erwähnt als Faltung zwischen der Objektfunktion $O(R)$ und dem Intensitätsverteilungsprofil der effektiven Elektronenprobe $P_{\text{eff}}^2(R)$ beschrieben werden (siehe Bild 2.13). In der Realität muss zum Bildsignal $I(R)$ noch ein Rauschsignal $N(R)$ additiv hinzugefügt werden:

$$I(R) = O(R) * P_{\text{eff}}^2(R) + N(R) \quad (2.35)$$

Durch die Anwendung des sogenannten Maximum-Entropie-Verfahrens (MEV) ist es möglich, die Objektfunktion aus einem experimentellen Bild zu rekonstruieren und somit wichtige zusätzliche Informationen über die abgebildeten atomaren Strukturen und die Streuintensität der Atomsäulen zu erhalten.

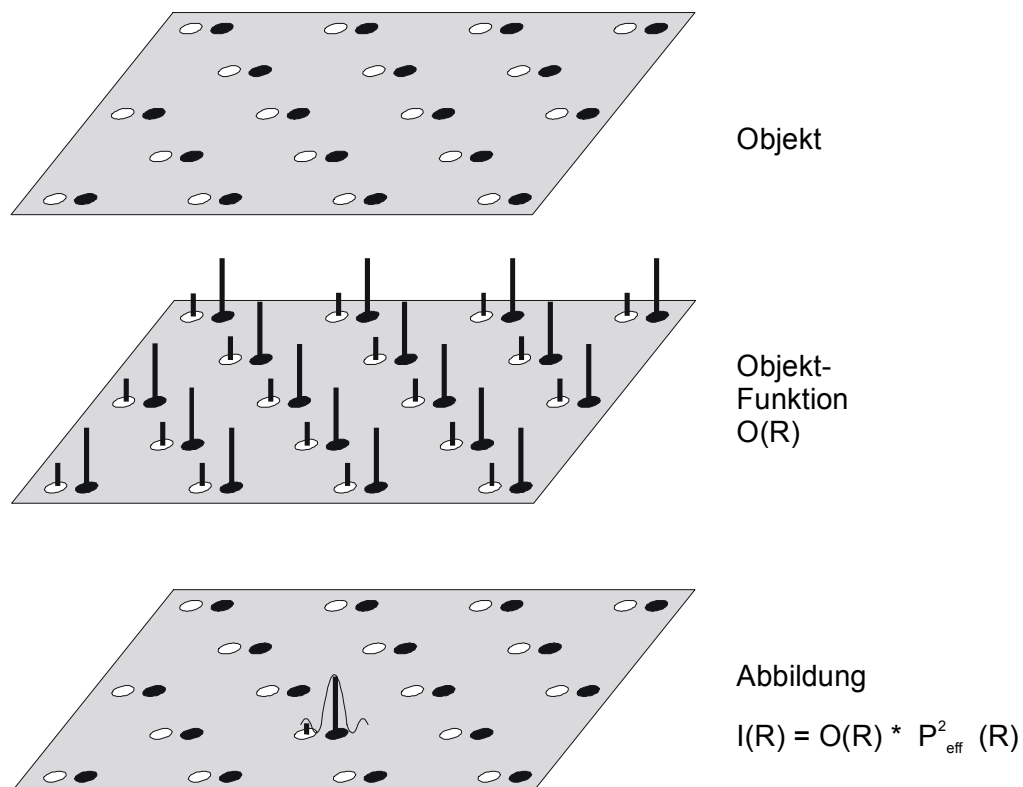


Bild 2.13 : Schematische Darstellung zur Interpretation über die Entstehung der Z-Kontrastabbildung nach McGibbon [43]

Die folgende Beschreibung des theoretischen Hintergrundes des MEV bezieht sich auf ein Datenbild, welches aus N Bildpunkten (engl.: Pixel) besteht. Ausgehend von einem experimentellen Hochauflösungs-Ordnungszahlkontrastbild ist es das Ziel, die wahrscheinlichste Lösung h der Objektfunktion m des Z-Kontrastbildes zu ermitteln. Bei einem gegebenen Bild gibt es aufgrund experimenteller Fehler eine Vielzahl möglicher Objektfunktionen. Um die wahrscheinlichste Objektfunktion zu bestimmen, ermittelt man die Wahrscheinlichkeit L , die durch die aus der Statistik bekannten χ^2 -Verteilung

$$L(h) = \frac{\chi^2(h)}{2} = \frac{1}{2} \cdot \sum_{i=1}^N \frac{[M(h)_i - I_i]^2}{\sigma_i^2} \quad (2.36)$$

gegeben ist. $M(h)$ repräsentiert sogenannte „Mock“-Daten (engl.: Schein-, Pseudo-), welche der Faltung der Objektfunktion, überprüft mit dem vorgegebenen Intensitätsprofil der Elektronensonde, entspricht. I gibt die experimentellen Bildintensitäten wieder und σ_i ist die Standardabweichung der Intensität des i -ten Pixels.

In einem einfachen Beispiel soll der Maximum-Entropie-Ansatz anhand eines zwei-dimensionalen Beispiels verdeutlicht werden. Bild 2.14 zeigt ein Diagramm, in dem zwei Punkte $m = (m_x, m_y)$ und $h = (h_x, h_y)$ eingetragen sind. Man erkennt, wie sich durch Anwendung des MEV der anfänglich angenommene Wert m zum Wert mit der bestmöglichen Wahrscheinlichkeit h entlang der sogenannten Maximum-Entropie-Trajektorie bewegt.

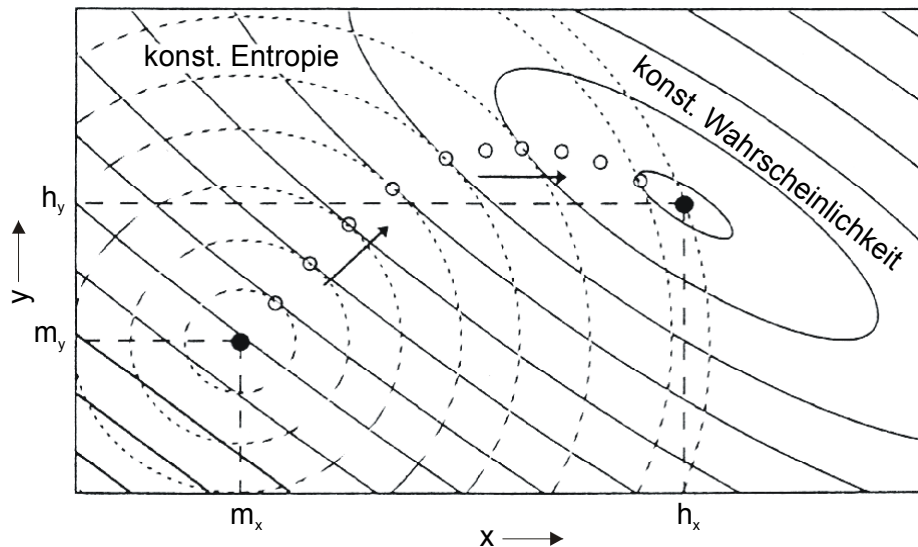


Bild 2.14: Schematische Darstellung der Maximum-Entropie-Trajektorie nach McGibbon [43]

Man erkennt ein Minimum von L im Punkt h und elipsenförmige, durchgezogene Linien, die das Minimum umgeben. An allen Punkten außer am Punkt h gibt es mehrere Lösungen gleicher Wahrscheinlichkeit L . Daher ist ein Kriterium zur Auswahl der bestmöglichen Objektfunktion notwendig. Ein solches Kriterium ist die Maximierung der Entropie S , die durch die Gleichung 2.37

$$S(h) = \sum_{i=1}^N h_i - m_i - h_i \log \left(\frac{h_i}{m_i} \right) \quad (2.37)$$

beschrieben wird [44]. m ist der Anfangswert der vorher definierten Objektfunktion, wobei alle m_i die gleiche Intensität besitzen. Daher ist m ein globales Maximum und alle Werte gleicher Entropie sind kreisförmig um m angeordnet.

Durch die Kombination der Kriterien Entropie S und Wahrscheinlichkeit L ergibt sich die Maximum-Entropie-Bedingung, nämlich die Maximierung des Wertes der „erzwungenen“ Entropie $Q(h)$

$$Q(h) = \alpha S(h) - L(h) \quad (2.38)$$

α ist ein Regulierungsparameter.

Die Beurteilung von $Q(h)$ geschieht mittels eines Iterationsverfahrens. Der Iterationsprozess beginnt mit der Annahme, dass die Objektfunktion maximale Entropie besitzt, d.h. $h = m$, wobei alle m_i den gleichen Wert besitzen. Zu Beginn wird der Wert des Regulierungsparameters α zu unendlich gesetzt. Für jeden Wert von α wird die Untermenge der wahrscheinlichen Objektfunktionen auf der Basis des χ^2 -Kriteriums bestimmt. Von dieser Teilmenge wird dann die Objektfunktion mit der maximalen Entropie iterativ bestimmt. Dies wird in Bild 2.13 durch die Kreise veranschaulicht, die dort positioniert sind, wo bei einem konstanten Wert von L die höchste Entropie auftritt. Der Iterationsprozess wird unter der Verwendung des neu „gefundenen“ Wertes für h und einem reduzierten Wert α nun wiederholt. Alle so gefundenen finalen Werte von h eines Iterationsschrittes ergeben die in Bild 2.14 dargestellte Maximum-Entropie-Trajektorie.

Zur Kontrolle des Iterationsprozesses finden Änderungen des Regulierungsparameters α und Änderungen der Iteration statt, um bei einem vorgegebenen Wert die maximale Entropie $S(h)$ zu finden, innerhalb einer sogenannten „Trust Region“, dessen Radius umgekehrt proportional zum Wert der Entropie ist. Konsequenterweise wird daher, je näher die Entropie

dem Wert Null kommt, die Schrittweite der sich schrittweise verringernden Werte von α immer kleiner, bis schließlich ein Endwert erreicht ist, nach dem es nur noch sehr wenig neue Informationen über die zu bestimmende Objektfunktion gibt. Das heißt, der Endwert korrespondiert mit dem Zustand, in dem die „Hauptmerkmale“ des experimentellen Bildes reproduziert werden und weitere Iterationsschritte nur minimale Verbesserungen erzielen würden, die in keinem Verhältnis zur benötigten Rechenzeit stehen [45]. An diesem Punkt besteht die Maximum-Entropie-Objektfunktion aus einem Feld von Datenpunkten, deren Intensität und Position direkt mit dem experimentellen Bild durch Faltung mit der Elektronensonde korreliert werden kann.

2.2 Thermodynamische Grundlagen zur Mischungslücke im III-V-Halbleitersystem

Zur Herstellung einkristalliner Halbleiterschichten werden Epitaxieverfahren eingesetzt. Das Wort Epitaxie leitet sich aus dem Griechischen ab (epi: auf, über und taxis: Ordnung) und bedeutet das Wachstum einer dünnen kristallinen Schicht auf ein Trägermaterial (Substrat), wobei die aufgebrachte Schicht die kristalline Ordnung des Substrats annimmt. Epitaxie bedeutet im Grunde die Erzeugung "kontrollierter Phasenübergänge", welche zu einem kristallinen Festkörper führen. Das Wachstum solcher aus unterschiedlichen Materialien bestehenden Halbleiterschichten bzw. allgemeiner, die Erstarrung eines Festkörpers aus einer fluiden Phase, ist thermodynamischen Gesetzen unterworfen. Es hat sich gezeigt, dass das kontrollierte Wachstum von Mischkristallen auf bestimmte Konzentrationsintervalle beschränkt ist. Bewegt man sich außerhalb dieser Materialkonzentrationen, wird die homogene Mischkristallphase instabil und es entstehen zwei stabile Phasen unterschiedlicher Zusammensetzung und Kristallstruktur. Dieses Intervall wird auch als Mischungslücke bezeichnet. Die Ausdehnung solcher Mischungslücken ist temperaturabhängig. Mit steigender Temperatur wird die Ausdehnung der Mischungslücke immer kleiner, bis sie bei der sogenannten kritischen Temperatur verschwindet. Treibende Kraft für die Entmischung ist die Differenz der kovalenten Bindungsradien der beteiligten Elemente. Zur Berechnung von Mischungslücken in quaternären III-V-Halbleitersystemen existieren verschiedene Ansätze. Da die in dieser Arbeit zugrunde gelegten Diagramme nach dem sogenannten „Delta-Lattice-Parameter-Modell“ (DLP-Modell) nach Stringfellow [46] berechnet wurden, soll nur dieses Modell vorgestellt werden.

2.2.1 Das Delta-Lattice-Parameter-Modell nach Stringfellow

2.2.1.1 Die freie Energie

Das chemische Phänomen der Mischung wird durch die freie Energie beschrieben [47], welche durch die Entropie S und die Enthalpie H (oder die innere Energie U) definiert ist. G.B. Stringfellow geht in seinem Ansatz von konstantem Druck innerhalb des Systems aus und legt seinen Berechnung damit die Gibbs'sche freie Energie G zugrunde:

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

$$\text{bzw.} \quad G = H - TS \quad (2.39)$$

Das einfachste Modell zur Beschreibung der freien Energie eines Mischkristall-halbleitersystems ist die sogenannte „reguläre Lösung“. Sie geht davon aus, dass die Atome statistisch auf den Untergittern verteilt sind. Eine weitere Annahme ist, dass die Wechselwirkung zwischen den beteiligten Atomsorten auf die nächsten und übernächsten Nachbarn begrenzt ist. Diese Wechselwirkungsenergien werden von Stringfellow als Funktionen der Gitterkonstanten der binären Komponenten definiert, welche in dem quaternären System vorkommen. Für Halbleitermaterialien mit einem hohen Grad an kovalenter Bindung wie die III-V-Halbleiter kann die Gibbs'sche freie Energie pro Mol für ein quaternäres System des Typs $A_xB_{1-x}C_yD_{1-y}$ durch folgenden Term ausgedrückt werden:

$$\underbrace{G^{\text{FK}} = -Ka_0^{-2,5}}_{\text{Enthalpie}} + \underbrace{RT[x\ln x + (1-x)\ln(1-x) + y\ln y + (1-y)\ln(1-y)]}_{\text{Mischentropie}} \quad (2.40)$$

mit

$$K = 1,15 \cdot 10^7 \frac{\text{cal}}{\text{mol}} \text{\AA}^{2,5} \quad \text{empirisch ermittelte Konstante}$$

$$a_0 \quad \text{Gitterkonstante nach dem Vegard'schen Gesetz}$$

$$R \quad \text{Gaskonstante}$$

Analog ergibt sich für ein binäres System:

$$G^{\text{FK}} = -Ka_0^{-2,5} + RT[x\ln x + (1-x)\ln(1-x)] \quad (2.41)$$

2.2.1.2 Spinodale und binodale Kurve

Bei einem zweiphasigen System besitzt G im thermodynamischen Gleichgewicht ein Minimum. Die Gleichgewichtsbedingung ist über die partielle Ableitung von G definiert und wird als chemisches Potential μ bezeichnet. Im Gleichgewichtsfall sind daher die chemischen Potentiale der Phase A und Phase B gleich. Es gilt:

$$\mu_A = \left. \frac{\partial G}{\partial x} \right|_{x_A} = \left. \frac{\partial G}{\partial x} \right|_{x_B} = \mu_B \quad (2.42)$$

Der Begriff des chemischen Potentials soll dem Verständnis zur Definition der binodalen Kurve dienen. Zunächst wird auf den einfacheren binären Fall eingegangen.

2.2.1.2.1 Spinodale und Binodale im binären System

Das obere Diagramm in Bild 2.15 zeigt die freie Gibbs'sche Energie als Funktion der Zusammensetzung. Die gezeigte Kurve besitzt zwei Wendepunkte S_1 und S_2 . Diese Wendepunkte korrespondieren mit der gestrichelten Kurve des unteren Diagramms, in der die Punkte S_1 und S_2 bei unterschiedlichen Temperaturen aufgetragen wurden. Diese Kurve wird als spinodale Kurve oder Spinodale bezeichnet. Den mathematische Ausdruck für die Spinodale liefert die zweifache partielle Ableitung von G :

$$\frac{\partial^2 G}{\partial x^2} = 0 \quad (2.43)$$

Die Punkte B_1 und B_2 erhält man durch Anlegen einer Doppeltangente an die Funktion. In den Punkten B_1 und B_2 ist die Steigung von G , also das chemische Potential, gleich und somit die Gleichgewichtsbedingung erfüllt. B_1 und B_2 korrespondieren mit der durchgezogenen Kurve des unteren Diagramms. Diese Kurve bezeichnet man als binodale Kurve oder Binodale.

Der Verlauf der Kurven zeigt drei unterschiedliche Stabilitätsbereiche des Materials in Abhängigkeit von der Zusammensetzung:

- stabile Bereiche in der Nähe der reinen Komponenten,
- metastabile Bereiche und einen
- instabilen Bereich.

Per Definition versteht man unter der thermodynamischen Mischungslücke den instabilen und metastabilen Bereich, der von der binodalen Kurve begrenzt wird [10]. Zwischen den Punkten A und B_1 sowie B_2 und C (bzw. außerhalb der Binodalen) liegt der stabile Bereich, die Materialkomponenten sind vollständig mischbar. Der Bereich zwischen den Punkten S_1 und S_2 (bzw. innerhalb der Spinodalen) ist thermodynamisch instabil. Es kommt zur spontanen Entmischung, da schon kleine Änderungen in der Zusammensetzung zur Verringerung der freien Gibbs'schen Energie führen. Zwischen B_1 und S_1 sowie S_2 und B_2 (Gebiet zwischen Binodale und Spinodale) liegt der metastabile Bereich. Dort kann durch

Entmischung die freie Gibbs'sche Energie verringert werden, es ist jedoch eine Anregungsenergie für die Entmischung nötig. Die Punkte B_1 und B_2 stellen die Zusammensetzungen dar, in die sich das Materialsystem entmischt.

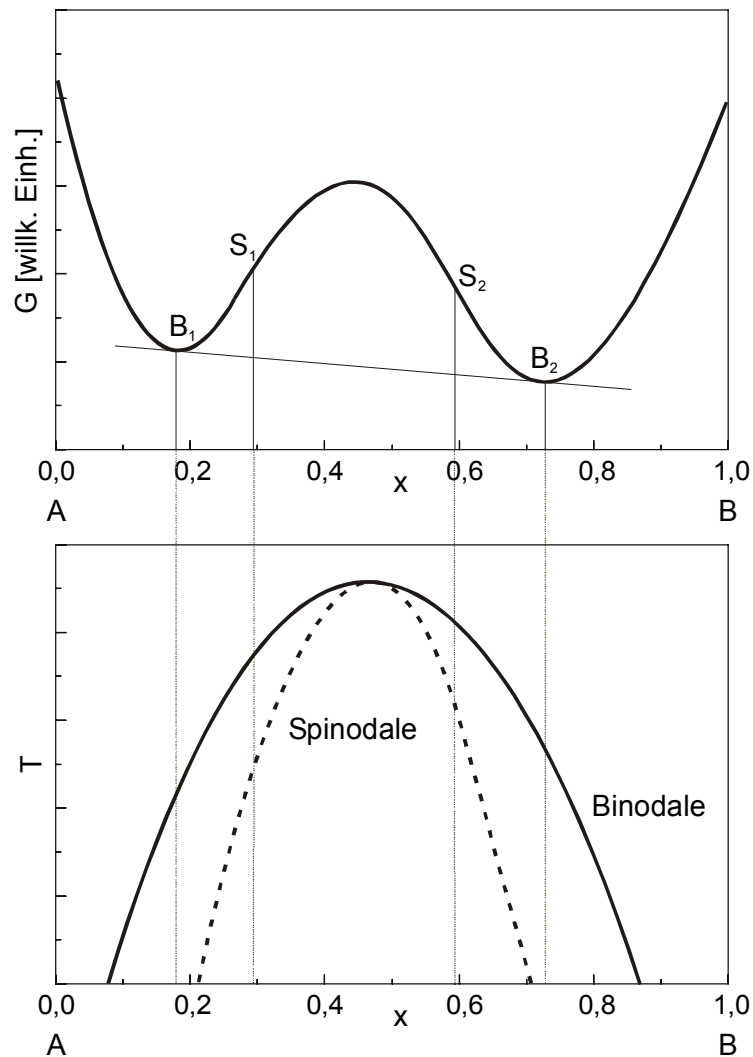


Bild 2.15 : Zur Definition von spinodaler und binodaler Kurve

2.2.1.2.2 Spinodale und Binodale im quaternären System

Für ein quaternäres Materialsystem ist die freie Gibbs'sche Energie eine Funktion von x und y . Daraus resultiert nunmehr eine Fläche, welche streckenweise eine negative Krümmung besitzt (siehe Bild 2.16). Analog zum binären Fall ergibt sich der Zusammenhang zwischen spinodaler und binodaler Kurve und der Zusammensetzung und es ergeben sich wiederum Wendepunkte und Tangentialpunkte. Die spinodale Kurve ist nun die Verbindungslinie

zwischen den Wendepunkten, die den Übergang zwischen einer rein konkaven und einer konkav-konvexen Fläche darstellen. Daraus ergibt sich für die Spinodale:

$$\left(\frac{\partial^2 G}{\partial x^2} \right) \left(\frac{\partial^2 G}{\partial y^2} \right) - \left(\frac{\partial^2 G}{\partial x \partial y} \right)^2 = 0 \quad (2.44)$$

Die Binodale erhält man durch das Anlegen von Tangentialebenen von unten an die Fläche der freien Gibbs'schen Energie. Sie ist die Verbindungslinie der einzelnen Tangentialpunkte, welche die Grenze zwischen dem stabilen (A) und dem metastabilen Bereich (B) darstellt. Die Beziehung hierfür lautet:

$$\left. \frac{\partial G}{\partial x} \right|_1 = \left. \frac{\partial G}{\partial x} \right|_2 \quad (2.45a)$$

$$\left. \frac{\partial G}{\partial y} \right|_1 = \left. \frac{\partial G}{\partial y} \right|_2 \quad (2.45b)$$

$$-\left. \frac{\partial G}{\partial x} \right|_1 x_1 - \left. \frac{\partial G}{\partial y} \right|_1 y_1 + G(x_1, y_1) = -\left. \frac{\partial G}{\partial x} \right|_2 x_2 - \left. \frac{\partial G}{\partial y} \right|_2 y_2 + G(x_2, y_2) \quad (2.45c)$$

Eine homogene Phase wird instabil, wenn ihre Zusammensetzung innerhalb der spinodalen Kurve (dem instabilen Bereich) liegt und wird in zwei Phasen entmischen. Diesen Prozess nennt man auch spinodale Entmischung. Die Zusammensetzung der sich im Gleichgewicht befindenden zwei Phasen ergibt sich aus der Lage der binodalen Kurve.

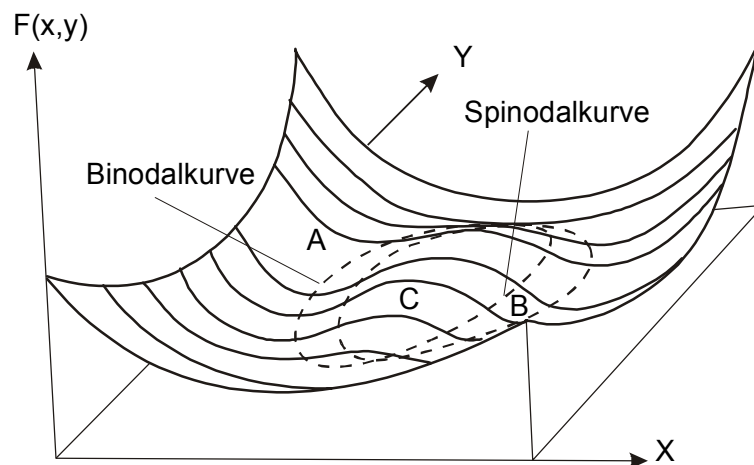


Bild 2.16 : Schematischer Verlauf der Fläche der freien Gibbs'schen Energie eines quaternären Materialsystems mit Mischungslücke [48]

Da eine Mischungslücke immer durch das Auftreten eines metastabilen und instabilen Bereiches gekennzeichnet ist, ist die Berechnung von Spinodalen sinnvoll, um Vorhersagen über die Existenz und sogar über die Ausdehnung einer Mischungslücke treffen zu können. Dies ist mit Hilfe der Gleichungen 2.45a, 2.45b und 2.45c möglich.

2.2.2 Die Mischungslücke des Materialsystems $\text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y}$

Alle im Folgenden gezeigten Diagramme von Mischungslücken wurden am Institut für Halbleitertechnik, Lehrstuhl I der RWTH Aachen berechnet und für diese Arbeit zur Verfügung gestellt. Aufbauend auf theoretischen Betrachtungen von Stringfellow [8], Onabe [49] und de Cremoux [10] wurde an diesem Institut ein Softwarepaket entwickelt, mit welchem die Ausdehnung der Mischungslücke in ternären und quaternären Materialsystemen berechnet werden kann. Zusätzlich zur Mischungslücke erlaubt das Programm die Berechnung der Spinodalen. Bild 2.17 zeigt beispielhaft für eine Temperatur von 640°C die berechneten Lagen der spinodalen und der binodalen Kurve für das quaternäre Materialsystem $\text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y}$. Außerdem sind die Geraden für gitterangepasstes Wachstum auf InP bzw. GaAs eingezeichnet.

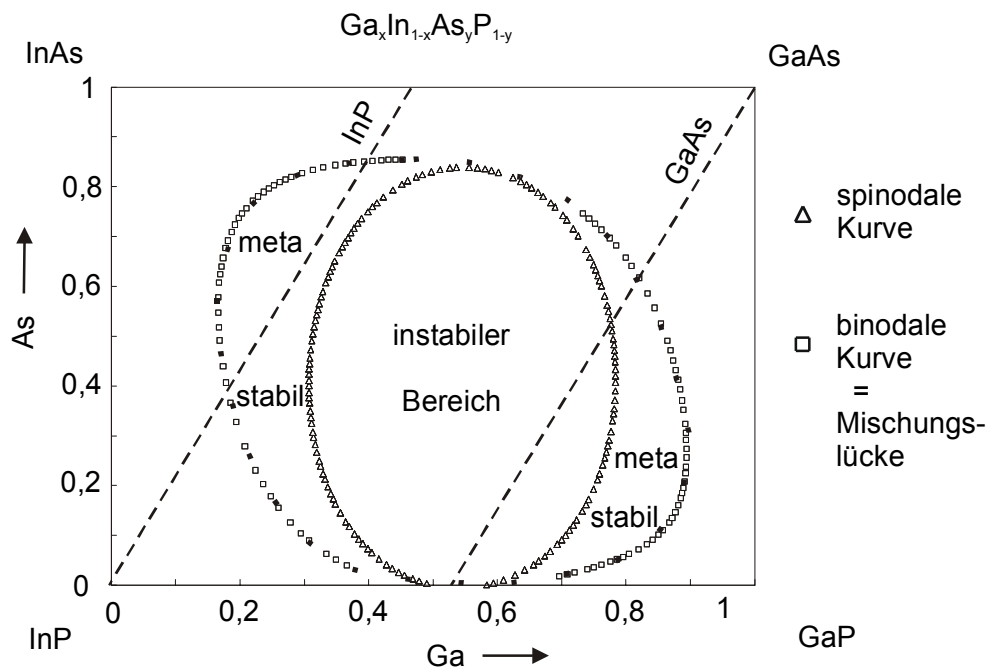


Bild 2.17: Spinodale und binodale Kurve im Materialsystem $\text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y}$ bei einer Temperatur von 640°C

Neben der Lage der Mischungslücke liefert das Programm zusätzlich die Lage der sogenannten Gleichgewichtsgeraden. Entlang dieser Geraden erfolgt der Zerfall der homogenen Mischkristallphase in zwei miteinander im thermodynamischen Gleichgewicht stehenden Phasen unterschiedlicher chemischer Zusammensetzung. Die Schnittpunkte der Gleichgewichtsgeraden mit der binodalen Kurve bestimmen jeweils die chemische Zusammensetzung der zwei Phasen. Bild 2.18 zeigt diese Gleichgewichtsgeraden für das gleiche Materialsystem und dieselbe Temperatur wie in Bild 2.17.

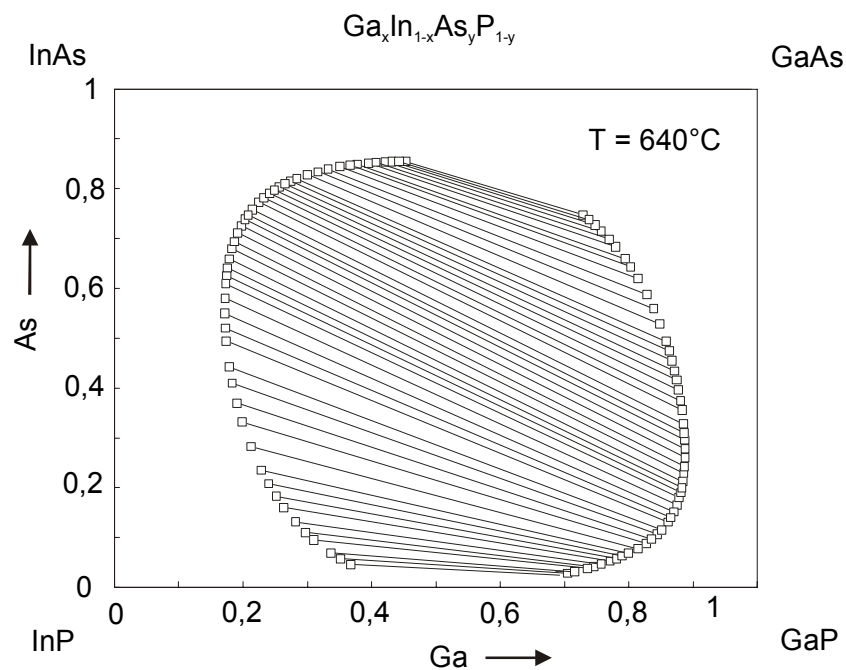


Bild 2.18: Lage der Gleichgewichtsgeraden für das Materialsystem $\text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y}$ bei einer Temperatur von 640°C .

Die Mischungslücke wird mit steigender Temperatur kleiner und verschwindet oberhalb der kritischen Temperatur vollständig. Die Mischungslücke ist in Bild 2.19 für verschiedene Temperaturen dargestellt.

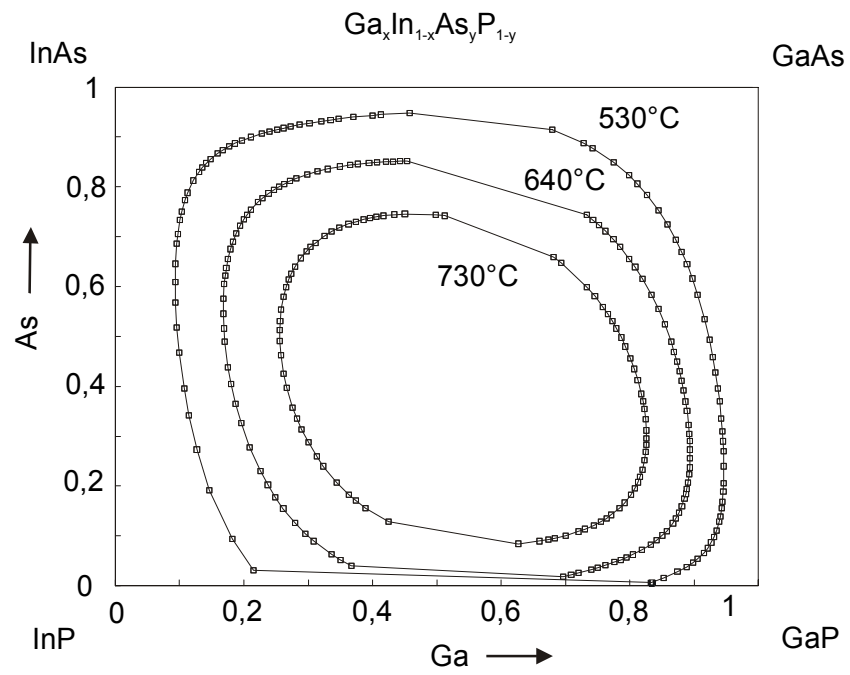


Bild 2.19 : Temperaturabhängigkeit der Mischungslücke im Materialsystem $\text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y}$.

2.3 *Eigenschaften und Herstellung des Materialsystems* *$Ga_xIn_{1-x}As_yP_{1-y}$*

Quaternäre Verbindungen aus den Hauptgruppen III und V des Periodensystems finden aufgrund ihrer speziellen physikalischen Eigenschaften im Bereich der Mikrowellentechnik und der Optoelektronik ihren Einsatz. Insbesondere das Interesse am Materialsystem $Ga_xIn_{1-x}As_yP_{1-y}$ hat in der vergangenen Zeit enorm zugenommen, da es sich theoretisch sowohl auf dem Substratmaterial GaAs als auch auf InP über weite Kompositionsbereiche gitterangepasst, d.h. mit der gleichen Gitterkonstante wie das Substrat aufwachsen lässt. Somit ist es möglich, den Bandabstand E_g des Halbleiters über die Zusammensetzung gezielt einzustellen und unabhängig von E_g die Gitterkonstante zu wählen. Man spricht bei der Herstellung von solchen maßgeschneiderten Bauelementeigenschaften auch vom sogenannten „Band Gap Engineering“.

Das Materialsystem $Ga_xIn_{1-x}As_yP_{1-y}$ setzt sich aus den binären Verbindungshalbleitern GaAs, GaP, InAs und InP zusammen. Es gehört der Gruppe der III-V-Halbleiter an und bildet unter üblichen Epitaxiebedingungen Kristalle mit einer Zinkblendestruktur. Das Zinkblendegitter besteht aus zwei kubisch-flächenzentrierten (kfz) Untergittern, welche um ein Viertel der Raumdiagonalen gegeneinander verschoben sind. Die Atome der III. Hauptgruppe (Ga, In) verteilen sich dabei statistisch auf das eine Untergitter mit der Gitterkonstanten a_0 . Die Atome der V. Hauptgruppe (As, P) besetzen die Gitterplätze des anderen Untergitters, ebenfalls mit der Gitterkonstanten a_0 (siehe Bild 2.20).

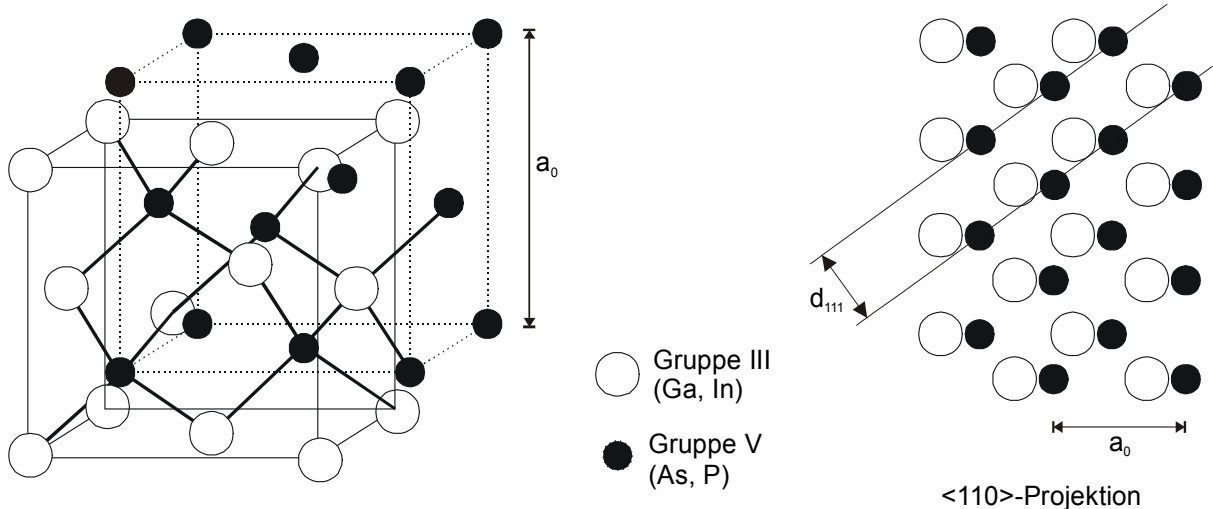


Bild 2.20 : Zinkblendestruktur des Materialsystems $Ga_xIn_{1-x}As_yP_{1-y}$ und ihre Ansicht in $\langle 110 \rangle$ -Projektion

Es kann in Einzelfällen auch zu einer Bildung geordneter Phasen in den Untergittern kommen, d.h. auf den Untergittern ordnen sich Elemente der III. bzw. V. Hauptgruppe in bestimmter Form regelmäßig an. Dieser Effekt wird als „Ordering“ bezeichnet und beeinflusst die optischen Eigenschaften des Halbleiters.

Die Gitterkonstante a_0 ist für alle drei Raumrichtungen gleich ($a_0 = b_0 = c_0$). Jedes Atom bildet mit den vier nächsten Nachbaratomen, die dem anderen Untergitter angehören, einen Tetraeder. Der kürzeste interatomare Abstand beträgt $a_0 \sqrt{3}/4$ [50].

Der für die Epitaxie wichtigster Parameter ist die Gitterkonstante der am Wachstum beteiligten Materialien, da sie primär für die kristalline Qualität des Halbleiters verantwortlich ist. Die Gitterkonstante a_0 des Mischkristalls $\text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y}$ ändert sich linear mit der Zusammensetzung x, y . Nach dem Vegard'schen Gesetz ergibt sich für a_0 :

$$a_0(x, y) = xy a_{\text{GaAs}} + x(1-y) a_{\text{GaP}} + y(1-x) a_{\text{InAs}} + (1-x)(1-y) a_{\text{InP}} \quad (2.46)$$

Hierbei sind a_{GaAs} , a_{GaP} , a_{InAs} und a_{InP} die Gitterkonstanten der jeweiligen binären Verbindungshalbleiter. Die Änderung des Bandabstandes in Abhängigkeit von der Zusammensetzung hingegen ist nichtlinear. Die Nichtlinearität findet im sogenannten „Bowingparameter“ b Berücksichtigung. Damit kann der Bandabstand E_g wie folgt errechnet werden [51]:

$$\begin{aligned} E_g(x, y) = & xy E_{g,\text{GaAs}} + x(1-y) E_{g,\text{GaP}} + y(1-x) E_{g,\text{InAs}} + (1-x)(1-y) E_{g,\text{InP}} \\ & - x(1-x)y \cdot b_{\text{GaInAs}} - x(1-x)(1-y) \cdot b_{\text{GaInP}} \\ & - x(1-y)y \cdot b_{\text{GaAsP}} - y(1-x)(1-y) \cdot b_{\text{InAsP}} \end{aligned} \quad (2.47)$$

Bild 2.21 zeigt die Darstellung der Gitterkonstante a_0 für $\text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y}$ über den Bandabstand, respektive der Emissionswellenlänge. Es zeigt, dass das Materialsystem bei nahezu jeder Zusammensetzung einen direkten Bandabstand liefert.

Auf GaAs-Substrat lassen sich $\text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y}$ -Schichten mit Emissionswellenlängen von ca. 620 nm (orange) bis 880 nm (rot/infrarot) herstellen. Das Interesse liegt hier im Einsatz von optoelektronischen Bauelementen wie Laser, Laserdioden oder auch HF-Bauelementen als Alternative zum AlGaAs-Materialsystem [52]. Bei auf InP-Substrat abgeschiedenen Schichten liegen die Emissionswellenlängen im tiefen Infrarotbereich (920 nm bis 1,65 μm). Dieser Wellenlängenbereich ist in der optischen Nachrichtentechnik von besonderer Bedeutung, da bei 1,3 μm und bei 1,55 μm die Dämpfung einer Glasfaser minimal ist [53].

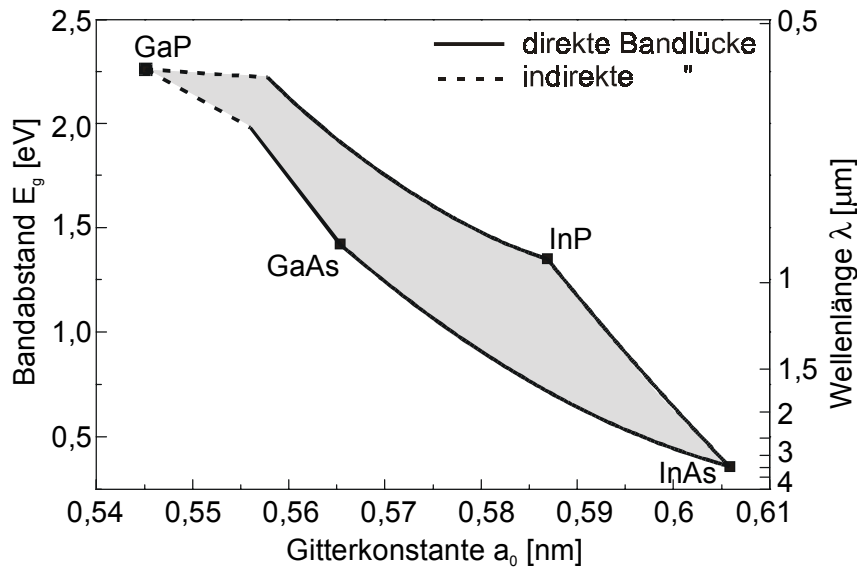
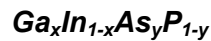


Bild 2.21 : Bandabstand und Gitterkonstante für das quaternären Materialsystem



Unter bestimmten Voraussetzungen ist es möglich, Schichten mit unterschiedlichen Gitterkonstanten aufeinander aufzuwachsen. Werden nur sehr dünne Schichten eines Materials mit kleinerer bzw. größerer Gitterkonstante auf ein kristallines Trägermaterial aufgebracht, wird der aufgewachsenen Schicht die Gitterkonstante des Substrats aufgezwungen und es entsteht innerhalb der Schicht eine mechanische Verspannung (vgl. Bild 2.22b). In Wachstumsrichtung ergibt sich eine Änderung der Gitterkonstanten bezogen auf die Gitterkonstante des Substrats, in lateraler Richtung bleibt die Gitterkonstante des Substratmaterials erhalten. Bis zu einer bestimmten nicht zu großen Schichtdicke, der kritischen Schichtdicke, ist ein fehlerfreies Kristallwachstum möglich. Man nennt derartig verspannte Schichten auch pseudomorphe Schichten [54]. Durch das Wachstum von verspannten Schichten hat man einen direkten Einfluss auf die Bandstruktur. Daher lassen sich gezielt elektrische und optische Eigenschaften eines Bauelementes beeinflussen. Mit zunehmender Schichtdicke und Gitterfehlانpassung nimmt die mechanische Spannung auf die verspannten Schichten zu bis sie schließlich relaxieren, wodurch Versetzungen entstehen (siehe Bild 2.22c). Solche Versetzungen zerstören lokal die kristalline Struktur und stellen Streuzentren für die Ladungsträger dar und verschlechtern optische und elektrische Eigenschaften des Materials [55].

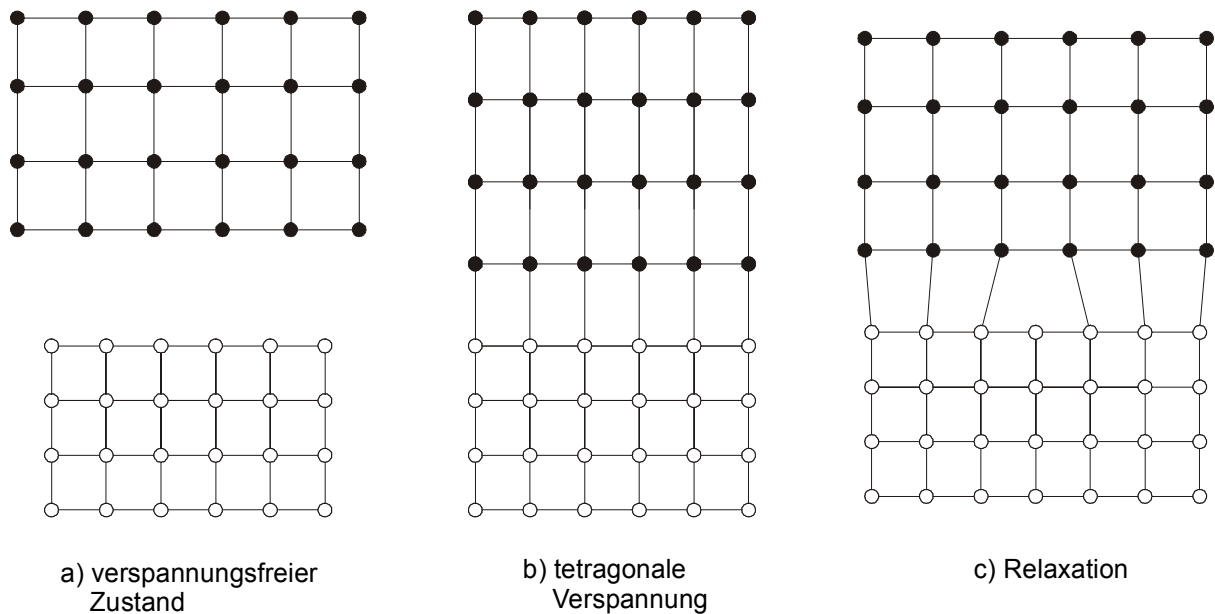


Bild 2.22 : Pseudomorphes Wachstum und Relaxation verspannter Schichten

Wie bereits erwähnt, stehen für die Epitaxie mit InP und GaAs zwei Substratmaterialien zur Verfügung, auf denen gitterangepasstes Wachstum möglich ist. Am häufigsten werden die Schichten auf den {100}-Substratflächen abgeschieden. Dabei ist für die Abscheidung aus der Gasphase eine Fehlorientierung von 2° bis etwa 5° in Richtung [110] erforderlich, um ausreichende Wachstumsgeschwindigkeiten und zufriedenstellende Oberflächenperfektion der Schichten zu erreichen [50]. Die wichtigsten Epitaxieverfahren zur Herstellung von Heteroschichtstrukturen mit hoher kristalliner Güte und Schichtdicken im Nanometerbereich sind die Molekularstrahl-Epitaxie (engl.: Molecular Beam Epitaxy, MBE) und die metallorganische Gasphasen-Epitaxie (engl.: Metall Organic Vapour Phase Epitaxy, MOVPE, auch Metal Organic Chemical Vapour Deposition, MOCVD). Alle in dieser Arbeit untersuchten Schichtstrukturen wurden mittels MOVPE hergestellt, so dass an dieser Stelle eine kurze Beschreibung dieses Herstellungsverfahrens erfolgt.

Das Kristallwachstum bei der MOVPE resultiert aus der thermischen Zerlegung von gasförmigen Ausgangsverbindungen (Precursor) und deren Abscheidung auf einem Substrat. Die Ausgangsstoffe für das Wachstum von quaternären III-V-Halbleitern sind metallorganische Verbindungen mit Elementen der dritten Hauptgruppe (z.B.: TMI: Trimethylindium, TMG: Trimethylgallium) und Hydride aus Elementen der fünften Hauptgruppe (AsH_3 : Arsin, PH_3 : Phosphin). Durch ein Trägergas (meist: H_2 : Wasserstoff) werden die Gruppe-III-Quellen (die in flüssiger oder fester Form vorliegen können) dem Reaktor, in dem sich das Substrat befindet, zugeführt. Die Gruppe-V-Quellen sind als

Hydride bereits gasförmig. Der Substrathalter (Suszeptor) wird induktiv über eine Hochfrequenzspule erwärmt. Im heißen Reaktorraum kommt es zur homogenen Gasphasereaktion, die zur Bildung der eigentlich am Wachstumsprozess beteiligten Ausgangsverbindungen führt. Diese werden zur Substratoberfläche transportiert (durch Diffusion) und werden dort adsorbiert, vollständig zerlegt und die Atome diffundieren entlang der Substratoberfläche zur Einbauposition (zweidimensionales Wachstum). Die während des Wachstums entstandenen Reaktionsprodukte werden desorbiert und abtransportiert. Die resultierende Wachstumsgeschwindigkeit wird dabei im Wesentlichen durch den langsamsten Prozess bestimmt [54]. Die wichtigsten Epitaxieparameter für ein kontrolliertes Wachstum sind

- Totaldruck im Reaktor (niedrige Drücke ermöglichen hohe Gasflussgeschwindigkeiten)
- V/III-Verhältnis (beeinflusst die Einbauwahrscheinlichkeit von Atomen an der Wachstumsoberfläche)
- Wachstumsrate
- Mischungsverhältnis
- Dotierung

Bild 2.23 zeigt die Prinzipdarstellung einer MOVPE-Anlage zur Herstellung von Halbleiter-Heterostrukturen. Detaillierte Angaben zur Theorie des MOVPE-Wachstumsprozesses sind in [56] zu finden.

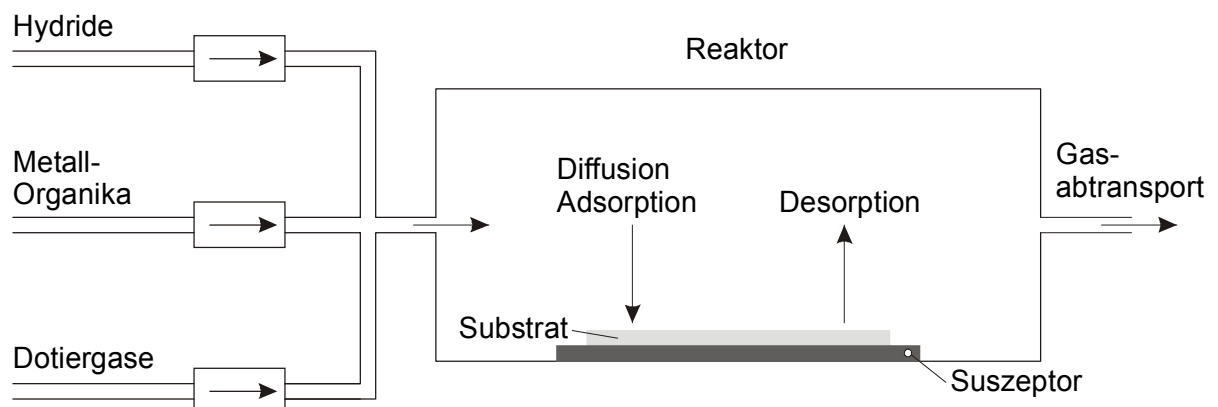


Bild 2.23 : Prinzipdarstellung einer MOVPE-Anlage