Kapitel 3

Die verwendeten Meßverfahren

In diesem Kapitel sollen einige grundlegende Eigenschaften der verwendeten Meßmethoden LEED, RHEED, SXRD und STM aufgeführt werden, um die Interpretation der in dieser Arbeit vorgestellten experimentellen Ergebnisse zu erleichtern. Ausführliche Darstellungen der Untersuchungsmethoden finden sich in eine Vielzahl ausgezeichneter Lehrbücher, Monographien und Übersichtsartikeln. Beispielhaft sollen hier auf die Lehrbücher von Henzler und Göpel [18] und Woodruff und Delchar [19] genannt werden und speziell für die Beugung langsamer Elektronen auf [20], für die Beugung schneller Elektronen auf [21], für die Oberflächenröntgenbeugung auf [22, 23] und für die Rastertunnelmikroskopie auf [24, 25, 26, 27] verwiesen werden.

3.1 Der reziproke Raum

Die Meßmethoden LEED, RHEED und SXRD sind beugende Verfahren. Die erzeugten Beugungsbilder sind Darstellungen des reziproken Gitters des Kristalls bzw. Abbildungen des reziproken Raums (Fourier-Raum, K-Raum). Der reziproke Raum eines idealen, unendlich ausgedehnten dreidimensionalen Kristalls setzt sich aus einer periodischen Anordnung von Bragg-Punkten zusammen. Die gebrochene Periodizität des Volumengitters an der Oberfläche des Kristalls bewirkt eine Verschmierung der Bragg-Reflexe des Volumens in Richtung der Oberflächennormalen zu Grundgitterstäben (*crystal truncation rods, CTR*). Eine periodische Anordnung der Atome der Oberfläche wird wegen der fehlenden Periodizität in der Normalenrichtung im reziproken Raum durch Bragg-Stäbe (*Bragg rods*) wiedergegeben. Sind die Periodizitätslängen der Überstruktur im Ortsraum größer als die des Volumens, rufen sie neben Stäben ganzzahliger Ordnung, deren Position mit der Position der CTR zusammenfällt, auch Stäbe gebrochenzahliger Ordnung hervor (*fractional order rods*). Eine schematische Darstellung des reziproken Raums eines Kristalls mit rekonstruierter Oberfläche gibt Abb. 3.1.

Die möglichen Richtungen der gebeugten Strahlen sind im Rahmen der kinematischen Näherung durch die Schnittpunkte der Ewald-Kugel mit den Stäben bzw. Punkten des reziproken Gitters gegeben (Abb. 3.2).

3.2 LEED

Die Bestimmung des reziproken Gitters aus dem erhaltenen Beugungsbild ist für die Beugung langsamer Elektronen (*low-energy electron diffraction*, LEED) besonders einfach, denn die LEED-Geometrie sieht im allgemeinen einen annähernd senkrechten Einfall des Elektro-

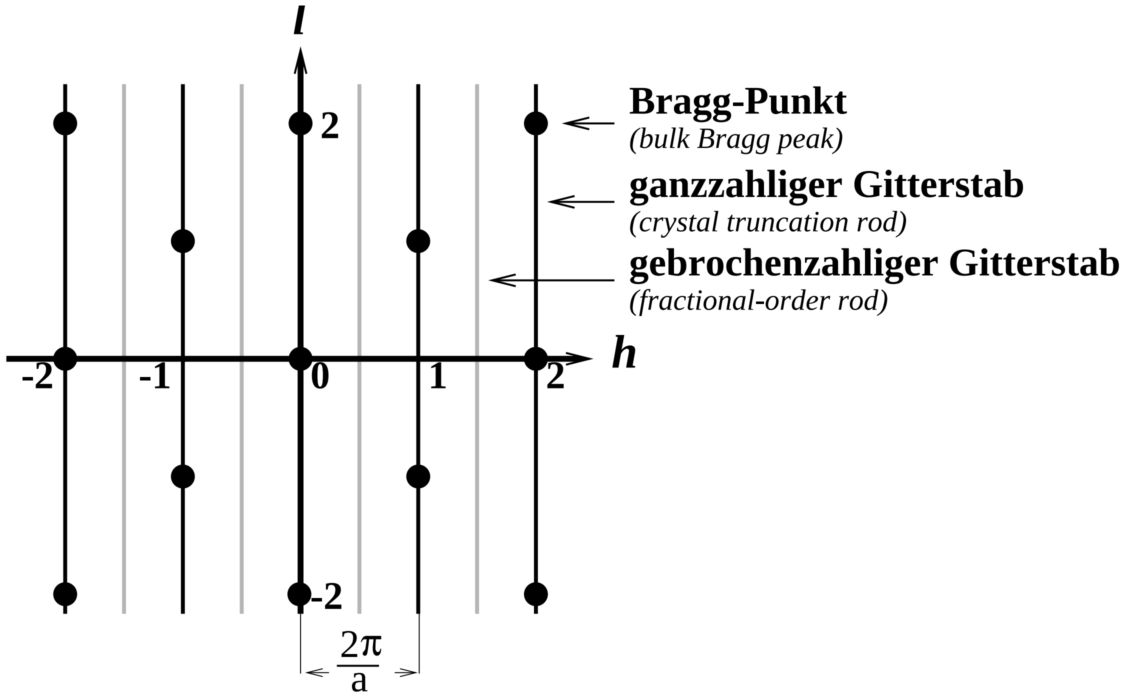

Abbildung 3.1: Schematische Darstellung des reziproken Raums eines semi-unendlichen Kristalls mit rekonstruierter Oberfläche.

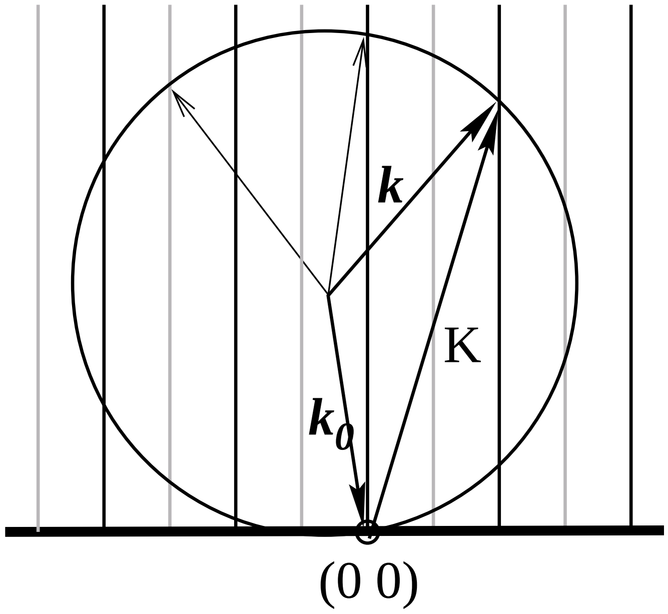

Abbildung 3.2: Ewald-Konstruktion für eine ideale Oberfläche in LEED-Geometrie. Der Vektor $\vec{k_0}$ bedeutet den Wellenvektor des einfallenden Strahls, $\vec{k}$ den Wellenvektor eines gebeugten Strahls und $\vec{K}$ den entsprechenden Streuvektor (reziproker Gittervektor).

nenstrahls auf die Probenoberfläche vor. Die Wölbung des Leuchtschirms kompensiert die Verzerrung der Abbildung durch die Krümmung der Ewald-Kugel, und das Beugungsbild gibt die Symmetrie des reziproken Gitters in Normalenrichtung maßgetreu wieder. Periodische Abstände wie Gitterkonstanten und Einheitszellen von Überstrukturen können einfach bestimmt werden. Das LEED-Experiment ist durch die Wahl niedriger Energien für den einfallenden monochromatischen Elektronenstrahl (10–500 eV) sehr oberflächenempfindlich. Die mittlere freie Weglänge und damit die Eindringtiefe der Elektronen beträgt nur wenige Atomlagen (∼5 Å). Gleichzeitig liegt die de Broglie-Wellenlänge der Elektronen im Bereich atomarer Abstände, was die Auflösung atomarer Strukturen unter großen Beugungswinkeln ermöglicht.

Als Kohärenzbreite oder Transferweite wird die instrumentenbedingte endliche Ausdehnung der Fläche bezeichnet, über die die Strahlen kohärent interferieren. Die endliche Kohärenzlänge bewirkt eine endliche minimale Halbwertsbreite der Beugungsreflexe. Für unser LEED-System wurde experimentell eine Kohärenzlänge von etwa 120 Å bestimmt. Periodische Abstände, die größer als die Transferweite sind, können nicht aufgelöst werden und Domänengrößen periodischer Strukturen unterhalb der Transferweite bedeuten eine Verbreiterung der Halbwertsbreite der Reflexe über die instrumentenbedingte Verbreiterung hinaus. Die Halbwertsbreite verhält sich umgekehrt proportional zur Ausdehnung der Domäne, so daß die mittlere Domänenbreite über bekannte Gitterabstände und zwei Längenmessungen einfach zu bestimmen ist. Da der Durchmesser des Elektronenstrahls mit 1 mm um viele Größenordnungen größer ist als die kohärent beleuchteten Bereiche, stellt das Beugungsbild eine inkohärente Überlagerung von Teilbildern vieler kohärenter Bereiche dar.

Die Eigenschaften des reziproken Raum in Normalenrichtung zur Oberfläche können durch die Variation der Energie der einfallenden Elektronen bzw. die Variation des Radiuses der Ewald-Kugel untersucht werden. Der Verlauf der Intensitäten der Reflexe als Funktion der Energie des Elektronenstrahl (I/V-Charakteristik) enthält prinzipiell die Information über die Position der Atome innerhalb der Einheitszelle. Die Auswertung der I/V-Kurven beinhaltet die Anpassung von Modellrechnungen an die experimentellen Daten. Sie ist numerisch sehr aufwendig und wurde in dieser Arbeit nicht durchgeführt.

Die Unterscheidung zwischen ebenen (1×1)-rekonstruierten Oberflächen und rauhen, ungeordneten Oberflächen erlaubt bereits eine grobe qualitative Beobachtung des Intensitätsverlauf mit dem Auge. Rauhe, ungeordnete Oberflächen zeichnen sich lediglich durch Crystal Truncation Rods aus, während ebene, (1×1)-rekonstruierte Oberflächen darüber hinaus ganzzahlige Gitterstäbe mit einem für die jeweilige Struktur charakteristischen modulierten Verlauf verursachen.

Eine Facettierung der Oberfläche ist durch die Variation der Elektronenenergie ebenfalls leicht nachzuweisen. Während die Beugungsreflexe einer ebenen Oberfläche bei Erhöhung der Energie kollektiv zentral auf den (0,0)-Reflex (Spekularreflex) zulaufen und das Beugungsbild lediglich "schrumpft", wird für Oberflächen, die aus mehreren Facetten bestehen, ein Gegeneinanderlaufen der Reflexe beobachtet. Tatsächlich bewegen sich alle Reflexe auf ihren jeweiligen Spekularreflex zu. Die Spekularreflexe bleiben während der Variation der Elektronenenergie auf dem Schirm ortsfest.

Die Abbildung 3.3 zeigt exemplarisch die Ewald-Konstruktion für eine {103}-facettierte (001)-Oberfläche. Die reziproken Gitterstäbe stehen jeweils senkrecht auf den zugehörigen Facettenflächen. Gehören die Facettenflächen zu einer gemeinsamen einkristallinen Unterlage, werden sich alle Stangensätze gemeinsam in Punkten des reziproken Raumes kreuzen.

Die Orientierung der Facetten läßt sich auf mehrere Arten bestimmen. Sofern es der Aufbau des Probenmanipulator erlaubt, genügt die Drehung der Probe aus einer definierten

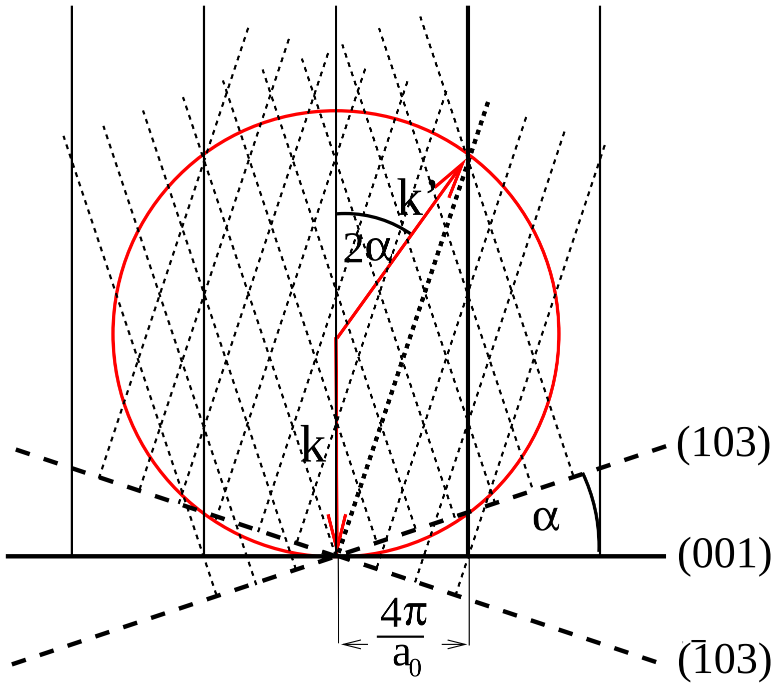

Abbildung 3.3: Ewald-Konstruktion einer facettierten Oberfläche am Beispiel von {103}-Facetten (gestrichelte Linien) auf einer (001)-Oberfläche (durchgezogene Linien). Mit Hilfe der Koinzidenz des Spekularreflexes der ($\bar{1}03$)-Facette mit dem (11)-Reflex der (001)-Oberfläche läßt sich der Neigungswinkel α der Facette berechnen.

Ausgangsstellung mit senkrechtem Einfall des Strahls auf eine bekannte Facette bis in die Position, in der sich der Spekularreflex der unbekannten Facette in der Mitte des Schirms befindet. Der Drehwinkel entspricht direkt dem Neigungswinkel der Facette zur Flächennormalen der Ausgangsstellung. Bei festgehaltener Probenposition kann der Neigungswinkel der Facette auch durch die Variation der Energie über die Koinzidenz des Spekularreflexes der Facette mit einem bekannten Reflex einer anderen Facette bestimmt werden. In Abb. 3.3 ist die Koinzidenz des Spekularreflexes einer {103}-Facetten mit der (1,1)-Reflex der (001)-Ebene dargestellt. Der Neigungswinkel α der ($\bar{1}03$)-Ebene zur (001)-Oberfläche berechnet sich aus

$$sin 2\alpha = \frac{\frac{4\pi}{a_0}}{\sqrt{2Em_e}\hbar^{-1}},$$

wobei m_e die Masse des Elektrons, $\hbar$ das Planck'sche Wirkungsquantum E die Energie des Elektronenstrahls und a_0 die Volumengitterkonstante des Substrats bedeuten. Das allgemeine Verfahren zur Bestimmung der Facettenorientierung besteht in der Auftragung der Reflexpositionen in den reziproken Raum für mehrere Energien. Durch das Verbinden der Punkte werden die Reflexe Gitterstäben zugeordnet und die Facetten identifiziert. Auf dieses Verfahren wird in Abschn. 10.6 an einem Beispiel näher eingegangen werden.

3.3 RHEED

Bei der Beugung schneller Elektronen mit Energien im Bereich 10–50 keV (*reflection high-energy electron diffraction*, RHEED) wird die nötige Oberflächenempfindlichkeit durch den

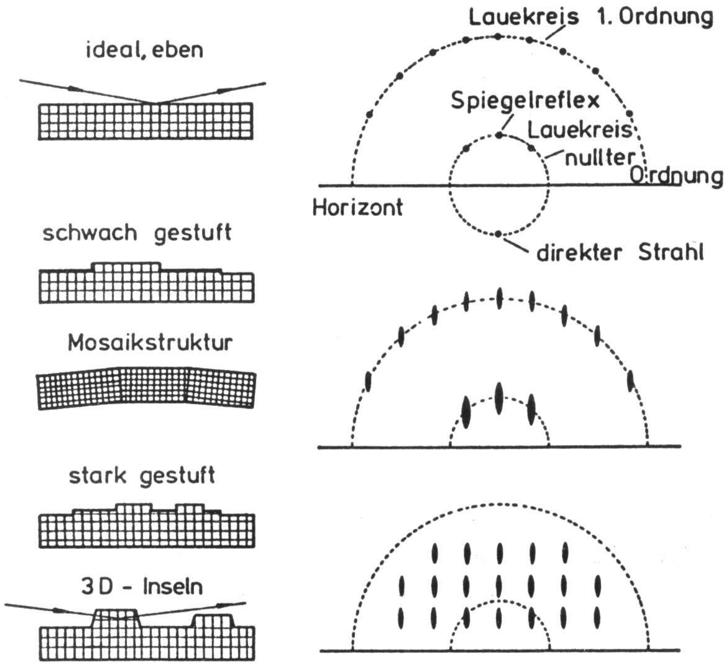

Abbildung 3.4: Beispiele für Oberflächenstrukturen und entsprechende RHEED-Bilder (aus [18]).

streifenden Einfall des Elektronenstrahls erreicht ($\Theta < 5°$). Die Eindringtiefe der Elektronen senkrecht zur Oberfläche ist trotz einer mittleren freien Weglänge der Elektronen von 100-1000 Å vergleichbar mit der typischen Eindringtiefe bei LEED-Experimenten. Der Geometrie des streifenden Einfalls bewirkt eine Verzerrung der Abbildung des reziproken Raum auf dem Leuchtschirm. In der Ewald-Konstruktion durchstoßen die Gitterstäbe die Ewald-Kugel unter kleinem Winkel, so daß Stäbe, die im reziproken Raum auf Geraden liegen, auf dem Bildschirm auf Kreisbögen, den sogenannten Laue-Ringen, abgebildet werden. Die Kohärenzlänge beträgt für ein konventionelles Gerät wie unseres etwa 100 Å senkrecht zur Strahlebene und etwa 1000 Å in der Strahlebene. Domänengrößen unterhalb der Kohärenzlänge weiten die ideal punktförmigen Reflexe durch den kleinen Einfallswinkel zu länglichen bis streifigen Reflexen (*streaks*) auf. Durch Variation des Einfallswinkels (*rocking curve*) können analog zur Variation der Elektronenenergie in der LEED-Geometrie Informationen über die Eigenschaften des reziproken Raums in Normalenrichtung zur Oberfläche gewonnen werden. Der streifende Einfall und die hohe Energie des Elektronenstrahls ermöglichen die Durchstrahlung von stark in Normalenrichtung hervorstehenden Strukturen und machen RHEED sensitiv für morphologische Eigenschaften wie Rauhigkeit und das Auftreten von 3D-Inseln. Ein fundamentaler Vorteil der RHEED-Geometrie besteht darin, daß der Halbraum vor der Probenoberfläche während der Messung für weitere Präparations- und Analysetechniken frei zugänglich bleibt und z. B. Wachstumsvorgänge während der Deposition beobachtet und gesteuert werden können.

3.4 SXRD

Die in dieser Arbeit aufgeführten Untersuchungen mit Oberflächenröntgenbeugung (*surface x-ray diffraction, SXRD*) wurden überwiegend in Zusammenarbeit mit O. Bunk sowie Arbeitsgruppen aus Risø und Delft durchgeführt. Die übrigen behandelten SXRD-Untersuchungen wurden von der Arbeitsgruppe von Prof. Moritz aus München durchgeführt und werden hier nur zitiert. Da die Auswertung der Daten nicht selbst vorgenommen wurde, sollen die experimentellen Grundlagen dieses Meßverfahrens nur kurz angerissen werden.

Die Oberflächenröntgenbeugung ist eine höchst sensitive Methode zur Bestimmung der atomaren Struktur von Oberflächen. Die Wechselwirkung von Röntgenstrahlen mit Materie ist vergleichsweise gering. Dadurch kann die Mehrfachstreuung in der Regel vernachlässigt werden und zur Beschreibung der elastischen Röntgenstreuung ist die numerisch weniger aufwendige kinematische Theorie ausreichend. Darüber hinaus ermöglicht die geringe Wechselwirkung die Untersuchung vergrabener Grenzflächen.

Bei der Auswertung der Daten besteht wie für alle beugenden Verfahren prinzipiell das Problem, daß mit der Messung der Intensität der Beugungsreflexe nur das Betragsquadrat der Amplitude der gebeugten Strahlung, nicht aber die Phase bestimmt werden kann. Eine einfache Fouriertransformation vom Implusraum zurück in den Ortsraum kann somit nicht ohne weiteres durchgeführt werden. Eine Möglichkeit, auch ohne Kenntnis der Phase Informationen über die atomare Struktur zu gewinnen, bietet die Pattersonanalyse. Die Pattersonfunktion ist definiert als die zweidimensionale Autokorrelationsfunktion der Ladungsdichte $\rho(\vec{r})$

$$P(\vec{r}) = V \int \rho(\vec{r'})\rho(\vec{r} - \vec{r'})d\vec{r'}, \tag{3.1}$$

dabei wird über das Volumen V der Einheitszelle integriert.

Wird die Ladungsdichte ρ als Fouriertransformierte des Strukturfaktors F dargestellt und die Messung im streifenden Einfall durchgeführt, so daß der Impulsübertrag in Normalenrichtung zur Oberfläche verschwindet (*in plane*-Geometrie), geht Gleichung 3.1 über in

$$P(x,y) = \frac{1}{A}\sum_{h,k} |F_{hk}|^2 \cos[2\pi(hx + ky)], \tag{3.2}$$

wobei A die Fläche der Einheitszelle bedeutet.

Da sich die Betragsquadrate der Strukturfaktoren direkt aus den gemessenen Intensitäten I_{hk} berechnen, läßt sich die Pattersonfunktion ohne Modellannahmen direkt aus den experimentellen Daten gewinnen. Jedes Maximum der Pattersonfunktion entspricht einem auf die Oberfläche projizierten interatomaren Vektor in der Ladungsdichteverteilung der Rekonstruktion. Die Höhe der Maxima ist mit dem Produkt der Kernladungszahlen der zugehörigen Atome und der Anzahl gleicher Abstandsvektoren innerhalb der Einheitszelle gewichtet.

Ein Problem für die Pattersonanalyse stellen die Beiträge der ganzzahligen Reflexe zur Pattersonfunktion dar, denn sie enthalten neben Streubeiträgen durch die Rekonstruktion der Oberfläche auch Beiträge der Grundgitterstäbe. Deshalb werden häufig nur die gebrochenzahligen Reflexe zur Pattersonanalyse herangezogen. Die Unvollständigkeit des Datensatzes ruft negative Maxima in der (Differenz-)Dichteverteilung und der Pattersonfunktion hervor, die die Interpretation der Pattersonfunktion erschweren.

Die weitere Strukturanalyse und Strukturverfeinerung beruht auf einem Vergleich zwischen den beobachteten und den für ein Strukturmodell berechneten Beiträgen der Strukturfaktoren. Als Maß für die Güte der Anpassung des Modells an die Meßdaten werden verschiedene Größen genutzt (χ^2, gewichteter und ungewichteter R-Wert, Goodness of Fit). Für die

Strukturanalyse wird auf der Grundlage von Vorüberlegungen oder der Pattersonfunktion ein Strukturmodell geraten. Die Parameter des Modells werden anschließend mittels eines Minimierungsprogramms an den Datensatz angepaßt. Zur Bestimmung der Atompositionen senkrecht zur Oberfläche ist es notwendig, für die Anpassung neben den Messungen der *in plane*-Reflexe auch Messungen des Intensitätsverlauf entlang der reziproken Gitterstangen der Überstruktur (*rod scans*) heranzuziehen.

3.5 STM

Für eine umfassende Charakterisierung der Struktur von Oberflächen sind neben dem Einsatz beugender Verfahren auch mikroskopische Verfahren notwendig. Die Abbildung der Oberfläche im Ortsraum erleichtert insbesondere die Untersuchung nicht-periodischer Eigenschaften und ermöglicht ein genaues Bild über die Morphologie, den Grad der Unordnung und die Natur von Defektstrukturen. Die klassischen Methoden der Mikroskopie, nämlich die optische Mikroskopie und die Rasterelektronenmikroskopie (*scanning electron microscopy*, SEM) sind in ihrer Ortsauflösung jedoch beschränkt. Atomare Auflösung im Ortsraum war bis in die 80'er Jahre speziellen Materialklassen und Probengeometrien vorbehalten (Nadeln für die Feldionenmikroskopie und dünne Schichten für die Transmissionselektronenmikroskopie (TEM)) . Die Entwicklung der Rastertunnelmikroskopie (*scanning tunneling microscopy*, STM) durch Binnig und Rohrer (1982) bedeutete eine Revolution für die mikroskopische Untersuchung von Oberflächen, denn sie ist auf allen leitende Oberflächen (Metalle, Halbleiter, Supraleiter) und in verschiedenen Umgebungen (Vakuum, Gase, Flüssigkeiten) einsetzbar und leistet höchste Auflösung bis hinunter zu atomaren Abständen.

3.5.1 Das Meßprinzip

Der Aufbau und das Operationsprinzip eines STM's ist sehr einfach. Den Kern des STM bildet ein angespitzter Metalldraht (W, PtIr), der mit höchster Präzision in drei Raumrichtungen über die Probenoberfläche bewegt werden kann (siehe Abb. 3.5). Zwischen der Spitze und der Probe wird eine niedrige Spannung von typischerweise wenigen mV bis wenigen V angelegt. Anschließend wird die Probe auf einen Abstand von einigen wenigen Ångstrom an die Spitze angenähert, bis auf Grund des quantenmechanischen Tunneleffekts Elektronen von der Spitze durch die Lücke zur Probe "tunneln" (oder umgekehrt) und ein Tunnelstrom in der Größenordnung von 1 nA zu fließen beginnt. Der Tunnelstrom ist exponentiell abhängig vom Abstand zwischen Probe und Spitze [28]. Die Verringerung des Abstandes um 1 Å entspricht einer Erhöhung des Tunnelstroms um eine Größenordnung.

Die Bewegung der Spitze über die Probe ermöglicht in unserem STM-Aufbau ein orthogonales Dreibein von Piezokeramiken. Die laterale Rasterbewegung erfolgt durch das Anlegen passender Steuerspannungen an die x- und y-Piezokeramiken. Im Konstantstrommodus, der ausschließlich in dieser Arbeit Verwendung findet, wird die Auslenkung der z-Piezokeramik über eine Rückkopplungsschleife so eingestellt, daß der Tunnelstrom einen zuvor definierten Wert annimmt und die Spitze damit der Korrugation der Probenoberfläche folgt. Unter idealisierten Bedingungen erhält man aus dem Meßsignal $U_z(U_x, U_Y)$ mit bekanntem Konversionsfaktor der Piezokeramiken eine Fläche konstanten Tunnelstroms, welche wegen der empfindliche Abhängigkeit des Tunnelstroms vom Abstand zwischen Spitze und Probe (in grober Näherung) einem topographischen Abbild der Probenoberfläche entspricht. Wie im folgenden erläutert wird, liefert ein Konstantstrombild (*constant current topography*, CCT) auf atomarem Maßstab tatsächlich kein getreues Abbild der geometrischen Anordnung der

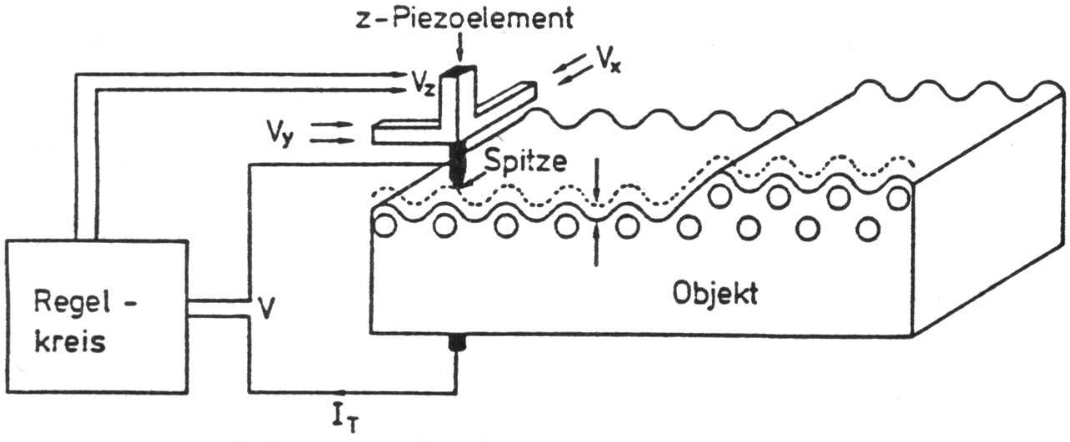

Abbildung 3.5: Schematischer Aufbau eines Rastertunnelmikroskops (aus [18]).

der Oberflächenatome, sondern ist darüber hinaus durch ihre elektronischen Eigenschaften bestimmt.

3.5.2 Theoretische Beschreibung

Obwohl das Funktionsprinzip eines STM's denkbar einfach ist, ist eine exakte theoretische Beschreibung des Tunnelstroms I($\vec{r}$,U) als Funktion des Ortes und der Tunnelspannung für ein reales Spitze-Probe-System nicht möglich. Ein großes Problem ist die Darstellung der Wellenfunktionen des Gesamtsystems, da auch bei einer perfekt periodischen Oberfläche das Gesamtsystem aperiodisch und die Symmetrie durch die Präsenz der Spitze verloren geht. Darüber hinaus ist die Anordnung der vordersten Atome der Spitze, über die der Tunnelstrom konzentriert ist, grundsätzlich unbestimmt und variiert unter Umständen während der Messung. Im wesentlichen sind zwei Klassen von Theorien zur Beschreibung des Tunnelprozesses zu unterscheiden, nämlich Störungsrechnungen erster Ordnung unter Verwendung des Transfer-Hamiltonoperators und Nicht-Störungerechnungen.

Die Rechnungen unter Verwendung des Transfer-Hamiltonoperators zeichnet sich durch ihre Einfachheit aus und sind von großer praktischer Relevanz, denn sie führen die beobachtete Korrugation in den STM-Bildern auf eine elementare Eigenschaft der Oberfläche zurück, nämlich der lokalen Zustandsdichte der Valenz- und Leitungselektronen. Der Transfer-Hamiltonoperator-Ansatz geht zurück auf die Beschreibung planaren Tunnelns von Bardeen [29]. In diesem Modell werden die Elektroden als zwei durch eine Potentialbarriere vollständig entkoppelte Systeme betrachtet. Wegen der schwachen Wechselwirkung der Elektroden untereinander kann der Tunnelstrom analog zur Störungsrechnung erster Ordnung geschrieben werden als

$$I = \frac{2\pi e}{\hbar} \sum_{\mu,\nu} [1 - f(E_\mu - eV)] f(E_\nu)\, |M_{\mu\nu}|^2 \delta(E_\mu - E_\nu). \tag{3.3}$$

Bezogen auf die Geometrie des STM gibt die Fermifunktion $f(E_\nu)$ die Verteilung der besetzten Zustände der Probe und $f(E_\mu - eV)$ die Verteilung der unbesetzten Zustände der Spitz an. V ist die Tunnelspannung, $M_{\mu\nu}$ das Matrixelement zwischen den ungestörten Wellenfunktionen ψ_μ und ψ_ν der beiden Elektroden und E_μ und E_ν sind die zugehörigen Energieeigenwerte. Die

δ-Funktion beschränkt das Tunneln auf elastische Übergänge, d. h. die Energie der Elektronen bleibt beim Tunnelprozeß erhalten.

Die Tunnelmatrixelemente $M_{\mu\nu}$ bzw. die Amplituden des Elektrontransfers sind durch den Überlapp der Wellenfunktionen der Spitze und der Probe gegeben und lassen sich als Stromdichteoperator ausdrücken:

$$M_{\mu\nu} = \frac{-\hbar^2}{2m} \int d\vec{S} \cdot (\psi_\mu^* \vec{\nabla} \psi_\nu - \psi_\nu \vec{\nabla} \psi_\mu^*), \tag{3.4}$$

dabei läuft die Integration über eine beliebige Fläche in der Vakuumregion zwischen Spitze und Probe. Zur Berechnung der Matrixelemente werden nun explizite Wellenfunktionen ψ_μ der Spitze und ψ_ν der Probe benötigt.

Tersoff und Hamann wählten in ihren berühmten Arbeiten [30, 31], die die erste mikroskopische Theorie für das STM darstellen, als einfaches Modell eine Spitze mit lokal sphärischer Symmetrie, so daß allein s-Wellenfunktionen zum Tunnelstrom beitragen. Unter den Nebenbedingungen, daß die Temperatur vernachlässigbar (kT klein gegen die Barrierenhöhe) und die Tunnelspannung klein ist (eV klein gegen kT), kommen sie zu folgendem Ausdruck für den Tunnelstrom:

$$I \propto \sum_\nu |\psi_\nu(\vec{r_0})|^2 \, \delta(E_\nu - E_F) \equiv \rho(\vec{r_0}, E_F). \tag{3.5}$$

Der Tunnelstrom ist demnach proportional zur lokalen Zustandsdichte der Probenoberfläche am Ferminiveau am Ort $\vec{r_0}$, dem Zentrum der Krümmung der Spitze, und STM-Abbildungen im Konstantstrommodus entsprechen nach dieser Theorie der Topographie konstanter elektronischer Zustandsdichte nahe dem Ferminiveau. Ein wesentlicher Vorteil dieser Theorie ist, daß die STM-Abbildungen auf eine anschauliche Größe zurückgeführt werden kann.

Leider ist diese einfach Interpretation der Konstantstrombilder für höhere Tunnelspannungen und Wellenfunktionen der Spitze mit $m \neq 0$ nicht länger gültig. Für normal dotierte Halbleiteroberflächen liegt das Fermi-Niveau innerhalb der Bandlücke des Volumens und Tunneln ist ohne metallische Oberflächenzustände erst aus den höchsten Valenzbandzuständen oder in die niedrigsten Leitungsbandzustände möglich. In der Praxis werden für das Tunneln auf Halbleiteroberflächen gewöhnlich Tunnelspannungen um 1–2 V gewählt, um stabile Tunnelbedingungen zu erreichen. Damit wird die Bedingung kleiner Tunnelspannungen in der Theorie von Tersoff und Hamann verletzt und es tragen nicht allein die Zustände am Fermi-Niveau zum Tunnelstrom bei, sondern alle besetzen Probenzustände in einem Energiefenster $(E_F, E_F - eV)$ bei negativer Probenspannung und alle unbesetzten Probenzustände im Energiefenster $(E_F + eV, E_F)$ bei positiver Probenspannung. Es ist nicht länger zu erwarten, daß die Wellenfunktionen der Spitze und der Probe voneinander ungestört bleiben. Hält man diese Bedingung dennoch aufrecht, kann der Tunnelstrom mit Hilfe der WKB-Näherung durch folgenden Ausdruck angenähert werden [35]:

$$I(\vec{r_0}, V) \propto \int_{E_F}^{E_F+eV} dE \rho_{Spitze}(E + eV) \rho_{Probe}(\vec{r_0}, E) T(\vec{r_0}, E, eV), \tag{3.6}$$

dabei bedeutet $\rho_{Spitze}(E+eV)$ die Zustandsdichte der Spitze, $\rho_{Probe}(\vec{r_0}, E)$ die Zustandsdichte der Probe am Ort $\vec{r_0}$ und $T(\vec{r_0}, E, eV)$ die Transmissionswahrscheinlichkeit der Elektronen. Die Tunnelwahrscheinlichkeit ist empfindlich energieabhängig. Wie in Abb. 3.6 durch die Pfeillänge angedeutet, tragen bei positiver Probenspannung die besetzten Zustände am Fermi-Niveau der Spitze und die unbesetzten Zustände mit Energie $E_F + eV$ der Probe am stärksten

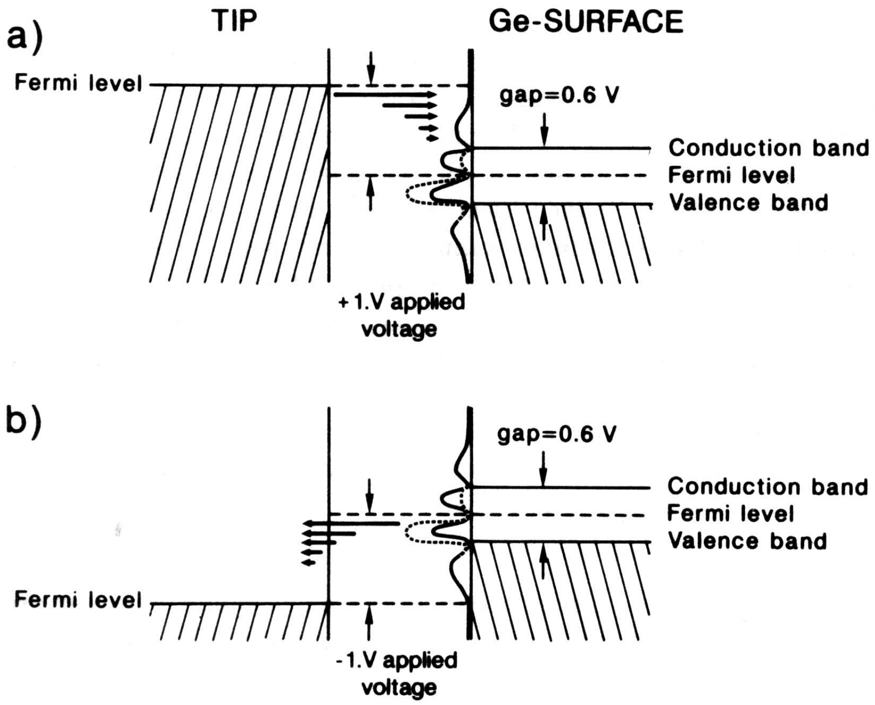

Abbildung 3.6: Schematisches Energiediagramm des Tunnelprozesses aus einer Metallspitze in eine Halbleiteroberfläche für (a) positive und (b) negative Probenspannung (aus [36]).

zum Tunnelstrom bei. Bei negativer Probenspannung kommt der größte Beitrag zum Strom aus denjenigen besetzten Zuständen der Probe, die dem Fermi-Niveau am nächsten liegen.

Neben der Tersoff-Hamann Theorie wurden eine Vielzahl weiterer Theorien des STM entwickelt, die mit Hilfe von den speziellen Problemen angepaßten Näherungen unterschiedliche Aspekte des Tunnelprozesses beschreiben. Störungsrechnungen erster Ordnung wurden über die Tersoff-Hamann Theorie hinaus erweitert. So konnte durch Einbeziehung von Orbitalen der Spitze mit höheren Bahndrehimpulsen die nach der Theorie von Tersoff und Hamann unerwartet hohe maximale Ortsauflösung in STM-Bildern erklärt werden [37, 38, 39]. Der Einfluß des konkreten Zustands der Spitze auf die Abbildung wurde anhand von Rechnungen unter Verwendung von Wellenfunktionen verschiedener realistischer Cluster als Mikrospitzen studiert [40]. Nicht-Störungsrechnungen besitzen den Vorteil, daß die Wechselwirkung zwischen Spitze und Probe nicht als klein angenommen werden muß. Garcia et al.[41] und Stoll [42] nutzten die quantenmechanische Streutheorie, um einen Ausdruck für den Tunnelstrom zu abzuleiten. Der Tunnelwiderstand eines einzelnen Atoms, welches zwischen zwei planaren *jellium*-Oberflächen adsorbiert ist, wurde von Lang [43] innerhalb der Dichtefunktionaltheorie berechnet. Eine vollständige Theorie zur Berechnung des Tunnelstroms zwischen zwei Elektroden wurde von Noguera entwickelt [44]. Sie basiert auf der Anpassung von Greens-Funktionen an die beiden Oberflächen. Übersichten über STM-Theorien mit und ohne Störungsrechnung finden sich in [40, 45, 46].

Ein bedeutender Vorteil der Tersoff und Hamann Theorie besteht darin, daß sie direkt mit den selbst-konsistenten Potentialen aus Totalenergierechnungen oder *ab initio* Molekulardynamikrechnungen kombiniert werden kann, welche gleichzeitig die atomare geometrische Struktur und die Ladungsdichte auch komplizierter Rekonstruktionen reiner oder adsorbatbedeckter Oberflächen liefern. Simulationen von STM-Bildern wurde bereits für eine größere Anzahl von Strukturen durchgeführt (z. B.[32, 33, 34]. Es ist zu vermuten, das Simulationsrechnungen in der näheren Zukunft noch an Bedeutung gewinnen und eine Standarthilfsmittel für die Interpretation von STM-Bildern werden. Einen Ansatz jenseits der Störungstheorie zur Berechnung von STM-Bildern komplexer Strukturen in Kombination mit*ab initio* Rechnungen geben Ness und Fisher [46].