

Kapitel 3

Stromtransport – Methoden und Theorie

Im Folgenden sollen die theoretischen und meßtechnischen Grundlagen der im Rahmen der vorliegenden Arbeit angewendeten Methoden zur Bestimmung der *Ladungstransporteigenschaften* des Materials CuGaSe_2 dargestellt werden.

Die Ladungstransporteigenschaften sind

- die Leitfähigkeit, die durch die Messung des Spannungsfalls an einer Materialprobe bei einem vorgegebenen Stromfluß bestimmt werden kann.
- die Dotierung, die entweder
 1. über den Hall-Effekt oder
 2. durch die Spannungsabhängigkeit der Kapazität einer Raumladungszone
 meßbar ist.
- die Beweglichkeit als Maß für die Leitfähigkeit pro Ladungsträger.

3.1 Leitfähigkeit und Beweglichkeit

Anders als eine freie Ladung Q , die im elektrischen Feld \mathbf{E} eine konstante Beschleunigung

$$\mathbf{a} = \frac{Q}{m} \mathbf{E} \quad (3.1)$$

erfährt – m ist ihre Masse, stellt sich im Leiter aufgrund der Streuung an den Atomen des Kristallgitters nach einer charakteristischen Zeit τ , die im Mittel zwischen zwei Streuereignissen vergeht, eine konstante Driftgeschwindigkeit

$$\mathbf{v}_d = -\mu_d \mathbf{E} \quad (3.2)$$

der Elektronen ein [27]. Die Proportionalitätskonstante μ_d heißt in diesem Zusammenhang Driftbeweglichkeit und ist eine Materialkonstante, die die Geschwindigkeit angibt, mit der sich ein Ladungsträger im Mittel in einem elektrischen Feld bewegt.

Mit der Anzahldichte n der Elektronen und ihrer Ladung e kann aus der Driftgeschwindigkeit die Stromdichte

$$\mathbf{j} = -en\mathbf{v}_d = \sigma\mathbf{E} \quad (3.3)$$

angegeben werden, wobei die Leitfähigkeit $\sigma = en\mu$ ist, wie mit (3.2) abgelesen werden kann. Ihre allgemeine Form bei der Anwesenheit zweier Ladungsträgersorten unterschiedlicher Konzentrationen n, p und Beweglichkeiten μ_n, μ_p ist

$$\sigma = e(\mu_n n + \mu_p p). \quad (3.4)$$

Für genügend hoch dotierte Halbleiter kann dabei der Ausdruck für die Minoritätsladungsträger jeweils vernachlässigt werden. Für einen p-Halbleiter wie das CuGaSe_2 lautet sie dann

$$\sigma = e(\mu_p p). \quad (3.5)$$

3.2 Ladungsträgerkonzentration

Hall-Messungen ermöglichen die direkte Bestimmung der effektiven Ladungsträgerkonzentration eines Halbleiters und in Verbindung mit der Leitfähigkeit auch die Bestimmung der Ladungsträgerbeweglichkeit. Auch aus Kapazitätsmessungen an Raumladungszonen, wie sie bei Schottky-Kontakten oder an p-n-Übergängen auftreten, sind unter gewissen Voraussetzungen, die im Folgenden erörtert werden sollen, Aussagen über die Dotierung des Materials im Bereich der betreffenden Raumladungszone möglich.

3.2.1 Der Hall-Effekt

Der Hall-Effekt [27, 104] folgt direkt aus der Lorentzkraft \mathbf{F} auf im Magnetfeld \mathbf{B} mit der Geschwindigkeit \mathbf{v} bewegte Ladungen Q

$$\mathbf{F} = Q(\mathbf{v} \times \mathbf{B}). \quad (3.6)$$

Für eine Stromdichte j_x in x -Richtung in ei-

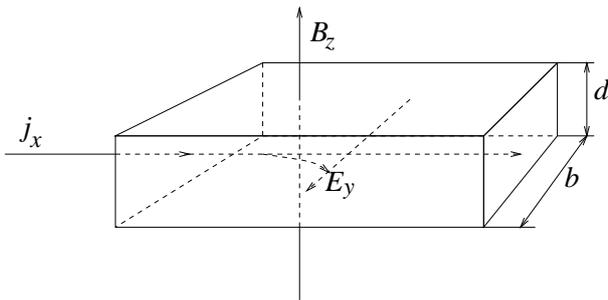


Abbildung 3.1: Hall-Spannung

nem Magnetfeld in z -Richtung gilt für das elektrische Feld in y -Richtung (Abbildung 3.1)

$$E_y = R_H j_x B_z. \quad (3.7)$$

Die Größen B_z , $E_y = \frac{U_y}{b}$ und $j = \frac{I}{b \cdot d}$ sind einer Messung direkt zugänglich — b und d bezeichnen jeweils die Breite und Dicke der vermessenen Probe¹. Die Proportionalitätskonstante

$$R_H = r \frac{1}{e} \frac{p - n \left(\frac{\mu_n}{\mu_p} \right)^2}{\left(p + n \frac{\mu_n}{\mu_p} \right)^2} \quad (3.8)$$

wird Hall-Koeffizient genannt. Wenn eine der Ladungsträgersorten n oder p überwiegt, folgt aus dem Massenwirkungsgesetz

$$n \cdot p = n_i^2 = N_C N_V \exp - \frac{E_G}{kT}, \quad (3.9)$$

sofort eine entsprechende Abnahme der Dichte der anderen Ladungsträger. n_i bezeichnet die intrinsische Ladungsträgerdichte, die unabhängig von einer Fremddotierung nur von den Zustandsdichten N_C und N_V an den Bandkanten, der Energielücke E_G und der Temperatur abhängt.

CuGaSe₂ neigt wie alle Chalkopyrite zur Dotierung durch Eigendefekte [68], das heißt Kristallfehler, die sich bei der Herstellung ergeben, ohne daß eine dem Material fremde Atomsorte anwesend ist. Es ist nun eine empirische Tatsache, daß CuGaSe₂ praktisch ausschließlich p-Leitung zeigt [89] – wenn auch neueste Arbeiten [91] eine n-Dotierung für möglich erachten.

Damit kann in (3.8) $n \frac{\mu_n}{\mu_p}$ gegen p vernachlässigt und vereinfacht

$$R_H \approx r \frac{1}{e} \frac{1}{p} \quad (3.10)$$

geschrieben werde.

Immerhin können, wie in Kapitel 6 gezeigt wird, in CuGaSe₂ erhebliche Donatorkonzentrationen auftreten, die einen Teil der Ak-

¹Bei den in der vorliegenden Arbeit betrachteten epitaktischen Schichten auf semiisolierendem GaAs sind dies die Dimensionen der Schicht

zeptoren *kompensieren* und die effektive Ladungsträgerkonzentration entsprechend absenken. Da jedoch auch bei erheblicher Kompensation durch Minoritätsladungsträger wegen (3.9) immer noch $p \gg n$ gilt, kann meist noch mit (3.10) gerechnet werden. Der Unterschied zur exakten Form (3.8) wird erst bei Annäherung der Ladungsträgerdichten $n \rightarrow p$ erheblich, wie zum Beispiel bei hohen Temperaturen, wenn die Ladungsträgerdichten durch Anregung über die Bandlücke rapide erhöht werden. Dieser Fall wird in Abschnitt 3.2.2.1 anhand einer Beispielrechnung nochmals quantitativ behandelt.

Damit steht einer Berechnung der Ladungsträgerdichte nur mehr die Größe r im Wege. r ist der Quotient aus dem Ensemblemittelwert $\langle \tau^2 \rangle$ und dem Quadrat $\langle \tau \rangle^2$ der mittleren freien Flugzeit τ zwischen zwei Streueignissen [104]. Sie hängt für verschiedene Streumechanismen unterschiedlich von der Energie der Ladungsträger ab. In nichtentarteten Halbleitern ergibt sich durch Integration der Energieabhängigkeit über die Boltzmannfunktion für die Streuung an neutralen Defekten $r = 1$, bei Elektron-Phonon Streuung $r = 1.18$ und bei Streuung an ionisierten Störstellen $r = 1.93$ [104]. Der maximale Fehler bewegt sich so zwar um einen Faktor zwei, was jedoch bei einer um viele Größenordnungen variierenden Größe wie der Ladungsträgerdichte oft vernachlässigt werden kann.

Versetzungen in Einkristallen oder gar Korngrenzen in polykristallinen Filmen können die Streuwahrscheinlichkeit erheblich erhöhen, wodurch die Energieabhängigkeit deutlich komplexer und in der Regel nicht mehr theoretisch vorhergesagt werden kann. Da wie schon erwähnt die Ladungsträgerdichte um viele Größenordnungen variieren kann und demgegenüber die Abweichung von r gegenüber 1 vernachlässigt werden kann, wird in der folgenden Diskussion $r = 1$ angenommen [63].

3.2.1.1 Messung der Hall-Spannung

Um die effektive Ladungsträgerkonzentration eines Halbleiters aus dem Hall-Effekt zu bestimmen, wird ein Strom bekannter Stärke I über ein Kontaktpaar längs durch ein quaderförmiges Probenstück getrieben, wie in Abbildung 3.2 schematisch gezeigt. Senkrecht zur Stromrichtung wird ein Magnetfeld B bekannter Flußdichte angelegt. An einem zum Strompfad und zum Magnetfeld senkrecht angeordneten Kontaktpaar wird dann eine Spannung U gemessen, die nach (3.7) proportional sowohl zu B_z als auch j_x ist, darüber hinaus jedoch nur noch vom Hall-Koeffizienten R_H und damit über (3.8) mittelbar von den Ladungsträgerkonzentrationen n und p abhängt. Dabei geht so-

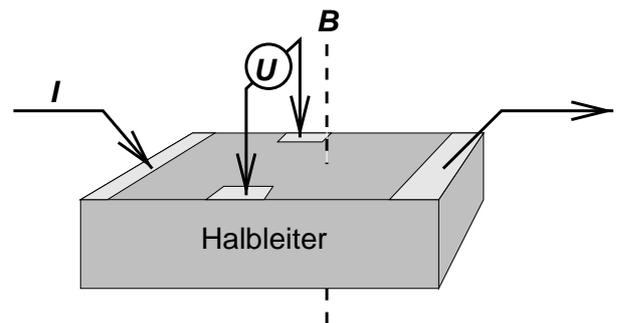


Abbildung 3.2: Geometrie zur Messung der Hall-Spannung

wohl die Probengeometrie explizit durch b und d als auch die Anordnung der Kontakte implizit in die Messung mit ein: Wenn zum Beispiel die Kontakte nicht exakt in senkrecht zueinander verlaufenden Richtungen angelegt sind, addiert sich zur meist sehr kleinen Hall-Spannung eine nicht zu vernachlässigende Zusatzspannung aufgrund des Ohmschen Spannungsfalls zwischen den Stromkontakten. Diese Spannung wird durch die Mittelung über mehrere Messungen unter Umkehrung der Strom- und Magnetfeldrichtungen eliminiert.

3.2.1.2 Nach Van der Pauw

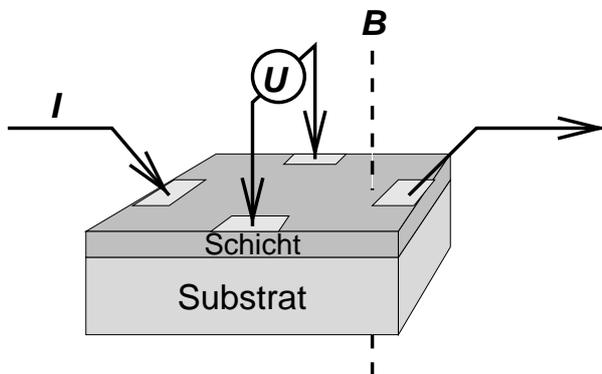


Abbildung 3.3: Geometrie zur Messung der Hall-Spannung nach Van der Pauw

Ist die benötigte Geometrie nicht mit angemessener Genauigkeit oder zum Beispiel bei nicht-kubischen Einkristallen überhaupt nicht herstellbar, kann die Methode nach Van der Pauw [2, 108, 94] angewendet werden: Dabei wird das Probenstück wie in Abbildung 3.3 mit einer Anordnung von vier Kontaktpunkten an dessen Rand versehen. Die Probe muß dazu lediglich eine einheitliche Dicke aufweisen und topologisch aus einer einfach zusammenhängenden Schicht bestehen.

Zusätzlich zu den Umkehrungen der Magnetfeld- und Stromrichtung werden nun noch Strom- und Spannungspfad vertauscht und durch Leitfähigkeitsmessungen über dieselben Kontakte, ebenfalls unter Vertauschung von Strom- und Spannungspfaden, ein Korrekturfaktor ermittelt, der leichte Abweichungen in der Symmetrie der Kontakte und Probe berücksichtigt. Damit lassen sich Proben beinahe beliebiger Form charakterisieren.

3.2.1.3 Meßplatz

Für die in Kapitel 6 gezeigten Daten aus Hall- und Leitfähigkeitsmessungen wurde ein beste-

hender Meßplatz am HMI verwendet, an dem mit einem normalleitenden Magnet ein Magnetfeld von bis zu 1 Tesla am Probenort angelegt werden kann. Der Eingangswiderstand des Spannungsmeßgerätes beträgt im Hochwiderstandsmodus durch Zwischenschalten eines Operationsverstärkers $100 \text{ T}\Omega$, im Niederwiderstandsmodus beträgt er immer noch $10 \text{ G}\Omega$ [44]. Damit sind Messungen von Schichtwiderständen bis zu etwa $100 \text{ G}\Omega$ bzw. $10 \text{ M}\Omega$ möglich [2]. Temperaturabhängige Messungen können von 15 K bis 480 K durchgeführt werden. Es wurde ausschließlich nach Van der Pauw (Abschnitt 3.2.1.2) [108] und in Dunkelheit gemessen und nach [2] ausgewertet. Kam es dabei zu Abweichungen zwischen den einzelnen zu mittelnden Werten von mehr als 10 %, wurde die betreffende Messung als ungültig verworfen.

Die Tieftemperaturmessungen wurden stets vor denjenigen über Raumtemperatur durchgeführt, da sich die Eigenschaften der Proben durch die hohe Temperatur mitunter bleibend verändern können.

3.2.2 Temperaturabhängigkeit

Aus Messungen der Temperaturabhängigkeit der Ladungsträgerdichte kann deren Anregungsenergie bestimmt werden, aus der wiederum Rückschlüsse auf den energetischen Abstand der Dotierniveaus von der jeweiligen Bandkante möglich sind. Eine große Anregungsenergie bedeutet einen großen Abstand oder einen entsprechend „tiefen“ Potentialtopf, aus dem heraus die Ladungsträger angeregt werden müssen. Entsprechend sind kleine Anregungsenergien mit einer flachen Potentialmulde verknüpft. Wenn diese Potentialmulde durch eine punktförmige Ladungsanordnung verursacht wird, kann ein in ihr befindlicher Ladungsträger in erster Näherung wie in einem Wasserstoffatom betrachtet werden, das in ein

Dielektrikum mit der Dielektrizitätszahl ϵ eingebettet ist, und sowohl von einem Elektron als auch von einem Loch umgeben sein kann, deren Masse m jeweils aufgrund der Struktur des Leitungs- beziehungsweise Valenzbandes modifiziert ist. Die Anregungsenergie ergibt sich dann zu [27]

$$E = E_i = -\frac{m e^4}{(4\epsilon\epsilon_0 h)^2} = 13.61 \frac{m}{\epsilon^2} \text{ eV}. \quad (3.11)$$

Für den Halbleiter gilt im Gleichgewicht die Ladungsneutralität

$$n + \sum_i N_{Ai}^- = p + \sum_i N_{Di}^+ \quad (3.12)$$

sämtlicher ionisierter Akzeptoren N_{Ai}^- und Donatoren N_{Di}^+ mit den freien Ladungsträgern n und p .

Für zunächst je einen Donator- und Akzeptorzustand mit den Konzentrationen N_D^- und N_A^+ lautet sie also

$$n + N_A^- = p + N_D^+, \quad (3.13)$$

wobei der Ionisierungsgrad durch die Verteilungsfunktionen

$$N_D^+ = \frac{N_D}{1 + g_D \exp \frac{E_F - E_D}{kT}} \quad (3.14)$$

$$N_A^- = \frac{N_A}{1 + g_A \exp \frac{E_A - E_F}{kT}} \quad (3.15)$$

aus den Konzentrationen der Akzeptoren und Donatoren und ihren energetischen Lagen E_A und E_D berechnet werden kann. E_F ist das Fermipotential, das als zentrale Größe alle an der Ladungsbilanz beteiligten Verteilungsfunktionen miteinander in Beziehung setzt.

In einem nicht entartet dotierten Halbleiter gilt für die Ladungsträgerdichten

$$n = N_C \exp -\frac{E_C - E_F}{kT} \quad (3.16)$$

$$p = N_V \exp \frac{E_V - E_F}{kT}, \quad (3.17)$$

worin

$$N_C = 2 \left(\frac{2\pi m_e kT}{h^2} \right)^{3/2} M_C \quad (3.18)$$

$$N_V = 2 \left(\frac{2\pi m_h kT}{h^2} \right)^{3/2} M_V \quad (3.19)$$

die effektiven Zustandsdichten der Elektronen und Löcher an den Bandkanten beschreiben. E_C und E_V bezeichnen die energetische Lage des Leitungs- und Valenzbandes. m_e ist die effektive Masse der Elektronen, m_h die der Löcher. Für beide gilt

$$N_{C,V} = (m_{e,h}^* T)^{3/2} M \cdot 4.8 \cdot 10^{15} \text{ cm}^{-3}, \quad (3.20)$$

mit den effektiven Massen $m_{e,h}^*$ der Ladungsträger in Einheiten der Elektronenmasse m_0 und der Anzahl äquivalenter Bandextrema M ; für direkte Halbleiter mit einem Extremum am Γ -Punkt pro Band gilt $M = 1$.

g gibt die Entartung der betreffenden Bandkanten an. Bei CuGaSe_2 gilt stets $g = 2$, da die energetische Entartung der Valenzbandkante wie in Abschnitt 2 erläutert, durch das Kristallfeld und die Spin-Bahn-Kopplung aufgehoben ist und die Zustände an der Bandkante von maximal zwei Elektronen mit entgegengesetztem Spin besetzt werden können.

Aus der Ladungsneutralität folgt dann eine allgemeine Bestimmungsgleichung für das Ferminiveau E_F :

$$N_C \exp -\frac{E_C - E_F}{kT} + \frac{N_A}{1 + g_V \exp \frac{E_A - E_F}{kT}} = N_V \exp \frac{E_V - E_F}{kT} + \frac{N_D}{1 + g_L \exp \frac{E_F - E_D}{kT}}. \quad (3.21)$$

Das Ferminiveau E_F ist damit implizit festgelegt und kann für einen Satz der darin enthaltenen übrigen Parameter bestimmt werden [104]. In vielen Fällen kann (3.21) so genähert werden, daß eine überschaubare Auflösung nach

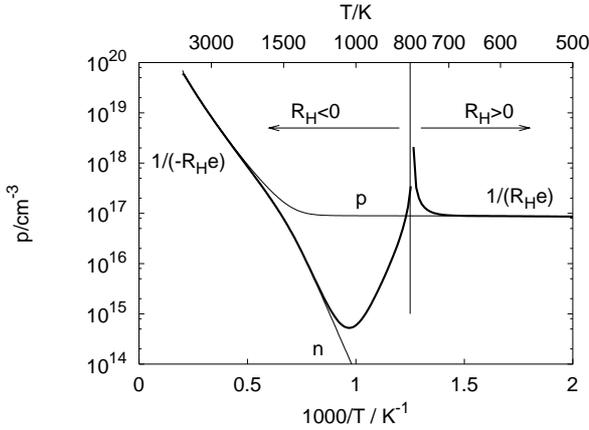


Abbildung 3.4: Ladungsträgerdichten und Hall-Koeffizient. Jeweils durchgezogen sind die Löcherdichte p und der umgerechnete Hall-Koeffizient mit positivem Vorzeichen, gestrichelt sind die Elektronendichte n und die Beträge aus dem negativen Hall-Koeffizient $-R_H$ eingezeichnet.

E_F möglich ist. Die für das CuGaSe_2 relevanten Fälle sollen im Folgenden näher betrachtet werden.

3.2.2.1 Hall-Koeffizient und Ladungsträgerdichten

Sobald beide Ladungsträgersorten im selben Material vorhanden sind, stellt Gleichung (3.8) prinzipiell ein Hindernis beim Vergleich von nach (3.17) *errechneten* Ladungsträgerdichten und dem aus *gemessenen* Hall-Spannungen nach (3.7) ermittelten Hall-Koeffizienten R_H dar. Wie jedoch in (3.10) diskutiert, wird der Fehler nur bei $n \approx p$ signifikant. Die Auswertung von (3.21) in Abbildung 3.4 ermöglicht eine quantitative Betrachtung dieses Falls:

Aus den PL-Messungen in [8] wird ein Donator 12 ± 5 meV unterhalb der Leitungsbandkante identifiziert. Mit dem Wasserstoffmodell (3.11) folgt daraus eine effektive Elektronenmasse von $m_e(0.11 \pm 0.04)m_0$. Daraus wird mit (3.18) $N_C = 1.8 \cdot 10^{14} \text{ cm}^{-3} \text{ K}^{-3/2}$. Mit der

im Rahmen dieser Arbeit ermittelten effektiven Löchermasse von $m_h = 1.5m_0$ gilt für die Zustandsdichte im Valenzband (3.19) $N_V = 1 \cdot 10^{16} \text{ cm}^{-3} \text{ K}^{-3/2}$. Für eine Streuung an akustischen Phononen, wie sie im Rahmen dieser Arbeit an den meisten der untersuchten CuGaSe_2 -Proben beobachtet wird (Kapitel 6), kann nach [104]

$$\mu \propto m^* - 5/2 \quad (3.22)$$

angenommen werden. Für das Verhältnis der Beweglichkeiten ergibt sich danach

$$\frac{\mu_n}{\mu_p} = \left(\frac{m_h}{m_e} \right)^{5/2} = 690. \quad (3.23)$$

Mit diesen Ergebnissen wurde in Abbildung 3.4 die Elektronen- und Löcherdichte mit (3.8) in den Hall-Koeffizienten R_H umgerechnet. Erst bei einer Temperatur von 800 K, wenn sich Löcher- und Elektronendichte aufgrund der Anregung über die Bandlücke hinweg annähern, weicht $\frac{1}{R_H e}$ merklich von p ab und wechselt je nach Verhältnis der Beweglichkeiten sein Vorzeichen. Für die Auswertung von Messungen bei Temperaturen über 300 K wird diese Beobachtung eine Rolle spielen. Für die folgenden Betrachtungen kann jedoch $p \approx \frac{1}{R_H e}$ gesetzt werden.

3.2.3 Mehrere Dotanden

In den meisten Fällen finden sich in Halbleitern sowohl Elektronen wie auch Löcher bindende Defekte unterschiedlicher Anregungsenergie, die sich entsprechend gegenseitig *kompensieren* können. Deshalb wurden im vorigen Abschnitt alle Gleichungen schon für beide Defektarten aufgeführt. Erst wenn eine der Defektarten die Oberhand gewinnt, kommt es zu einer Nettodotierung und zu einem Überschuss an freien Ladungsträgern, die sich auch elektrisch nachweisen lassen.

Chalkopyrite neigen, wie bereits in Abschnitt 3.2.1 erwähnt, zur Eigendotierung, wobei CuGaSe₂ vornehmlich Elektronenhaftstellen, also Akzeptoren ausbildet und damit p-leitend wird. Eine n-Dotierung ist bis auf einen eher als exotisch zu bezeichnenden Fall [91], bei dem nach Ionenimplantation eine schwache n-Dotierung gemessen wurde, nicht beobachtet worden; jedoch kann, wie sich in Kapitel 6 herausstellen wird, der Anteil der Kompensation durch donatorische Zustände bis zu 96 % der Akzeptorkonzentration ausmachen. Da das Fermi-niveau in p-dotierten Halbleitern stets weit von der Valenzbandkante und damit den Donatorzuständen entfernt ist, kann in Gleichung (3.14) die e -Funktion gegen 1 vernachlässigt und allgemein für CuGaSe₂ $N_D^+ = N_D$ gesetzt werden.

Im Folgenden soll gezeigt werden, wie aus dem Temperaturverlauf der Nettoladungsträgerdichte Aussagen über die tatsächlichen Dotierkonzentrationen gewonnen werden können. Davon wird der Unterschied zwischen dem gemessenen Hall-Koeffizienten R_H und der Nettoladungsträgerkonzentration quantitativ betrachtet.

3.2.3.1 Kompensation

$$\begin{array}{ll} E_V=0 \text{ eV} & N_V/T^{3/2}=10^{16}\text{K}^{-3/2}\text{cm}^{-3} \\ E_C=1.68 \text{ eV} & N_C/T^{3/2}=10^{16}\text{K}^{-3/2}\text{cm}^{-3} \\ N_A=10^{17} \text{ cm}^{-3} & N_D=10^{15}, 6 \cdot 10^{16} \text{ cm}^{-3} \\ E_A=0.110 \text{ eV} & \end{array}$$

Tabelle 3.1: Parameter für die Berechnung der Ladungsträgerdichte in Abbildung 3.5

In Abbildung 3.5 wurde mit (3.17) und den Parametern in Tabelle 3.1 jeweils die effektive Ladungsträgerdichte p als Funktion der reziproken Temperatur bei drei verschiedenen Donatorkonzentrationen N_D berechnet: Bei verschwindender Donatorkonzentration $N_D <$

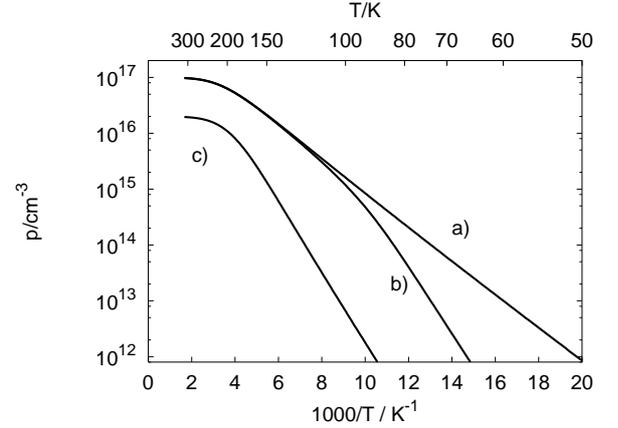


Abbildung 3.5: Auswirkung der Kompensation auf die Ladungsträgerkonzentration. Parameter siehe Tabelle 3.1.

10^{12} cm^{-3} , Kurve a), bei im Vergleich zur Akzeptorkonzentration niedriger, Kurve b) und bei hoher von $6 \cdot 10^{16} \text{ cm}^{-3}$, Kurve c), also einer Kompensation von 60 %.

Die Bandlücke von CuGaSe₂ ist $E_C - E_V = E_G = 1.68 \text{ eV}$.

Bei verschwindender Kompensation

$$N_A \gg N_D, \quad (3.24)$$

kann in Gleichung (3.21) sowohl der Summand mit N_D , als auch die Elektronenkonzentration (3.16) vernachlässigt werden. Da in einem p-Halbleiter das Fermi-niveau weit von der Leitungsbandkante entfernt ist, gilt $E_C - E_F \gg kT$ womit die Exponentialfunktion in (3.16) verschwindet, so daß schließlich Gleichung (3.21) genähert nach

$$E_F = \ln \frac{1}{4} E_A \left(\sqrt{1 + \frac{8N_A}{N_V} \exp \frac{E_A - E_V}{kt}} - 1 \right) \quad (3.25)$$

aufgelöst werden kann [60].

In (3.17) eingesetzt folgt

$$p \approx \sqrt{\frac{N_A N_V}{2}} \exp -\frac{E_A}{2kT}. \quad (3.26)$$

Die Steigung bei tiefen Temperaturen und verschwindender Donatorkonzentration entspricht also der *halben* Aktivierungsenergie. In dem in Abbildung 3.5 gewähltem Beispiel der Kurve a) $\frac{1}{2}E_A = 55$ meV. Sie mündet bei $N = N_A = p_{max} = 10^{17}$ cm⁻³ in die Akzeptorerschöpfung.

Ist eine nicht verschwindende aber im Vergleich zu den Akzeptoren geringere Dichte ionisierter Donatoren vorhanden $N_A^- > N_D^+$, und außerdem diese Donatorkonzentration größer als die Boltzmann-gewichtete Zustandsdichte des Valenzbands

$$N_D \gg \frac{1}{2}N_V \exp\left(-\frac{E_A}{kT}\right), \quad (3.27)$$

knickt die effektive Ladungsträgerkonzentration auf der Höhe dieser Dichte (in Abb. 3.5 Kurve b) bei der Donatorkonzentration $N_D = 10^{15}$ cm⁻³) nach unten ab und verläuft dann wie

$$p \approx \frac{N_A - N_D}{2N_D} N_V \exp\left(-\frac{E_A}{kT}\right), \quad (3.28)$$

($p = N_A^- - N_D^+$) [60] mit der *vollen* Aktivierungsenergie $E_A = 110$ meV. Die Anregungsenergie der *Donatoren* ist dabei unerheblich, da sie ihre Elektronen nicht an das energetisch höher liegende Leitungsband, sondern an die tiefer liegenden Akzeptoren abgeben.

Liegt die Donatordichte in derselben Größenordnung wie die Akzeptordichte (in Abb. 3.5 Kurve c), so verläuft die Ladungsträgerdichte insgesamt niedriger und von der Akzeptorerschöpfung an mit der vollen Aktivierungsenergie $E_A = 110$ meV.

Aus dem Vorfaktor $\frac{N_A - N_D}{2N_D} N_V$ kann aus einer einfachen Anpassung dieser Gleichung an die Messung das Verhältnis von dotierenden Akzeptoren N_A zu kompensierenden Donatoren N_D bestimmt werden, wenn die effektive Masse der Löcher bekannt ist².

²Eine kurze Literaturstudie zur Frage nach der *effekti-*

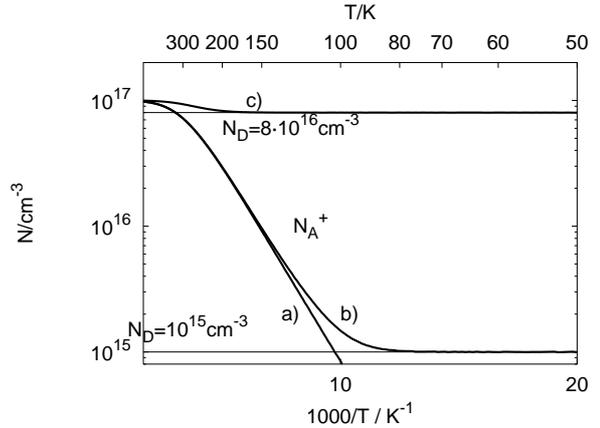


Abbildung 3.6: Auswirkung von Kompensation auf die Ladungsträgerkonzentration. Parameter siehe Tabelle 3.1.

In Abbildung 3.6 ist für dieselben Fälle wie in Abbildung 3.5 mit (3.15) die Konzentration ionisierter Akzeptoren N_A^- berechnet worden. Die Donatoren bleiben immer vollständig ionisiert. Kurve 3.6a) beschreibt den Fall ver-

ven Masse von Löchern in CuGaSe₂ m_h fördert ein leider allzu häufig beobachtetes Phänomen zutage:

Während [79] einen Wert von $1.2m_0$ aus elektrischen Messungen von [115] angibt, nennt dieser denselben Wert Resultat eines „theoretical estimate of the density of states“ und zitiert wiederum [59]. Dieser schließlich grenzt den Wert der Löchermasse durch die *gemessenen* Massen in CuGaS₂[117] und CuGaTe₂[70] zwischen 0.7 und $1.3m_0$ ein. Seine *Überlegung* führt dann zu einer Masse von $m_h(\text{CuGaSe}_2) = 1.2m_0$.

Der Wert für CuGaS₂ basiert dabei wiederum auf Hall-Messungen an Einkristallen, der Wert für CuGaTe₂ auf ebensolchen an epitaktischen Schichten. Beide Werte resultieren aus der Anpassung von

$$\frac{p(p + N_D)}{N_A - N_D - p} = \frac{1}{\beta} \frac{2(2\pi m_h kT)^{3/2}}{h^3} \exp(-E_A/kT),$$

mit den freien Parametern N_A, N_D, E_A und m_h an die gemessenen $p(T)$ Verläufe. β ist darin die Entartung des obersten Valenzbandes und wird mit 2 angenommen. [70] erhält damit für CuGaTe₂ Werte zwischen 1.16 und $1.45m_0$ und untermauert diese mit einer theoretischen Abschätzung von $1.26m_0$, [117] gibt für CuGaS₂ nur lapidar $0.69m_0$ an.

schwindender Donatorkonzentration, die Akzeptoren frieren exponentiell aus. Bei Kurve 3.6b) biegt die Konzentration ionisierter Akzeptoren bei der Konzentration der Donatoren ab, gleichzeitig ändert sich in Abbildung 3.5 die Steigung der Ladungsträgerkurve. Die Neutralitätsbeziehung (3.13) $p = N_A^- - N_D^+$ wirkt sich hier aus, die von den Donatoren freigesetzten Elektronen werden sofort von den Akzeptoren eingefangen.

Bei hoher Kompensation (Kurve 3.6c) erfolgt dieser Übergang direkt im Anschluß an die Akzeptoreerschöpfung.

3.2.3.2 Zwei Akzeptoren

Wie schon in Abschnitt 2.3.2 angesprochen, wurden durch Photolumineszenzuntersuchungen in CuGaSe₂ zwei akzeptorartige Zustände neben einem donatorartigen gefunden [8]. Auch das läßt sich durch (3.21) simulieren, indem in (3.13) ein weiterer Summand N_{A2}^- (3.15) hinzugefügt wird. Zunächst soll dabei von einer verschwindenden Donatorkonzentration ausgegangen werden, so daß (3.13)

$$n + N_{A1}^- + N_{A2}^- = p$$

wird. Abbildung 3.7 zeigt diesen Fall für $E_{A2} = 80$ meV und drei verschiedene Konzentrationen $N_{A2} = 10^{18}$ (3.7a), 10^{15} (3.7b) und 10^{12} cm⁻³ (3.7c). Alle anderen Parameter wurden von Tabelle 3.1 übernommen.

Besitzt der Akzeptor mit der niedrigeren Anregungsenergie E_{A2} die größere Dichte $N_{A2} > N_{A1}$, bestimmt er das Verhalten im gesamten Temperaturbereich bis zu seiner Erschöpfung, wie Kurve 3.7a) zeigt. Die Steigung entspricht der halben Anregungsenergie, also 40 meV.

Ist dagegen der mit der höheren Anregungsenergie E_{A1} häufiger $N_{A1} > N_{A2}$, bestimmt er die Ladungsträgerdichte bei hoher Temperatur mit einer größeren Steigung entsprechend

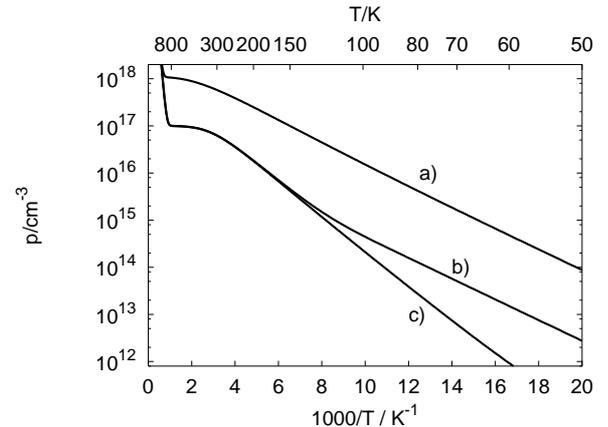


Abbildung 3.7: Löcherkonzentration bei zwei Akzeptorniveaus

110/2 meV, bis er unter das Niveau des „flacheren“ Akzeptors ausgefroren ist, der dann das Regime übernimmt (Kurve 3.7b). Am linken oberen Bildrand tritt wieder die intrinsische Ladungsträgerdichte in Erscheinung.

Für Kurve 3.7c) ist $N_{A2} < 10^{12}$ cm⁻³ und damit bleibt die Steigung im gezeigten Temperaturbereich bis zur Erschöpfung bei hohen Temperaturen gleich.

3.2.3.3 Mehrere Akzeptoren und Kompensation

Dieser Fall ist gerade bei Chalkopyriten wegen ihrer Vielzahl an Eigendefekten zu erwarten [68]. Dazu muß zum vorherigen Fall wieder eine endliche Donatorkonzentration N_D zugelassen werden und die Ladungsneutralität

$$n + N_{A1}^- + N_{A2}^- = p + N_D^+$$

verwendet werden.

Abbildung 3.8 stellt somit eine Kombination aus 3.5 und 3.7 dar: Zusätzlich zu den Parametern in 3.7 wurde für die Kurven 3.8a) b) und c) noch ein Akzeptor bei $N_D = 10^{13}$ cm⁻³ eingesetzt. Bei Kurve 3.8d) wurde eine hohe Kompensation mit einer Donatordichte von

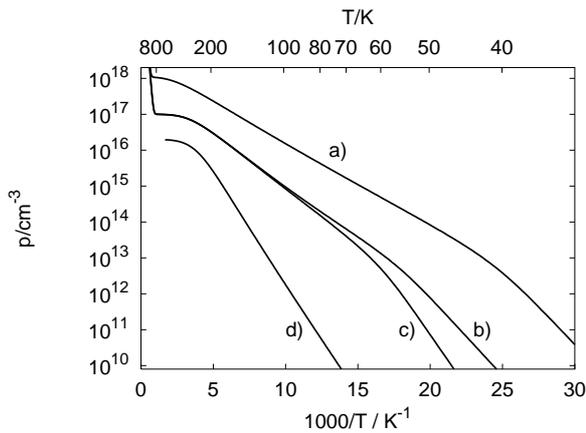


Abbildung 3.8: Zwei Akzeptor- und ein Donatorniveau – Variation des flachen Akzeptors

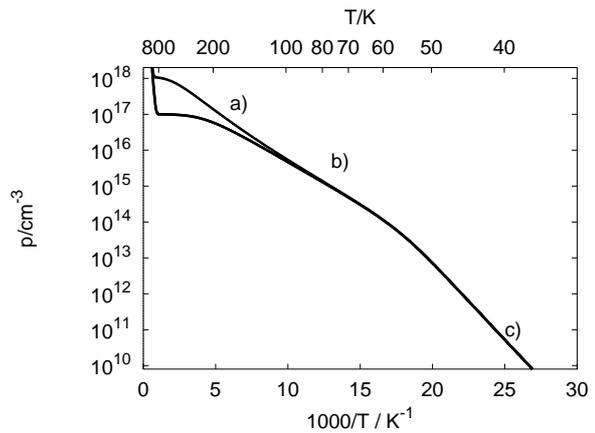


Abbildung 3.9: Zwei Akzeptor- und ein Donatorniveau – Variation des tiefen Akzeptors

$N_D = 8 \cdot 10^{16} \text{ cm}^{-3}$ angenommen. Dadurch werden sämtliche Einflüsse des flachen Akzeptors vollständig verdeckt.

Solange die Akzeptoren jeder für sich stärker ionisiert sind, als die Anzahldichte der Donatoren, gelten die Kriterien wie in Abbildung 3.7. Ist der flachere Akzeptor häufiger als der tiefere, so bestimmt ersterer das Verhalten wie in Kurve 3.8a). Friert er unter die Dichte des Donators aus, verdoppelt sich die Steigung entsprechend seiner Anregungsenergie.

Friert zunächst der tiefe Akzeptor A_1 unter die Dichte des Donators aus, wechselt die Steigung vom Wert für diesen Akzeptor ohne Kompensation $\frac{1}{2}E_{A1}$ zum Wert für den flachen Akzeptor, der je nach Konzentration des Donators $\frac{1}{2}E_{A2}$ oder, wie im gewählten Beispiel (Kurve 3.8b), E_{A2} beträgt.

Friert der flachere unter die Dichte des Donators aus, bevor der tiefere diesen Wert unterschreitet, nimmt die Steigung nur Werte an, die dem tieferen entsprechen (Kurve 3.8c). Der Akzeptor mit der niedrigeren Aktivierungsenergie bleibt in diesem Fall bei allen Temperaturen für die Hall-Messung unsichtbar.

In Abbildung 3.9 wurde die Konzentration des tiefen Akzeptors variiert. Solange er sel-

tener ist als der flache, wird er von diesem vollständig verdeckt (3.9b und c); nur wenn er häufiger ist, bestimmt er die Steigung bei hohen Temperaturen (3.9a), überläßt das Regime aber mit abnehmender Temperatur wieder dem flacheren.

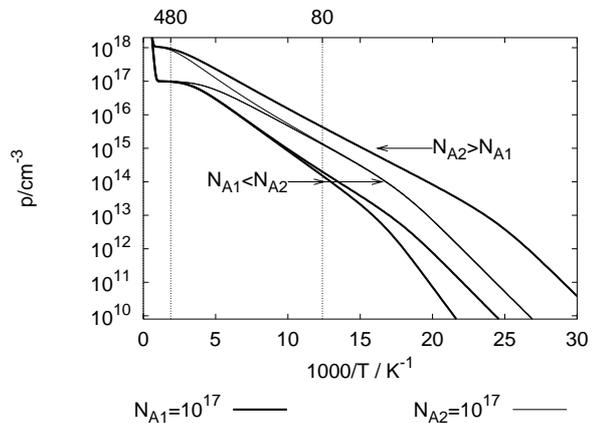


Abbildung 3.10: Löcherkonzentration bei Kompensation und zwei Akzeptoren. Zusammenfassung der Abbildungen 3.8 (—) und 3.9(---)

Abbildung 3.10 gibt im direkten Vergleich die Löcherkonzentrationen aus den Abbildungen 3.8 und 3.9 wieder. Sie veranschau-

licht die Schwierigkeit, aus der einen Meßgröße $R_H(T)$ die zugrundeliegenden Dotandenverhältnisse zu bestimmen, insbesondere, wenn sich der zugängliche Temperaturbereich nur zwischen 80 und 400 K erstreckt, wie bei den in Kapitel 6 vorgestellten Messungen.

Es können damit jedoch zumindest Aussagen zur prinzipiellen Nachweisbarkeit einzelner Defektenergien und -Konzentrationen angegeben werden:

So legt die Konzentration des kompensierenden Donators die untere Nachweisgrenze $N_A > N_D$ für alle Akzeptoren fest. Seine Konzentration läßt sich nur in günstigen Fällen aus dem Ort der Steigungsänderung der Löcherkonzentration ablesen.

Untereinander verdeckt der jeweils flachere den tieferen Akzeptor $E_{A2} < E_{A1}$ im gesamten Temperaturbereich, wenn seine Anzahldichte größer als die des steileren ist $N_{A2} > N_{A1}$.

Ist dagegen der Akzeptor mit der größeren Anregungsenergie häufiger, beherrscht er das Verhalten bei hohen Temperaturen. Da er schneller ausfriert, als ein Akzeptor mit niedrigerer Energie, gibt es immer eine Temperatur, bei der letzterer das Regime übernimmt.

3.2.4 Beweglichkeit und Streumechanismen

Aus dem Temperaturverhalten der Beweglichkeit kann qualitativ auf unterschiedliche Streumechanismen im Halbleiter zurückgeschlossen werden. [104] nennt hier insgesamt drei grundsätzlich unterschiedliche Verhalten:

1. Bei teilweise kompensierten Halbleitern ist, wie die vorstehende Diskussion gezeigt hat, ein Teil der Dotanden *unabhängig* von der Temperatur immer ionisiert, da die von ihnen freigesetzten Ladungsträger sofort von den komplementären Dotanden eingefangen werden. Bei niedrigen Temperatu-

ren dominiert dann die Streuung an diesen ionisierten Störstellen und mindert die Beweglichkeit der verbleibenden freien Ladungsträger. Es kann theoretisch gezeigt werden, daß sie dann wie $T^{3/2}$ mit der Temperatur steigen sollte.

2. Umgekehrt weist eine mit der Temperatur wie $T^{-3/2}$ fallende Beweglichkeit auf die Streuung an akustischen Phononen hin, die mit steigender Temperatur zunehmend angeregt werden.
3. In polaren Halbleitern, wie CuGaSe₂ spielt darüber hinaus auch die Streuung an optischen Phononen eine Rolle, die mit einem Temperaturverhalten von $T^{1/2}$ verknüpft ist.

Schon bei vergleichsweise idealen Halbleitern wie Silizium treten jedoch in der Praxis bedingt durch weitere nicht berücksichtigte Streumechanismen [104] davon abweichende, betragsmäßig meist größere Exponenten auf [39], so daß in komplexeren Materialien wie den Chalkopyriten allenfalls die Größenordnung und das Vorzeichen als Hinweis auf den dominierenden Mechanismus gedeutet werden kann.

3.2.5 Störbandleitung

Bei hohen Akzeptorkonzentrationen kombiniert mit starker Kompensation $> 50\%$ kann mit sinkender Temperatur statt eines weiteren Abfalls eine Sättigung der Leitfähigkeit beobachtet werden. Gleichzeitig stagniert der Abfall der Ladungsträgerdichte oder sie steigt scheinbar wieder an. Da dieses Phänomen bei den in dieser Arbeit untersuchten Proben nur bei genau zwei der quasi stöchiometrischen Proben aufgetreten ist, konnte seine Ursache nicht weiter ergründet werden. Dieser Abschnitt ist daher als Plausibilitätsbetrachtung anzusehen.

Zur Erklärung einer mit der Temperatur verlangsamtten Leitfähigkeitsabnahme bei Halbleitern wird von einer Akzeptordichte ausgegangen, die zur Ausbildung eines energetisch schmalen Subbandes nur mehr schwach lokalisierter Lochzustände führt, in denen die Leitung, ähnlich wie im Valenzband, durch Drift- und Diffusionsvorgänge zu beschreiben ist, die jedoch keine Aktivierungsenergie zur Freisetzung der Ladungsträger benötigt. Die effektive Masse der Ladungsträger in diesem Störband ist um ein Vielfaches höher als die in den Bändern [94].

Nach dem Wasserstoffmodell für Akzeptoren kann ihre Ausdehnung mit der effektiven Masse der Löcher m_h und der Dielektrizitätszahl des Halbleiters ϵ zu

$$r = \frac{\epsilon\epsilon_0 h^2}{\pi m_h e^2} = 0.5292 \text{ \AA} \frac{\epsilon}{m_h^*} \quad (3.29)$$

berechnet werden [27]. Für CuGaSe_2 werden Werte von $\epsilon \approx 11$ und $m_h \approx m_0$ berichtet³ [59], womit sich der kritische Abstand zweier Akzeptoren zu etwa 6.4 \AA ergibt. Damit ist er nur etwas größer als die kubische Gitterkonstante $a = 5.62 \text{ \AA}$ des CuGaSe_2 (Abschnitt 2.1)[61]. Bei einer Defektdichte von etwa 10 % der Gitterplätze würden sich also die Defekte gerade alle berühren, da in einem Volumen von $(5.3 \text{ \AA})^3 \approx 8$ Atome enthalten sind (s.a. Abschnitt 2.1). Dies würde zum Beispiel mit einem Galliumüberschuß, der nach dem Phasendiagramm 2.4 noch zu einphasigem Material führt, erreichbar erscheinen.

Die Beweglichkeit μ nimmt gleichzeitig mit der Temperatur ab, da immer mehr Ladungsträger in diesem Störband gefangen werden, wobei ein Temperaturverlauf ähnlich wie bei der Streuung an ionisierten Störstellen (s.o.) beobachtet wird.

Die Hall-Spannung fällt ebenso drastisch, da die Bewegungsfreiheit der Ladungsträger im

Störband stark eingeschränkt ist und sie nicht mehr beliebig auf die Lorenzkraft reagieren können, was eine scheinbare Erhöhung der Ladungsträgerdichte mit fallender Temperatur bewirkt. Ladungsträgerdichte und Leitfähigkeit sind nicht mehr gekoppelt, vielmehr wird die Leitfähigkeit von der Defektdichte bestimmt.

³Siehe auch Fußnote auf Seite 40

3.3 Die Kapazität von Verarmungszonen

Treffen in einem Halbleiter unterschiedlich dotierte Bereiche aufeinander oder gar ein Halbleiter mit einem Metall zusammen, setzt durch die unterschiedlichen Ladungsträgerkonzentrationen ein Diffusionsstrom ein, der diese auszugleichen versucht. Die Ladungsverchiebung führt zur Ausbildung von *Raumladungszonen*, die dem Diffusionsstrom einen Driftstrom im elektrischen Feld (3.3) entgegensetzen. Dies sind Bereiche mit verminderter oder sogar verschwindender Konzentration freier Ladungsträger. Elektrisch gesehen wirken diese Bereiche isolierend und bilden zusammen mit den angrenzenden leitfähigen Abschnitten einen Plattenkondensator. Aus dessen Kapazität läßt sich auf die Ladungsträgerdichte in der betreffenden Schicht zurückschließen, was in Ergänzung zu der in Abschnitt 3.2.1 vorgestellten Hall-Messung eine weitere Charakterisierungsmethode zur Bestimmung der elektrischen Eigenschaften von Halbleitern darstellt.

Im Folgenden wird zunächst die zur Messung der Kapazität als Funktion der Temperatur und Frequenz benötigte Apparatur beschrieben, die im Rahmen der vorliegenden Arbeit aufgebaut wurde. Anschließend werden die grundlegenden Beziehungen zur Auswertung der Kapazitätsmessungen diskutiert und auf Besonderheiten bei der Betrachtung von Solarzellen aus Chalkopyriten eingegangen.

3.3.1 Meßaufbau

Im Rahmen der vorliegenden Arbeit wurde ein Aufbau zur frequenz- und temperaturabhängigen Kapazitätsmessung zusammengestellt und in Betrieb genommen sowie eine Auswertungsstrategie entwickelt.

Dazu wurden folgende Geräte beschafft: Eine

Admittanzmeßbrücke HP4192 A, ein stickstoffgekühlter Kryostat mit Temperaturregelung und Armaturen zur Flüssigstickstoffzuführung, eine für Hochfrequenzsignale geeignete elektrische Verkabelung bis zur Probe einschließlich geeigneter Kontakthalter und ein fernabfragefähiges Voltmeter Keithley 2000.

Abbildung 3.11 gibt einen schematischen Überblick über die Komponenten und den Aufbau.

Die Admittanzmeßbrücke und das Voltmeter werden über einen GPIB-Bus durch einen Rechner mit einem in LabView⁴ geschriebenen Programm gesteuert und ausgelesen. Er übernimmt ebenfalls die Programmierung des autarken Temperaturreglers über dessen serielle RS232-Schnittstelle.

Zur Kühlung wird flüssiger Stickstoff in einem 50 l-Vorratsgefäß bereitgestellt und mit einer Membranpumpe durch den Kühlblock des Kryostaten gesaugt, auf den die Probe thermisch leitend montiert wird. Die Temperaturregelung erfolgt mit einem PID-Regler⁵ zweistufig einerseits durch elektrisches Heizen gegen die Dauerkühlung mit flüssigem Stickstoff, andererseits durch zeitweise Überbrückung des Stickstoffdrosselventils vor der Membranpumpe und damit Erhöhung des Stickstoffstroms und der Kühlleistung. Je nach vorgewählter Drosselung läßt sich dadurch die Temperatur bis zu ± 0.5 K stabil halten.

Die Meßbrücke wird über vier getrennt abgeschirmte Kabel und feinfühliges Kontaktnadelhalter mit der Probe verbunden. Je ein Paar ist dabei für die Zuführung des Meßstroms über die Vorder- und Rückseitenkontakte der Solarzelle und für den Abgriff der Spannung vorgesehen. Das Meßsignal läuft dabei zur Störsignalunterdrückung gegensinnig durch die Abschirmung, die deshalb nach der Spezifikation

⁴National Instruments

⁵engl. Proportional Integral Differential

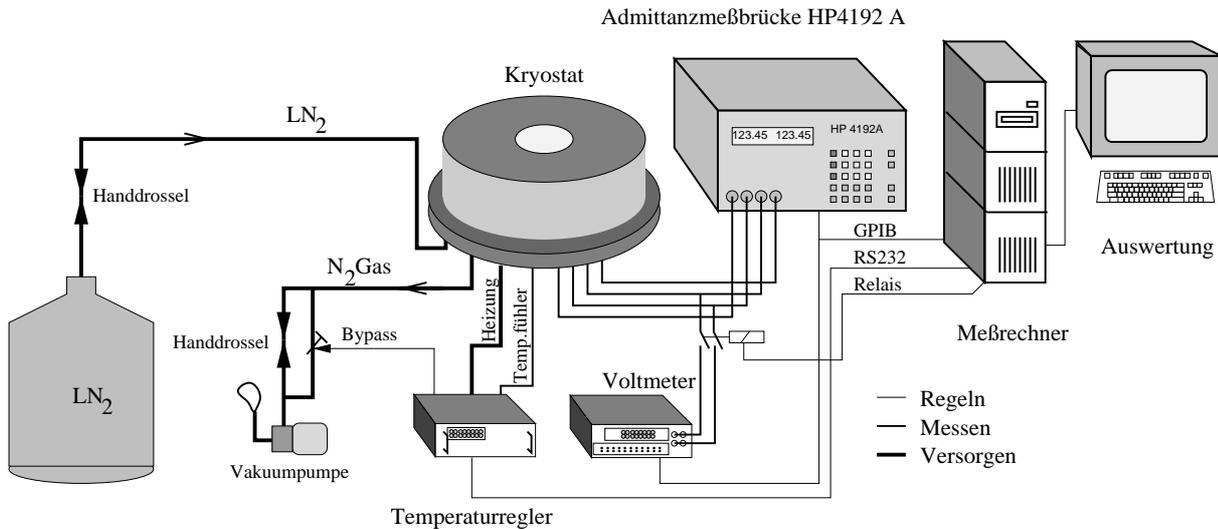


Abbildung 3.11: Schematischer Aufbau des Admittanzmeßplatzes

des Gerätes ausschließlich am probenseitigen Ende aller vier Kabel miteinander, aber nicht mit Masse verbunden sein muß.

Der zur Kapazitätsmessung angelegte Wechselstrom muß möglichst groß sein, um ein möglichst großes Meßsignal und damit eine ausreichende Auflösung des Ausgabewertes zu gewährleisten, die vom Gerät selbstständig begrenzt wird, er muß jedoch noch so klein sein, daß die Diodenkennlinie für die aus dem Meßstrom resultierende Spannungsdifferenz als näherungsweise linear betrachtet werden kann. Ansonsten würde sich auch ihre Gleichstromleitfähigkeit im Takt des Wechselstroms ändern, was fälschlicherweise als Wechselstromleitfähigkeit interpretiert würde. Als Kompromiß hat sich eine Wechselspannung von ± 40 mV Spitze-Spitze bewährt [35].

Zur spannungsabhängigen Messung von Bauteilen besitzt die Meßbrücke eine integrierte Spannungsquelle. Da sie allerdings einen Innenwiderstand von ca. $1\text{ k}\Omega$ besitzt⁶ bricht die

⁶weshalb im Grund *nicht* von einer Spannungsquelle gesprochen werden kann

Spannung bei den Solarzellen in Vorwärtspolung schon vor 1 V erheblich ein, was vom Gerät nicht registriert werden kann. Deshalb ist eine unabhängige Kontrolle der Spannung durch ein externes Voltmeter nötig. Dieses darf sich jedoch während der eigentlichen Admittanzmessung *nicht* im Meßstromkreis befinden, da seine Impedanz um mehrere Größenordnungen über der des Bauteils liegt, sie also völlig überdecken würde. Zu diesem Zweck wurden Hochfrequenzrelais in den Meßkreis eingeschleift, die vor jeder Messung vom Rechner gesteuert das Voltmeter kurz an die Probe anschließen, bis eine Spannungsmessung erfolgt ist und sogleich wieder wegschalten, bevor die Admittanzmessung ausgelöst wird.

3.3.2 Meßablauf

Dabei wurden vor allem drei Varianten verwendet:

1. Vorspannungsabhängige Kapazitätsmessung bei fester Frequenz (standardmäßig 100 kHz) und Raumtemperatur zur Ermittlung des Ladungsträgerprofils nach

Gleichung (3.47) und als Funktionskontrolle vor einer temperaturabhängigen Messung. Die Vorspannung wurde in einem Bereich von -1 bis 1 V in Schritten von 50 mV variiert.

Zur Kontrolle der tatsächlich am Bauteil anliegenden Vorspannung wurde vor jeder Einzelmessung computergesteuert ein Voltmeter über HF-Relais an die Spannungskabel angeschaltet und der Wert mitprotokolliert. Auf eine Nachregelung der Spannung wurde zugunsten einer verkürzten Meßzeit verzichtet.

2. Temperatur- und frequenzabhängige Kapazitätsmessung ohne Vorspannung zur Auswertung der Störstellendichte. Der Frequenzbereich lag meist zwischen 100 Hz und 10 MHz, logarithmisch unterteilt in 100 Intervalle, die Temperatur wurde von -180°C bis 100°C in 5-K-Schritten variiert.
3. Temperatur- und spannungsabhängige Kapazitätsmessung bei ausgewählten Frequenzen zur Untersuchung grenzflächen-naher Zustände.

Um eine möglichst hohe Auflösung zu erhalten, wurde bei maximaler Integrationszeit von 3 s je Punkt gemessen. Auf eine zusätzliche Akkumulation und Mittelung über mehrere Meßpunkte wurde zugunsten eines schnelleren Meßablaufs und damit der weitestmöglichen Schonung der Proben verzichtet. Auch so dauerte eine Messung unter Punkt 2. etwa 5 Stunden.

3.3.3 Kapazität

Die Kapazität ist über die Wechselstromleitfähigkeit oder Admittanz des betreffenden Bauteils bestimmbar. Admittanz bedeutet den

komplexen Leitwert

$$Y = \frac{I}{U} \quad (3.30)$$

eines Bauteils. Einige Beispiele sind in Tabelle 3.2 zusammengestellt.

Zur Ermittlung der elektrischen Kenngrößen eines Bauteils genügt dann genau eine Messung der Admittanz, wenn es sowohl frequenz- als auch spannungsunabhängig arbeitet. Dies gilt zwar für die in Tabelle 3.2 aufgeführten Bauteile in sehr guter Näherung. Eine Solarzelle stellt jedoch ein komplexes Netzwerk aus kapazitiven und resistiven Komponenten dar, das einen charakteristischen Frequenzgang aufweist.

Darüber hinaus ändert sich zum einen ihre Gleichstromleitfähigkeit je nach angelegter Spannung um viele Größenordnungen, zum anderen wird durch Variieren dieser Spannung die elektronische Geometrie beträchtlich verändert, was sich zusätzlich auf die Wechselstromleitfähigkeit auswirkt.

Somit müssen für eine vollständige Beschreibung der elektrischen Eigenschaften einer Solarzelle sowohl die angelegte Spannung als auch die Frequenz variiert werden. Zusätzliche Informationen können aus temperaturabhängigen Messungen gewonnen werden.

Um die Kapazität eines Bauteils wie zum Beispiel einer Solarzelle zu bestimmen, wird ein Strom durch dieses getrieben, der sich aus einem Wechselstromanteil geringer, aber fester Amplitude und variabler Frequenz sowie einem Gleichstromanteil⁷ j variabler Amplitude zusammensetzt.

$$j(t) = \bar{j} + \hat{j} \sin \omega t; \quad \bar{j} \gg \hat{j} \quad (3.31)$$

⁷Um im Folgenden Verwechslungen mit der komplexen Einheit $i = \sqrt{-1}$ zu vermeiden, wird die Wechselstromstärke entgegen der Konvention mit j bezeichnet, da der Unterschied zur *Stromdichte* weniger fundamental ist.

Bauteil	Leitwert	Phase ϕ [Grad]	Komplex
Ohmscher Widerstand	$G = \frac{1}{R} = \frac{I^2}{P}$	0	$Y = \frac{1}{R}$
Kondensator	$B = \omega C$	-90	$Y = i\omega C$
Spule	$B = \frac{1}{\omega L}$	90	$Y = -\frac{i}{\omega L}$

Tabelle 3.2: Admittanz einfacher elektronischer Bauteile

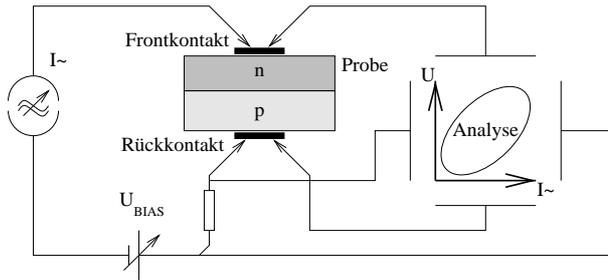


Abbildung 3.12: Meßbrücke zur Bestimmung der Admittanz

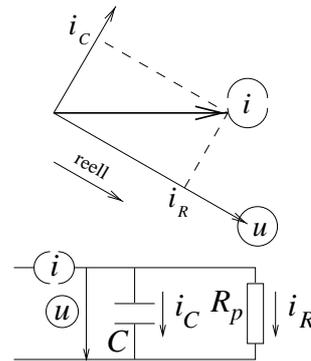


Abbildung 3.14: Parallelschaltung

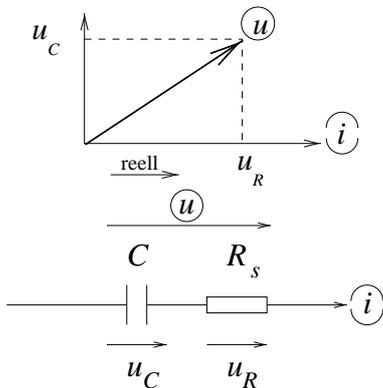


Abbildung 3.13: Reihenschaltung

Die dabei am Bauteil auftretende Wechselspannung

$$u(t) = \hat{u} \sin(\omega t + \phi) \quad (3.32)$$

wird phasenrichtig gegen den Strom gemessen (Abbildung 3.12). Aus den gemessenen Verhältnis von j (3.31) zu u (3.32) läßt sich unter der Annahme eines bestimmten Schaltungstyps

auf die Aufteilung der Spannung (Reihenschaltung, Abbildung 3.13) oder Ströme (Parallelschaltung, Abbildung 3.14) schließen, und damit

- entweder bei einer Parallelschaltung mit der reellen Achse in u -Richtung und

$$j = j_R + i j_C$$

(Abb. 3.14) die komplexe Admittanz

$$Y = G + i \cdot B \quad (3.33)$$

in deren Komponenten

$$G = \frac{1}{R} = \frac{j_R}{u} = \frac{\text{Re}(j)}{u} \quad (3.34)$$

$$B = \omega C = \frac{j_C}{u} = \frac{\text{Im}(j)}{u} \quad (3.35)$$

- oder bei einer Reihenschaltung mit der reellen Achse in j -Richtung und

$$u = u_R + i u_C$$

(Abb. 3.13) die komplexe Impedanz

$$Z = R + i \cdot X \quad (3.36)$$

in den Komponenten

$$R = \frac{u_R}{i} = \frac{\text{Re}(u)}{i} \quad (3.37)$$

$$X = \frac{1}{\omega C} = \frac{u_C}{i} = \frac{\text{Im}(u)}{i} \quad (3.38)$$

berechnen. Die entsprechende Einstellung muß an der HP4192A-Meßbrücke vorgenommen werden. Sie bietet für Messungen an Solarzellen ohne angelegte Vorwärtsspannung das Parallelersatzschaltbild an, da dann die Diodenstrecke hochohmig ist und ihr gegenüber der Serienwiderstand vernachlässigt werden kann⁸. Für spannungsabhängige Messungen, bei denen die Diode bis in die Vorwärts- oder Durchlaßrichtung betrieben wird, wird die Wahl des Ersatzschaltbildes kritisch. Insbesondere im Übergangsbereich, wenn der Serienwiderstand vergleichbar mit dem Widerstand der Diodenstrecke wird ($R_s \approx R_d$), ist ohne absolute Kenntnis wenigstens eines der beiden Widerstandswerte keine konsistente Aussage über die tatsächliche Kapazität des Bauteils möglich. Eine ausführliche Diskussion der sich damit stellenden Probleme wird in Abschnitt 3.3.6 gegeben.

Da das Gerät *ausschließlich* diese vier Größen zur Verfügung stellt, ist die *Wahl* des Ersatzschaltbildes eng mit der weiteren *Auswertung* verknüpft:

So lautet der Phasenwinkel

$$\tan \phi = \frac{B}{G} = \frac{X}{R}, \quad (3.39)$$

⁸Tatsächlich wird bei einigen Zelltypen der Serienwiderstand zumindest bei tiefen Temperaturen schon ohne Vorwärtsspannung bemerkbar, weshalb sein Einfluß gegebenenfalls besonders behandelt werden muß

und die Kapazität

$$C = \frac{B}{\omega} = \frac{B}{2\pi\nu} \text{ bei Parallel-} \quad (3.40)$$

$$\text{und} = -\frac{1}{\omega X} = -\frac{1}{2\pi\nu X} \quad (3.41)$$

bei Serienschaltung.

3.3.4 Schottky-Kontakt

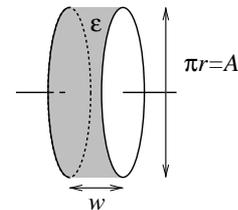


Abbildung 3.15: Zur Kapazität einfacher Geometrien

Für einfache geometrische Verhältnisse läßt sich die Kapazität direkt aus den geometrischen Gegebenheiten ablesen: Bei einem aus zwei gegenüberliegenden, voneinander durch ein Dielektrikum der relativen Dielektrizitätszahl ϵ getrennten Platten gebildeten *Plattenkondensator* wie in Abbildung 3.15 gilt für dessen Kapazität

$$C = \epsilon \epsilon_0 \frac{A}{w}. \quad (3.42)$$

Ein idealer Metall-Halbleiter Übergang, der aufgrund der Bandverbiegung nach Schottky (Abbildung 3.16 a) sperrendes Verhalten zeigt, kommt dem Fall des Plattenkondensators (3.42) sehr nahe, wenn darin w mit der Raumladungsweite

$$w = \sqrt{\frac{2\epsilon\epsilon_0(V_{bi} - V)}{e} \frac{1}{N_A^-}} \quad (3.43)$$

identifiziert wird. Darin ist N_A^- die Anzahldichte der ionisierten Akzeptoren und V_{bi} die eingebaute Potentialbarriere, die durch die Differenz

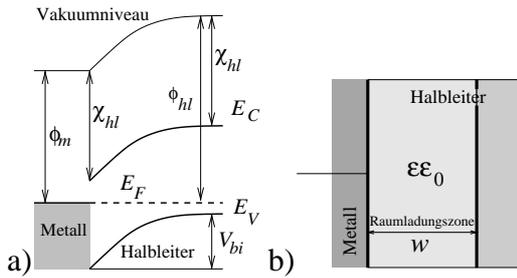


Abbildung 3.16: Schichtstruktur eines Metall-p-Halbleiterkontakts und entsprechende Bandverbiegung nach Schottky

Es ist E_V Valenzband-, E_C Leitungsbandkante, E_F Fermienergie. Die Bandverbiegung wird bewirkt durch die unterschiedlichen Lagen der Austrittsarbeit ϕ_m des Metalls und des Halbleiters ϕ_{hl} . Die resultierende Barrierenhöhe für Löcher ist V_{bi} . Die Breite der Verarmungszone w (3.43) bildet eine Kapazität, die sich nach (3.42) berechnen läßt.

zwischen Austrittsarbeit des Metalls ϕ_m und der des Halbleiters ϕ_{hl} festgelegt wird.

Damit wird die inverse flächenbezogene Kapazität⁹ der Raumladungszone

$$C^{-2} = \frac{2}{eN_A^- \epsilon\epsilon_0} (V_{bi} - V). \quad (3.44)$$

In p-dotierten Halbleitern gilt außerhalb der Raumladungszone in guter Näherung $n \approx 0$ und $N_A^- \gg N_D^+ \approx 0$, weswegen mit der Ladungsneutralität (3.13)

$$n + N_A^- = p + N_D^+,$$

am Rand der Raumladungszone $N_A^- \approx p$ geschrieben werden kann und im Folgenden synonym verwendet wird.

Eine lineare Regression $C^{-2} = aV + b$ der sogenannten „Mott-Schottky-Auftragung“ $C^{-2}(V)$ liefert über (3.44) sowohl Aufschluß

⁹Eine gemessene Kapazität C' muß deshalb durch die effektive Fläche A dividiert werden. Im Folgenden wird C stets auf die Einheitsfläche 1 m^2 bezogen sein.

über die effektive Ladungsträgerkonzentration

$$p = -\frac{2}{e\epsilon\epsilon_0} \frac{1}{a} \quad (3.45)$$

als auch aus der Extrapolation für $C^{-2} \rightarrow 0$ über das eingebaute Potential

$$V_{bi} = -\frac{b}{a} = \frac{e\epsilon\epsilon_0 p}{2} b. \quad (3.46)$$

Das Differential

$$p = \frac{2}{e\epsilon\epsilon_0} \left(\frac{d(C)^{-2}}{dV} \right)^{-1}, \quad (3.47)$$

beschreibt die Ladungsträgerkonzentration am Rande der Verarmungszone in Abhängigkeit von der angelegten Spannung.

Durch die Messung von C steht über (3.42) auch die jeweilige Weite der Verarmungszone zur Verfügung, so daß sich

$$p(V) = p(C) = p(w) = N_A^-(w) \quad (3.48)$$

in Abhängigkeit von w schreiben läßt: Die $C(V)$ -Messung ermöglicht daher ein $N_A(w)$ Profil. Aus (3.44) läßt sich zusätzlich auch eine tiefenabhängige Aussage über das eingebaute Potential

$$V_{bi} = \frac{p\epsilon\epsilon_0 e}{2} C^{-2} + V \quad (3.49)$$

machen. Sie liefert eine Obergrenze für das tatsächlich in der Verarmungszone abfallende Potential.

3.3.5 p-n-Übergang

Die vorstehende Auswertung ist strenggenommen nur gültig für Schottky-Kontakte, bei denen die Raumladungszone sich einseitig nur im Halbleitermaterial ausdehnt. Bis heute ist jedoch kein Material bekannt, das auf CuGaSe_2 einen stabilen Schottky-Kontakt ausbilden würde. Praktisch alle Metalle bilden

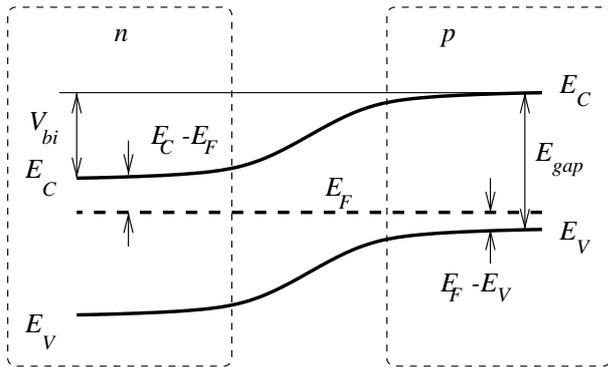


Abbildung 3.17: Das eingebaute Potential

einen wenn auch „schlechten“ aber Ohm’schen Kontakt. Deshalb sind Messungen an Raumladungseffekten in Chalkopyriten grundsätzlich auf Heterostrukturen angewiesen, die durch einen metallurgischen p-n-Übergang eine Raumladungszone in das Material implementieren.

Für einen homogenen p-n-Übergang gilt in Abwandlung von Gleichung (3.43)

$$w = \sqrt{\frac{2\epsilon\epsilon_0(V_{bi} - V)}{e} \frac{N_A^- + N_D^+}{N_A^- N_D^+}}, \quad (3.50)$$

wenn die Dotierkonzentrationen $N_A \approx N_D$ vergleichbar sind. Das eingebaute Potential

$$V_{bi} = (E_C - E_V) - (E_F - E_V)_p - (E_C - E_F)_n. \quad (3.51)$$

entspricht, wie in Abbildung 3.17 skizziert, der Bandlücke $E_{gap} = E_C - E_V$ abzüglich des Abstandes des Ferminiveaus von den Bandkanten in den neutralen Gebieten der angrenzenden Halbleiter, in denen die Bänder praktisch flach und parallel zum Ferminiveau verlaufen [104].

Diese Abstände hängen wiederum über die Gleichungen (3.14) bis (3.17) von den Konzentrationen der Akzeptoren N_A und Donatoren N_D in den Halbleitern beiderseits des p-n-Überganges und von der Temperatur ab.

In dem bei der Solarzellenherstellung verwendeten hochdotierten ZnO kann jedoch in guter Näherung ($E_C - E_F \approx 0$) gesetzt werden, womit es einem Schottky-Kontakt-Material ähnlich ist: Seine Dotierung im Bereich einiger 10^{20}cm^{-3} liegt um zwei Größenordnungen über der der höchstdotierten Chalkopyrite mit einigen 10^{18}cm^{-3} wie in Kapitel 6 noch gezeigt wird. Entsprechend verhält sich die Ausdehnung der Raumladungszonen in den beiden Materialien umgekehrt proportional, kann also im ZnO getrost vernachlässigt werden. Insbesondere wird sie sich durch Anlegen einer Spannung V praktisch nur im niedriger dotierten Material verschieben, so daß nach Gleichung (3.47) die Dotierung des Chalkopyrits am Rande der Verarmungszone bestimmt werden kann.

Eine Unsicherheit könnte noch das verwendete Puffermaterial darstellen, dessen elektrische Eigenschaften nur ungenau bekannt sind. Als *Einzelschicht* ist es undotiert und mithin semiisolierend. Die Wechselwirkung mit der Chalkopyritschicht und seine Rolle beim Stromtransport ist aktueller Forschungsgegenstand [25, 56]. Unter der *Annahme*, der Puffer sei nach wie vor undotiert, spielt er für die Kapazitätsmessung keine Rolle, da er dann von den umgebenden höherdotierten Schichten praktisch vollständig seiner Ladungsträger beraubt wird und nur die Dicke w der Raumladungszone geringfügig erhöht, also die Kapazität entsprechend (3.42) senkt.

3.3.6 Serienwiderstand

Neben der Sperrschichtkapazität liegt in einem p-n-Übergang die Diode mit ihrem exponentiell spannungsabhängigen Leitwert im Strompfad, der sich im Idealfall durch die Shockleygleichung [104]

$$j = j_0 \left(\exp \frac{eV}{kT} - 1 \right) \quad (3.52)$$

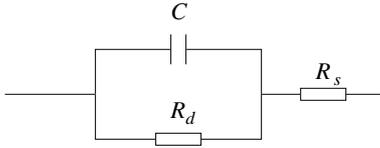


Abbildung 3.18: Ersatzschaltbild mit Serienwiderstand. Es muß beachtet werden, daß R_d den Parallelwiderstand *und* die Diodencharakteristik in Form eines spannungsabhängigen Widerstands gemäß (3.52) umfaßt, im Gegensatz zur üblichen Darstellung der realen Diode.

beschreiben läßt. Eine Betrachtung der Diodenkennlinien war nicht Gegenstand der vorliegenden Arbeit, ausführliche Diskussionen der Diodeigenschaften von Chalkopyritzellen findet sich zum Beispiel bei [32, 67, 106]. Für die folgende Diskussion ist alleine ihre spannungsabhängige Leitfähigkeit von Interesse.

Wie in Abschnitt 3.3.3 erwähnt, spielt der Serienwiderstand bei den Solarzellen dann eine Rolle, wenn die Diodenstrecke zunehmend leitfähig wird. Das dafür zunächst naheliegende Ersatzschaltbild ist in Abbildung 3.18 dargestellt. Der Gesamtwiderstand lautet dafür

$$Z = R_s + \frac{1}{\frac{1}{R_d} + i\omega C_d} \quad (3.53)$$

mit dem Serienwiderstand R_s , dem spannungsabhängigen Widerstand der Diode R_d , und der Sperrschichtkapazität C_d . Die *gemessene* Kapazität ist dann

$$C = \frac{B}{\omega} = \text{Im} \frac{1}{\left(R_s + \frac{1}{\frac{1}{R_d} + i\omega C_d} \right) \omega} \quad (3.54)$$

$$= \frac{C_d}{\left(\frac{R_s}{R_d} + 1 \right)^2 + (R_s \omega C_d)^2} \quad (3.55)$$

Der Ohmsche- oder Realteil der Leitfähigkeit

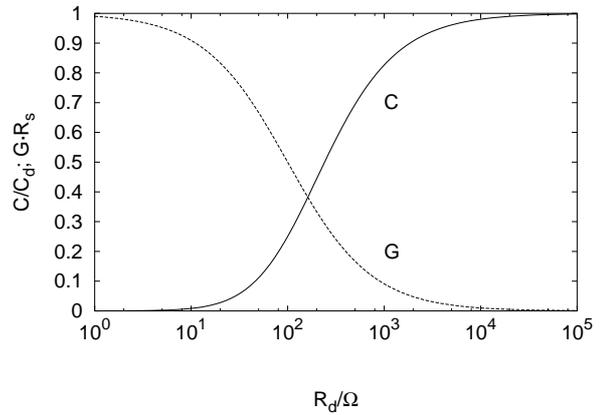


Abbildung 3.19: Kapazität und Leitwert des Schaltbildes 3.18 als Funktion des Widerstands R_d des p-n-Übergangs

lautet entsprechend

$$G = \frac{\frac{R_s}{R_d^2} + R_s(\omega C)^2 + \frac{1}{R_d}}{\left(\frac{R_s}{R_d} + 1 \right)^2 + (R_s \omega C_d)^2} \quad (3.56)$$

Mit anderen Worten: Die *gemessenen Größen* C und G sind frequenzabhängig. Dabei enthalten Real- und Imaginärteile dieselbe Information, da sie über die Kramers-Kronig-Relation miteinander verknüpft sind. In Abbildung 3.20 wird diese Komplementarität augenscheinlich: Hier sind die Verläufe von C und G für variierte R_s, R_d beziehungsweise C_d aufgetragen. Sowohl in der Kapazität als auch im Leitwert sind die charakteristischen Strukturen – Maxima oder Stufen – bei denselben Parameterkonstellationen zu finden.

Für Vorwärtsspannungen wird der p-n-Übergang der Solarzelle leitfähig, was nach der Shockleygleichung (3.52) einer drastischen Abnahme des Widerstands R_d entspricht. Demzufolge sinkt der Spannungsfall über R_d und C_d und verlagert sich auf den Serienwiderstand R_s , womit der imaginäre Anteil der Leitfähigkeit vermindert, der Realteil erhöht wird. Wie

3.3. DIE KAPAZITÄT VON VERARMUNGZONEN

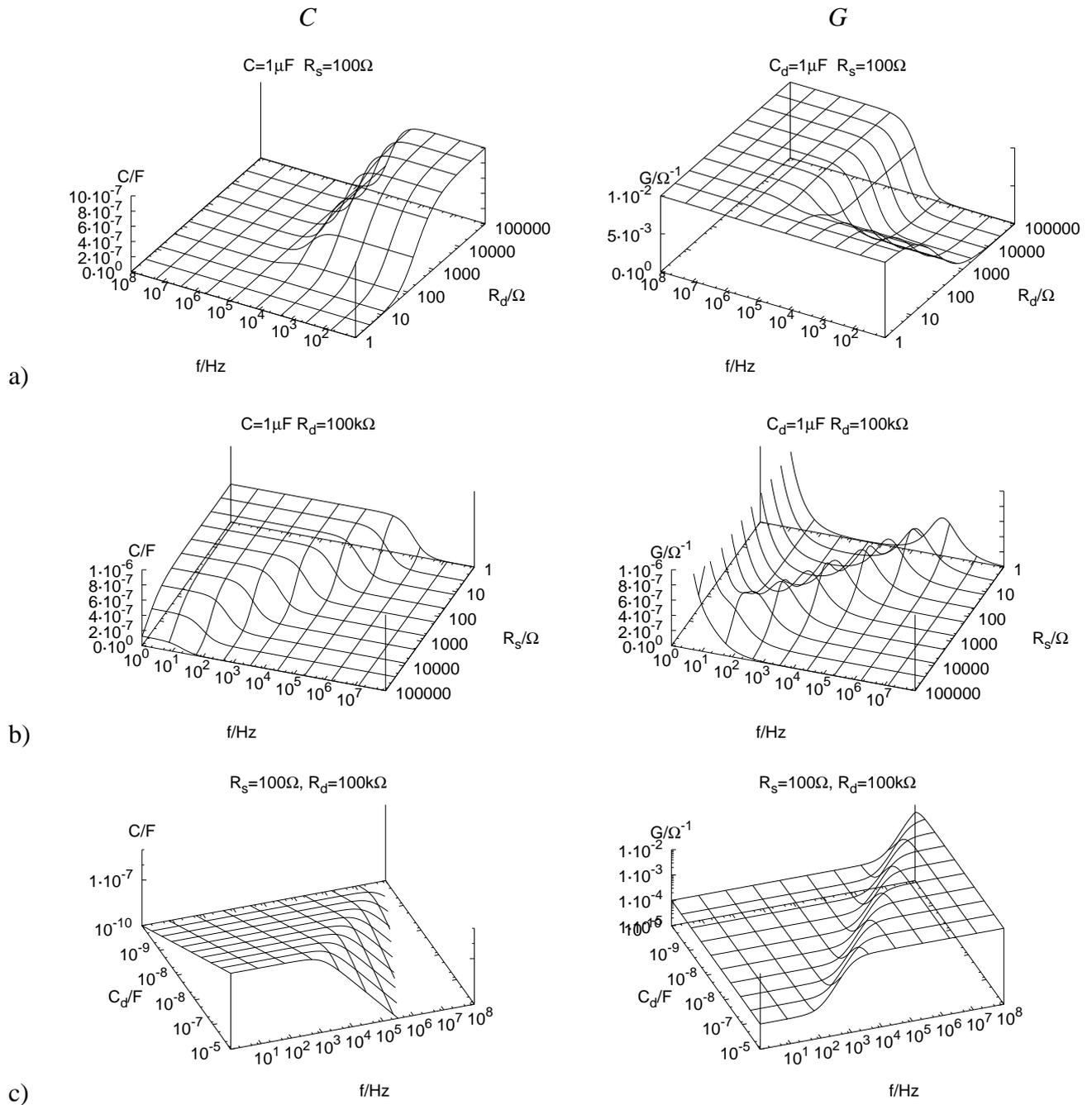


Abbildung 3.20: Effektive Kapazität und Leitfähigkeit des Schaltbildes 3.18 als Funktion von R_d , R_s und C_d nach (3.55) und (3.56). Zwei von drei Parametern wurden jeweils konstant $R_d = 100k\Omega$, $R_s = 100\Omega$ beziehungsweise $C_d = 1\mu F$ gehalten, der dritte variiert.

Abbildung 3.19 verdeutlicht, führt dies zu einer Abnahme der gemessenen Kapazität, da $\left(\frac{R_s}{R_d} + 1\right)^2$ groß gegen $(R_s \omega C_d)^2$ wird, und zu einem Anstieg des Leitwertes G . Es ist also zu erwarten, daß eine Auswertung nach (3.44) oder (3.45) für Vorwärtsspannungen, bei denen der Realteil der Leitfähigkeit groß wird und die gemessene Kapazität verringert, ungünstig ist. Eine Korrektur ist durch die Berücksichtigung des Realteils der komplexen Admittanz zumindest qualitativ möglich: Mit dem Verlustfaktor

$$d = \frac{1}{R\omega C} = \frac{G}{B}, \quad (3.57)$$

der den Anteil des in Phase fließenden Stroms beschreibt, kann die nach dem Parallelschaltbild gemessene Kapazität C' in die tatsächlich in der Reihenschaltung von R_s und C_d auftretende

$$C_d = \frac{C'}{\left(1 + \frac{G}{B}\right)} \quad (3.58)$$

umgerechnet werden. Im Übergangsbereich $\frac{1}{\omega C_d} \approx R_s \approx R_d$ bleibt die Abweichung von der tatsächlichen Kapazität jedoch unvermeidlich. Eine vollständige Auswertung wäre nur durch Kenntnis der separaten Werte der beiden Widerstände möglich. Deren Auswertung gestaltet sich aufgrund der in der Realität oft deutlich komplizierteren und nur teilweise verstandenen Kennlinienverläufe als mindestens genauso schwierig [67].

Teilbild 3.20 b) ähnelt in seiner Form stark den tatsächlich gemessenen temperaturabhängigen Kapazitätsspektren, deren Ursprung im folgenden Abschnitt 3.4 beschrieben wird. Derartige Messungen sind in Kapitel 5, Abschnitt 5.3.2.3 zu finden.

Wird der Serienwiderstand als temperaturabhängig angenommen — was für das Bahngebiet der Absorber durch das Ausfrieren der

Ladungsträger nach (3.26) sicherlich nicht abwegig ist, wie die Leitfähigkeitsmessungen in Abschnitt 6.1.3.1 zeigen — ergäbe sich auch ohne eine Beteiligung irgendwelcher Störstellen allein durch die Veränderung des Serienwiderstandes dieses Bild. Allerdings erstreckt sich die gemessene Temperaturabhängigkeit des Leitwertes nur über wenige Größenordnungen, so daß sie alleine nicht für die gemessenen Kapazitätsspektren verantwortlich sein kann. Wohl aber zeigen vornehmlich CuInS₂-basierte Zellen (Kapitel 5, Abschnitt 5.3.2.3¹⁰), aber auch andere Zellen mit schwefelhaltigen Komponenten eine Kapazitätsstufe, die gerade wegen ihrer extrem geringen Temperaturabhängigkeit mit dem Ansatz eines temperaturabhängigen Widerstandes am einfachsten zu verstehen ist [99].

Messungen bei möglichst hohen Frequenzen sind speziell für die $C(V)$ -Auswertung wünschenswert, da zum einen der Rauschabstand mit der Frequenz größer wird, zum anderen, wie im nächsten Abschnitt ausgeführt, nur die freien Ladungsträger einen Beitrag zur Kapazität liefern. Zu hohen Frequenzen hin wird der Blindleitwert der Kapazität groß gegen den Leitwert aller übrigen Widerstände: Die gemessene Kapazität sinkt wie

$$C = \frac{1}{C_d(R_s \omega)^2} \quad (3.59)$$

und setzt damit der Meßfrequenz eine obere Schranke.

Eine weitere Schranke ergibt sich aus der dielektrischen Relaxationszeit

$$\tau = \frac{\epsilon \epsilon_0}{\sigma}, \quad (3.60)$$

wenn auch die freien Ladungsträger der angelegten Wechselspannung nicht mehr folgen

¹⁰CuInS₂-basierte Zellen waren nicht primärer Gegenstand der vorliegenden Arbeit.

können und der Halbleiter in seiner geometrischen Ausdehnung zum Isolator für den Wechselstrom wird. Bei den epitaktischen Proben mit einer Leitfähigkeit von mindestens 0.1 bis $10 \text{ } (\Omega\text{cm})^{-1}$ (Kapitel 6) tritt dieser Fall jedoch erst bei Frequenzen um 1 GHz und damit deutlich später ein, als bei vielen Zellen in Kapitel 5 beobachtet wird.

3.4 Verlustmechanismen

Neben den Ladungstransporteigenschaften der einzelnen Materialien, die in Abschnitt 3.2 diskutiert wurden, treten bedingt durch den Schichtaufbau in Heterostrukturen und Solarzellen weitere Effekte auf, die den Stromfluß beeinträchtigen können.

Die meisten hängen eng mit der *Rekombination* photogenerierter Ladungsträger an *Störstellen* zusammen. Der Begriff „Störstelle“ überschneidet sich in seiner Bedeutung mit dem des Dotanden: Sorgt letzterer dafür, daß überhaupt Ladungsträger für den Stromtransport zur Verfügung stehen, sind erstere für den Verlust derselben verantwortlich, indem sie als *Ladungsträgerfallen*¹¹ diese Ladungsträger einfangen. Können sie aus beiden Bändern sowohl Elektronen wie auch Löcher aufnehmen, fungieren sie darüber hinaus als *Rekombinationszentren* und können damit zum Beispiel den Photostrom teilweise kurzschließen.

Ein mögliches quantitatives Kriterium stellt die energetische Lage des *flachsten Majoritätsdotanden* dar: Alle *tieferen Defekte* können als *Störstelle* bezeichnet werden.

Es wird eine Vielzahl möglicher *Rekombinationspfade* diskutiert. Abbildung 3.21 skizziert einige davon:

1. Störstelle im Bahngebiet [88]

¹¹Meist mit engl. Trap bezeichnet

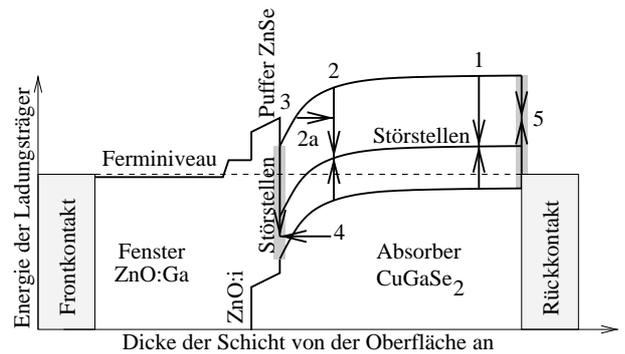


Abbildung 3.21: Rekombinationspfade in einer ZnO/ZnSe/CuGaSe₂-Solarzelle [6, 66]

2. Störstelle in der Verarmungszone, auch tunnelunterstützt (2a) [66, 88]
3. Grenzflächenzustände am Übergang zwischen Absorber und Puffer [107]
4. Tunneln in Grenzflächenzustände [66, 81]
5. Rekombination am Rückkontakt [104]

Während der letzte Punkt in der Praxis kaum eine Rolle spielt, da bis zum Rückkontakt die meisten der einfallenden Photonen bereits absorbiert wurden und demzufolge dort kaum Elektronen zur Rekombination anstehen, kann bei Heterostrukturen mit einer Absorberdicke von unter 500 nm sehr wohl Licht bis zum GaAs vordringen und dort Elektronen erzeugen, die an der Grenzfläche rekombinieren können.

3.4.1 Störstellen in der Raumladungszone

Kristalldefekte können Elektronen und Löcher an sich binden und ihnen damit den Aufenthalt in der Bandlücke ermöglichen. Liegt das betreffende chemische Potential energetisch in der Nähe solcher Störstellen, wie in Abbildung 3.22 durch den Kreis markiert, so können sie entsprechend durch Anlegen einer kleinen Wech-

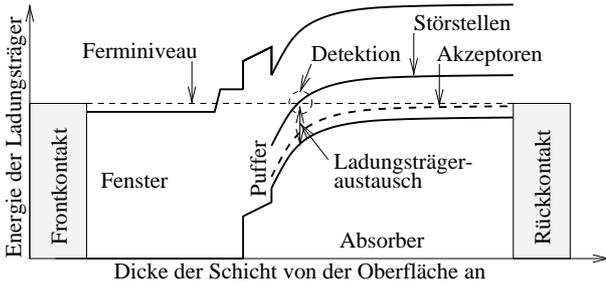


Abbildung 3.22: Schematischer Bandverlauf in einer CuGaSe₂-Solarzelle mit einem Störstellenniveau [6]

selspannung $u \ll 1$ V ge- und entladen werden und tragen somit zur Kapazität des Materials bei. Dabei ist diese Kapazität jedoch im Gegensatz zur rein geometrisch bedingten stark frequenzabhängig. Damit ist klar, daß von den oben diskutierten Rekombinationspfaden im wesentlichen nur der 2. mit der Kapazitätsspektroskopie direkt nachweisbar ist, da nur in der Raumladungszone das Ferminiveau einen nennenswerten Teil der Bandlücke und damit möglicher Störstellenniveaus überstreicht.

Nur wenn von einer homogenen Störstellenverteilung in der gesamten Absorbiertiefe ausgegangen wird, deckt diese Messung auch den 1. Fall ab.

Die Nachweisbarkeit von Grenzflächendefekten wird in Abschnitt 3.4.2 noch ausführlicher diskutiert.

3.4.1.1 Diskrete Störstellen

Diskrete Störstellen geringer Dichte $N_t \ll N_A$ tragen zur Kapazität wie [113]

$$C_t \propto \frac{\omega_0^2}{\omega_0^2 + \omega^2}, \quad (3.61)$$

bei, was in Abbildung 3.23 veranschaulicht ist.

Die charakteristische Frequenz der Störstelle ω_0 oder die Häufigkeit, mit der ein La-

dungsträger die Störstelle verlassen kann, ist dabei mit deren Anregungsenergie E_t durch die Boltzmannverteilung der Ladungsträgergeschwindigkeiten,

$$\omega_0 = 2N_{C,V} v_{th} \sum_{n,p} e^{-E_t/kT} \quad (3.62)$$

$$= \xi_0 T^2 e^{-E_t/kT} \quad (3.63)$$

verknüpft, die in diesem Zusammenhang häufig Arrheniusgesetz genannt wird.

Die *Emissionsrate* ξ_0 nimmt dabei sämtliche Parameter mit Ausnahme der expliziten Temperaturabhängigkeit der Zustandsdichte an der betreffenden Bandkante $N_{C,V}$ (3.18) bzw. (3.19) und der thermischen Geschwindigkeit

$$v_{th} = \sqrt{\frac{3kT}{m_{e,h}}} \quad (3.64)$$

auf. Bei einem typischen Einfangquerschnitt einer tiefen Störstelle von $\Sigma_p = 10^{-20}$ m² für Löcher [104], ist bei Raumtemperatur mit Werten von einigen 10^{10} s⁻¹K⁻² zu rechnen.

Im Prinzip kann umgekehrt der Einfangquerschnitt $\Sigma_{n,p}$ der Störstellen aus ξ_0 ermittelt werden, wenn $N_{C,V}$ und v_{th} hinreichend genau bekannt sind. Wie sich jedoch bei der praktischen Anwendung in Kapitel 5 herausstellen wird, ist ξ_0 nur dermaßen unsicher zu bestimmen, daß die damit ermittelten Streuquerschnitte keine sinnvolle Interpretationsgrundlage liefern können.

Wie Abbildung 3.23 zeigt, weist (3.61) eine charakteristische Stufe um ω_0 auf, die sich nach Gleichung (3.62) mit der Temperatur verschiebt.

Damit läßt sich aus einer Reihe frequenzabhängiger Messungen bei unterschiedlichen Temperaturen umgekehrt die Anregungsenergie E_t und der Vorfaktor ξ_0 bestimmen.

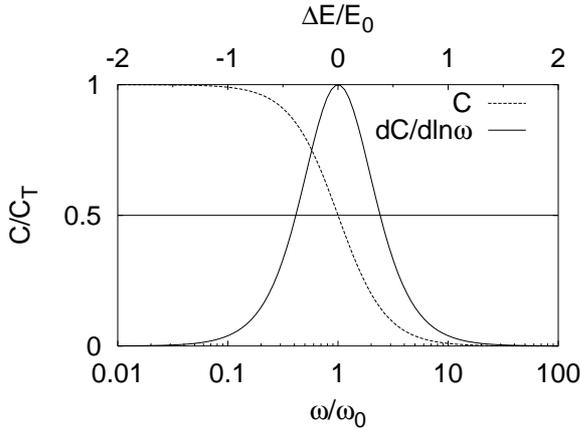


Abbildung 3.23: Beitrag einer diskreten Störstelle zur Kapazität der Raumladungszone, Gleichung (3.61)

3.4.1.2 Kontinuierliche Störstellen

Existiert in der Bandlücke eine energetisch und räumlich kontinuierliche Verteilung von Störstellen, so tragen sie je nach ihrer energetischen Lage bezüglich

- der Verteilungsfunktion der mit ihnen wechselwirkenden Ladungsträger
- des Abstands von den Bandkanten
- des Abstands zur Mitte der Bandlücke

zur Gesamtkapazität bei [113], solange für ihre Energie

$$E_t < kT \ln \left(\frac{2N_V v_{th} \Sigma_p}{\omega} \right) = kT \ln \left(\frac{\xi_0 T^2}{\omega} \right) \quad (3.65)$$

gilt, sie also noch mit der Frequenz ω der zur Messung angelegten Wechselspannung ge- und entladen werden können.

Walter et al. [113, 35] haben einen Formalismus vorgeschlagen, der diesem unterschiedlichen Beitrag Rechnung trägt und mit dem aus den gemessenen Kapazitätsspektren Störstellenspektren berechnet werden können.

Er soll im Folgenden kurz beschrieben werden.

Der Kapazitätsbeitrag sämtlicher Störstellen $N_t(E)$ bei einer Meßfrequenz ω und der Amplitude der angelegten Messspannung \tilde{u}_{ext} ist durch das Integral

$$C = \frac{q^2}{\tilde{u}_{ext}} \int_{E_{Fp}(\infty)}^{E_t} \tilde{u}_p(x(E)) N_t(E) \frac{w}{qU_D} dE \quad (3.66)$$

über die Energie von der Fermienergie E_{Fp} im Bahngebiet bis zur Grenzerie E_t nach (3.65) gegeben. Über den Verlauf des Spannungsfalls $\tilde{u}_p(x(E))$ in der Raumladungszone, deren Weite w durch (3.42) berechnet werden kann, ist zunächst nichts bekannt, genauso, wie über den Spannungsfall über die Verarmungszone U_D . Dies wird gleich noch eine Rolle spielen.

Da meßtechnisch zunächst nur $C(\omega)$ vorliegt, $N_t(E)$ aber gesucht wird, muß der Integrand isoliert werden. Dazu wird er zunächst mit

$$\frac{dC}{d\omega} = \frac{dC}{dE} \frac{dE}{d\omega} \quad (3.67)$$

erweitert, worin

$$\frac{dE}{d\omega} = \frac{kT}{\omega}, \quad (3.68)$$

aus der Verallgemeinerung von (3.65)

$$E = kT \ln \left(\frac{\xi_0 T^2}{\omega} \right) \quad (3.69)$$

ersetzt wird. Damit wird die zugehörige Störstellendichte zu

$$\begin{aligned} N_t(E) &= -c(E) \frac{\omega}{kT} \frac{dC(\omega)}{d\omega} \\ &= -\frac{c(E)}{kT} \frac{dC(\omega)}{d \ln \omega}, \end{aligned} \quad (3.70)$$

berechnet. Die darin enthaltene logarithmische Ableitung von C nach ω ist in Abbildung 3.23 mit eingezeichnet. Ihre Halbwertsbreite entspricht etwa dem Bereich $\omega_0/2$ bis $\omega_0 \cdot 2$ um

die Mittenfrequenz ω_0 . Im Energiemaß ist die Angabe der Halbwertsbreite abhängig von der individuellen Emissionsrate ξ_0 .

(3.67) kann außerdem nach $\frac{dC}{dT} \frac{dT}{d\omega}$ umgeschrieben werden, womit

$$N_t(E) = \frac{c(E)}{k \ln(2N_V v_{th} \Sigma_p \omega^{-1})} \frac{dC}{dT} \quad (3.71)$$

lautet. Letztere Form wird seltener angewendet, da sie das Differential nach der Temperatur T enthält, was kleinere Temperaturschritte bei der Messung wünschenswert macht. Die jedoch verlängern die Meßzeit wesentlich stärker, als eine Verdichtung der Frequenzschritte, die für (3.70) benötigt wird.

Zur Berechnung der Gewichtsfunktion $c(E)$, die den Beitrag der Störstellen bei E berücksichtigt, wird der Feldverlauf des Meßsignals $\tilde{u}_p(x(E))$ benötigt, von dem a priori nichts bekannt ist. Eine Abschätzung unter der, wie sich in Kapitel 5 herausstellen wird, plausiblen Annahme einer homogenen Dotierung ergibt bei Kontakt mit einem hochdotierten n-Halbleiter wie ZnO einen parabolischen Verlauf der Bänder [104]. Dann berechnet sich $c(E)$ zu [112]

$$c(E) = \frac{2U_D^{3/2}}{w \sqrt{q(E - E_{Fp}(\infty))}}. \quad (3.72)$$

Theoretisch würde zur Aufnahme eines Defektspektrums eine Messung in Abhängigkeit der Frequenz oder Temperatur *alleine* genügen: Gleichung (3.70) ist schon bei einer Temperatur vollständig bestimmt, wenn der Zusammenhang (3.69) zwischen Frequenz, Temperatur und der Energie bekannt ist. Da dies jedoch nicht von vornherein der Fall ist, sind zumindest Messungen bei einigen verschiedenen Temperaturen notwendig. Außerdem hängt die Störstellenenergie nur logarithmisch von der

Frequenz ab, so daß sich selbst unter Ausnutzung des gesamten Frequenzbereichs der Meßbrücke HP4192A von 5 Hz bis 13 MHz nur ein Energiebereich von

$$E_{LF} - E_{HF} = kT \ln \frac{13 \cdot 10^6}{5} = 14kT, \quad (3.73)$$

also bei Raumtemperatur etwa 360 meV, erfassen läßt. Tatsächlich ist der nutzbare Frequenzbereich wegen des Serienwiderstandes meist deutlich kleiner und der Bereich wird mehr als halbiert. Liegt ω_t außerhalb des Frequenzbereichs, ist die zugehörige Störstelle nicht detektierbar. Durch zusätzliche Variation der Temperatur von $T_l = 80$ bis $T_h = 400$ K, wie es mit der im Rahmen dieser Arbeit aufgebauten Apparatur möglich ist, kann dieses „Energiefenster“ nochmals wegen

$$\begin{aligned} E_h - E_l &= kT_h \ln \frac{\xi_0 T_h^2}{\omega} - kT_l \ln \frac{\xi_0 T_l^2}{\omega} \\ &\approx 2 \ln \frac{\xi_0 T_h}{\omega} \cdot 27 \text{ meV}, \end{aligned} \quad (3.74)$$

$$1 < \ln \frac{\xi_0 T_h}{\omega} \quad (3.75)$$

je nach Frequenz und Emissionsrate um mehrere 10 bis 100 meV verschoben werden.

3.4.2 Störstellen an den Grenzflächen

Wie eingangs erwähnt, ist der Ort der Detektion nur durch die Überschneidung des Fermi-niveaus mit den Störstellenniveaus festgelegt. Je nach Störstellenenergie liegt der Detektionsort also in unterschiedlichen Tiefen des Absorbers und ist selbst nicht genauer bestimmbar. Insbesondere kann daher zunächst nicht zwischen Störstellen im Volumen und solchen an der Grenzfläche unterschieden werden. Auch ist eine Unterscheidung zwischen Minoritäts-

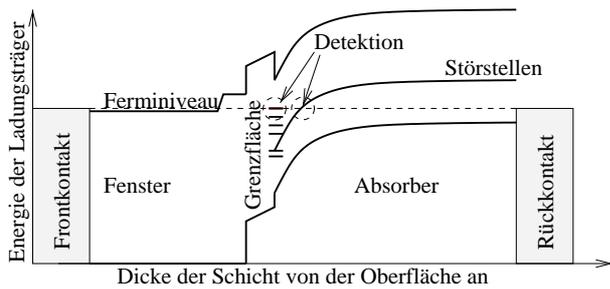


Abbildung 3.24: Bandverlauf und Grenzflächenzustände. Ohne angelegte Spannung verläuft das Fermi-Niveau glatt durch die Struktur. Die Stellen, an denen es Störstellen schneidet sind mit Kreisen markiert.

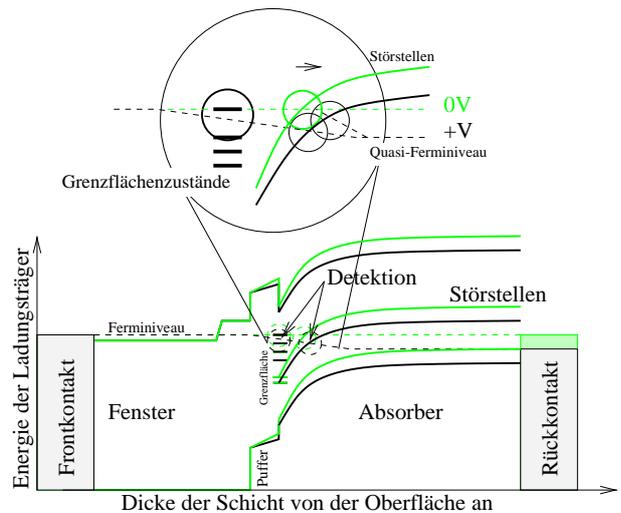


Abbildung 3.25: Bandverlauf unter angelegter Vorwärtsspannung. Bei einer angelegten Spannung verläuft das Quasi-Fermi-Niveau um so steiler, je niedriger die betreffende Schicht dotiert ist. Die Stellen, an denen das Quasi-Fermi-Niveau Störstellen schneidet, sind mit Kreisen markiert. Zum Vergleich ist das spannungslose Schema 3.24 grau unterlegt.

und Majoritätsträgerstellen nicht möglich. Bei einem Bandverlauf, wie er für Chalkopyritsolarzellen als typisch angenommen wird [64], erscheint jedoch zumindest ein Anhaltspunkt für die Existenz von Grenzflächenzuständen denkbar. In Abbildung 3.24 ist der Bandverlauf ohne angelegte Spannung und in Dunkelheit gegen die Schichttiefe aufgetragen. Die Stellen, an denen das Fermi-Niveau die willkürlich eingezeichneten Störstellen schneidet und damit durch Kapazitätsmessungen nachweisbar macht, sind mit Kreisen markiert. Wird nun eine kleine Vorwärtsspannung angelegt, wie in Abbildung 3.25 gezeigt, verschieben sich diese Stellen gegenüber dem spannungslosen Fall durch die Aufspaltung des Fermi-Niveaus in zwei Quasi-Fermi-Niveaus für die beiden Ladungsträgerarten:

1. Die Detektion der Volumendefekte verschiebt sich *räumlich* (\rightarrow). Es wird nach wie vor ein Kapazitätsbeitrag der Störstellen gemessen.
2. Das Quasi-Fermi-Niveau an der Grenzfläche verschiebt sich *energetisch* gegenüber den Grenzflächenzuständen. Es verläßt in dem Beispiel den obersten in

Bild 3.25 eingezeichneten Defekt, der daher nicht mehr zur Kapazität beitragen kann.

Eine *Veränderung* des Kapazitätsspektrums mit der angelegten Vorspannung ist daher ein Indiz für das Vorhandensein von Grenzflächenzuständen, kann aber auch Ausdruck einer *räumlich inhomogenen* Verteilung von Defekten in der Raumladungszone sein. Die *Entzerrung* der gemessenen $C(f, T)$ -Verläufe nach Gleichung (3.70) ist bei Grenzflächenzuständen selbstverständlich nicht mehr korrekt, da sie für Volumendefekte hergeleitet wurde und demzufolge falsch skalierte Konzentrationen $N_t(E)$ liefert.

Tatsächlich zeigen manche $C(V)$ -Verläufe einen entsprechenden „Knick“, der mit dem Simulationsprogramm SCAPS [16, 73, 72] qualitativ nachvollzogen werden kann. Dieses Pro-

gramm löst die Ladungstransportgleichungen in Halbleitern [104] in einer Dimension durch Unterteilung des Halbleiterbauteils in Abschnitte konstanter Eigenschaften und anschließende numerische Iteration unter gegebenen Randbedingungen, wie anliegender Spannung und Lichteinstrahlung. Als Bauelement kann eine Schichtstruktur mit Haft- und Rekombinationsdefekten sowohl an den Grenzflächen zwischen den Materialien, als auch im Volumen des Materials vorgegeben werden.

In Abbildung 3.26 wurden unter der Annahme eines Defektzustandes an der Grenzfläche zwischen CdS und CuGaSe₂ beispielhaft¹² drei derartige $C(V)$ -Kurven berechnet. Für die Schichtdicken wurden die Werte in Abbildung 2.7 übernommen, die übrigen Parameter für die beteiligten Materialien ZnO, CdS und CuGaSe₂ stammen aus [73, 72, 103, 104]. Nur die Art des Defekts, nicht jedoch seine energetische Lage 600 meV oberhalb des Valenzbandes des CuGaSe₂ wurde variiert: Die obere Kurve entspricht einem donatorartigen, die in der Mitte einem neutralen, der also sowohl ein Elektron aufnehmen wie abgeben kann, aber nicht von sich aus zur Ladungsträgerdichte beiträgt, und die untere einem akzeptorartigen Defekt.

Die Steigung bei Vorwärtsspannung ist bei allen Kurven gleich und kann daher ohne weitere Annahmen zur Auswertung nach Mott-Schottky herangezogen werden.

Ein neutraler Defekt zeigt unter den gegebenen Bedingungen keinen Effekt, ein Akzeptor dagegen erhöht die scheinbare Dotierung unter Sperrspannung erheblich. Gleichzeitig verschiebt sich der Schnittpunkt mit der Abszisse

¹²Selbstverständlich sind eine Fülle weiterer und komplexerer Defektverteilungen denkbar, die ein mitunter völlig anderes Bild ergeben würden. Aus einzelnen $C(V)$ -Messungen sind nicht genügend Informationen extrahierbar, als daß die Fülle der schon für dieses einfache Beispiel benötigten Modellparameter eindeutig zu ermitteln wäre.

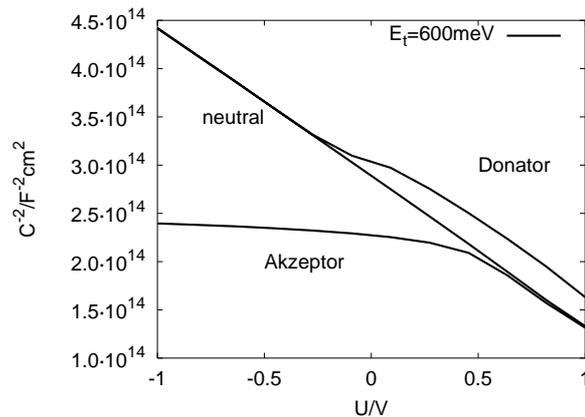


Abbildung 3.26: Mit SCAPS [16] simulierte $C(V)$ -Kennlinie. Für den Defekt an der Grenzfläche zwischen CuGaSe₂ und CdS wurde eine Dichte von $1 \cdot 10^{14} \text{ cm}^{-2}$ und der energetische Abstand 600 meV vom Valenzband des CuGaSe₂ angenommen. Akzeptor: Kann ein Elektron aufnehmen und setzt damit ein Loch frei $\rightarrow -e$ geladen. Donator: Kann ein Elektron abgeben $\rightarrow +e$ geladen. Neutral: Wirkt nur als Rekombinationszentrum, muß also gleichzeitig ein Elektron und Loch aufnehmen.

$C^{-2} \rightarrow 0$ zu unphysikalisch hohen Werten. Dies kann umgekehrt als Indiz für eine von Oberflächenzuständen dominierte Kapazität gedeutet werden.

Ein Oberflächendonator erniedrigt die Kapazität bei Vorwärtsspannung insgesamt, daher ergibt die Extrapolation auf $C^{-2} \rightarrow 0$ nach (3.46) für einen solchen Zustand ebenfalls ein – allerdings nur leicht erhöhtes – eingebautes Potential. Die Steigung und die daraus ermittelbare Dotierung bleibt, bis auf einen kleinen Bereich um 0 V herum, die gleiche wie ohne Defekt.

Umgekehrt läßt sich aus der Unabhängigkeit der Steigung der $C(V)$ -Kurve von der Spannung jedoch nicht die Existenz von Grenzflächenzuständen ausschließen: Letztere sind bei genügend hoher Dichte dazu in der Lage, Ladungsträger, die durch die Verschiebung des

Ferminiveaus eigentlich abfließen müßten, aufzunehmen und damit das Ferminiveau selbst energetisch festzuhalten¹³.

Da außerdem im Bahngebiet ein nicht unerheblicher Teil der Vorspannung fällt, muß diese um so größer sein, um auch an der Grenzfläche noch eine nennenswerte Verschiebung des Ferminiveaus zu erreichen. Größere Vorspannungen als ± 1 V werden von den bislang untersuchten Chalkopyritzellen selten über eine längere Zeit ohne signifikante Veränderung der Zellparameter verkraftet. Bei einem Dickenverhältnis des Absorbers zum Puffer von 40:1 und einem Ladungsträgerverhältnis von schätzungsweise 10:1 teilt sich der Spannungsfall etwa wie 4:1 zwischen Absorber und Puffer auf. Somit müßte wenigstens eine jeweils vierfach höhere Spannung angelegt werden, um einen nennenswerten Teil der Bandlücke auf Grenzflächenzustände abzutasten. Oder umgekehrt: Mit einer Spannungsvariation von ± 1 V ließe sich ein Bereich von bestenfalls ± 250 meV abdecken. [51] schlägt auch deshalb Spannungsbereiche von mehreren 10 V zur Spektroskopie von Grenzflächenzuständen vor.

¹³was engl. mit „pinning“ umschrieben wird