Anleitung zum Versuch "Deep Level Transient Spectroscopy" (DLTS) von PD K. Thonke, Abt. Halbleiterphysik, Universität Ulm

Inhalt:

1	Vorbemerkungen	1
2	Die Kennlinien von pn-Übergängen	2
3	Die Kennlinien von Schottky-Übergängen	
4	Das C(V)-Meßverfahren	
5	Wie wird die Kapazität gemessen?	
6	Rekombination über tiefe Störstellen	
7	Die Kapazitätstransiente	
8	Prinzip der Kapazitätstransientenmessung DLTS	
9	Gold in Silizium: Simulation von DLTS-Kurven	
-		

1 Vorbemerkungen

In diesem Versuch geht es darum, mit elektrischen Meßmethoden die wichtigsten Eigenschaften von Halbleiter/Metall- und Halbleiter-pn-Übergängen zu bestimmen, sowie Informationen über flache und tiefe Störstellen in den beteiligten Halbleiterschichten zu erhalten.

Zunächst werden an diversen Dioden die Strom/Spannungs- (I(V)-) Kennlinien aufgenommen und mit theoretischen Kurven verglichen.

Dann werden mit Kapazitäts/Spannungs- (C(V)-) Messungen die Dotierprofile der flachen Majoritätsstörstellen ermittelt.

Mit temperaturabhängigen Kapazitätstransientenmessungen (= DLTS) werden schließlich die Konzentration und Bindungsenergie von absichtlich oder unabsichtlich vorhandenen tiefen Störstellen erfasst. Für diesen Zweck wird das DLTS-Meßverfahren auch standardmäßig bei Halbleiterherstellern eingesetzt.

1

2 Die Kennlinien von pn-Übergängen

2.1 Bandstruktur und Ladungsträgerdichten beim pn-Übergang im thermischen Gleichgewicht

Vereinfachende Annahmen:

- Die Konzentrationen n und p von Elektronen und Löchern in der Raumladungszone werden zu Null gesetzt
- das Dotierprofil ist auf beiden Seiten homogen und ändert sich abrupt
- alle Dotieratome sind vollständig ionisiert: $N_D = N_{D+}$ = n_{no} , $N_A = N_{A-} = p_{po}$ d.h. es gibt als ortsfeste Ladungen nur die Rümpfe der Dotieratome

Das Feld E(x) und der Potentialverlauf $\varphi(x)$ können aus der gegebenen Raumladungsdichte $\rho(x)$ durch Lösen der Poisson-Gleichung berechnet werden.



Poisson-Gleichung:

$$-\frac{\P^2 \boldsymbol{j}}{\P x^2} = \frac{\P E_x}{\P x} = \frac{\boldsymbol{r}(x)}{\boldsymbol{e}\boldsymbol{e}_0} = \frac{q}{\boldsymbol{e}\boldsymbol{e}_0} \left\{ p(x) - n(x) + N_D - N_A \right\}$$
(1)

Für die Ladungen gilt nach der Voraussetzung des abrupten Übergangs:

$$\begin{cases} \frac{\P E_x}{\P x} = -\frac{q}{\boldsymbol{e}\boldsymbol{e}_0} N_A & \text{für } -\mathbf{x}_p \le x \le 0\\ \frac{\P E_x}{\P x} = +\frac{q}{\boldsymbol{e}\boldsymbol{e}_0} N_D & \text{für } 0 \le x \le x_n \end{cases}$$
(2)

Für das elektrische Feld ergibt sich:

$$E_{x}(x) = -\frac{qN_{A}}{ee_{0}}(x + x_{p}) \quad f\ddot{u}r \quad -x_{p} \le x \le 0$$

$$E_{x} = \frac{qN_{D}}{ee_{0}}(x - x_{n}) \quad f\ddot{u}r \qquad 0 \le x \le x_{n}$$
(3)

Die max. Feldstärke entsteht bei x = 0:

$$E_{x.\max} = \frac{-qN_D x_n}{\boldsymbol{e}\boldsymbol{e}_0} = \frac{-qN_A x_p}{\boldsymbol{e}\boldsymbol{e}_0} \tag{4}$$

Durch nochmalige Integration ergibt sich das Potential; die Randbdingungen:

$$\begin{cases} \boldsymbol{j}(x) = \frac{qN_A}{2\boldsymbol{e}\boldsymbol{e}_0} (x + x_p)^2 & \text{für } - x_p \leq x \leq 0 \\ \boldsymbol{j}(x) = -\frac{qN_D}{\boldsymbol{e}\boldsymbol{e}_0} \left(\frac{1}{2}x^2 - x_nx\right) + \frac{qN_A}{2\boldsymbol{e}\boldsymbol{e}_0}x_p^2 & \text{für } 0 \leq x \leq x_n \end{cases}$$
(6)

Für die Diffusionsspannung UD folgt:

$$U_D = \mathbf{j} \left(x_n \right) = \frac{q N_D}{2 \mathbf{e} \mathbf{e}_0} x_n^2 + \frac{q N_A}{2 \mathbf{e} \mathbf{e}_0} x_p^2 \tag{7}$$

Die Neutralitätsbedingung $N_D x_n = N_A x_p$ ist eine zweite Beziehungsgleichung für x_n und x_p , so daß man nach x_n und x_p auflösen kann:

$$x_n^2 = \frac{2ee_0}{q} U_D \frac{N_A}{N_D(N_D + N_A)}$$
 und $x_p^2 = \frac{2ee_0}{q} U_D \frac{N_D}{N_A(N_D + N_A)}$

Daraus folgt die Weite $W = x_n + x_p$ der Raumladungszone (RLZ):

$$W = \sqrt{\frac{2\boldsymbol{e}\boldsymbol{e}_0}{q} \cdot \frac{N_D + N_A}{N_D \cdot N_A} \cdot \boldsymbol{U}_D} \tag{8}$$

Zahlenbeispiel:

$$N_D = 10^{15} \text{cm}^{-3}$$
, $N_A = 10^{20} \text{cm}^{-3}$ (wie oben): $W = 1.05 \ \mu$

Bei einer genaueren Betrachtung muß die "Aufweichung" der RLZ an den Rändern durch die Abschirmeffekte der Majoritätsladungsträger mitgenommen werden. Hier kommt die Debyelänge L_D ins Spiel. Für den Fall N_A >> N_D ändert sich die Weite damit zu:

$$W = \sqrt{\frac{2\boldsymbol{e}\boldsymbol{e}_0}{qN_D} \left(\boldsymbol{U}_D - \frac{2kT}{q}\right)} = L_D \sqrt{2\left(\frac{qU_D}{kT} - 2\right)}$$
(9)

Hierbei ist die Debye-Länge L_D :

$$L_D = \sqrt{\frac{ee_0}{qN_D} \cdot \frac{k_B T}{q}}$$
(10)

2.2 Die Strom-Spannungs-Kennlinie I(V)

Die einfachste Beschreibung für die I(V)-Kennlinie eines abrupten pn-Überganges erhält man mit diesen Idealisierungen:

- Boltzmann-Näherung zur Beschreibung der Besetzungswahrscheinlichkeit
- "low injection": $n_p \ll p_p$,, $p_n \ll n_n$
- keine Generation und Rekombination in RLZ; d.h. Ströme I_p und I_n sind konstant durch RLZ

Durch Betrachtung der Ladungsträgerdichte-Gradienten an den Rändern der RLZ gelangt man zur Shockley-Glg. (1949) [1]

$$j = j_p + j_n = j_s (e^{qV/kT} - 1);$$

$$j_s = \frac{qD_p p_{n0}}{L_p} + \frac{qD_n n_{p0}}{L_n}$$
Sättigungs strom (11)

mit $L_p = \sqrt{D_p t_p}$ =Diffusionslänge der Löcher im feldfreien n-Gebiet,

$$L_n = \sqrt{D_n t_n}$$
 =Diffusionslänge der Elektronen im feldfreien p-Gebiet,
 $D_p = \frac{\mathbf{m}_p kT}{q}; \quad D_n = \frac{\mathbf{m}_n kT}{q}$ = Diffusionskonstanten der Ladungsträger,

 τ_p bzw. τ_n = Minoritätsträger-Lebensdauern,

 p_{n0} bzw. n_{p0} = Gleichgewichts-Minoritätsträgerdichten

Ein Zahlenbeispiel für die Sperrstromdichte js:

$$\begin{split} p_{po} &= N_A = 10^{20} \ cm^{-3}, \quad n_{no} = N_D = 10^{15} \ cm^{-3} \\ p_{no} &= p_{po} e^{-qV_{bi}/kT} = 7.7 \times 10^4 \ cm^{-3} \\ n_{po} &= n_{no} e^{-qV_{bi}/kT} = 0.77 \ cm^{-3} \\ Mit \ D_n &= 33.7 \ cm^2/s, \ D_p = 12 \ cm^2/s \ und \ \tau_n = \tau_p = 10^{-6} \ s \ folgt \\ L_n &= 6 \times 10^{-3} \ cm, \ L_p = 3.5 \times 10^{-3} \ cm \ und \ j_s \approx 4 \times 10^{-11} \ A/cm^2. \\ Bei \ einer \ Diodenfläche \ von \ 1 \ mm^2 \ folgt \ I_s \approx 4 \times 10^{-13} \ A. \\ Realität: \ kleinste \ Sperrströme \ von \ ca. \ 20 \ nA \end{split}$$

Von dieser Gleichung gibt es in der Realität Abweichungen durch:

1. Generation/Rekombination in RLZ

vorwärts: RLZ "voll" \rightarrow Rekombination wichtig rückwärts: RLZ "leer" \rightarrow Generation wichtig

a) rückwärts

 $n, p < n_i$

Generationsrate $G = -\frac{n_i}{t_{eff}}; \frac{1}{t_{eff}} \approx \frac{1}{N_i s v_{ih}}$

für den in der RLZ generierten Strom ergibt sich für einen abrupten Übergang

$$j_{gen} = \int_{0}^{W} q |G| dx \approx \frac{q n_i W}{t_{eff}}; W \sim \sqrt{U_D + U_{ext}}$$

Der Gesamtstrom wird (mit $p_{n0} >> n_{p0}, \, |V| > {^{2kT}/_q})$

$$j_{r,ges} = j_{gen} + j_{diff} = \underbrace{q_{\sqrt{\frac{D_p}{t_p}}} n_i^2}_{Diff,-Strom} + \underbrace{\frac{qn_i W}{t_{eff}}}_{Generat.}$$

b) vorwärts

zusätzlich zu Diffusionsstrom tritt Rekomb.-Strom j_{rek} auf Rekomb.-Rate R = R(σ , N_t, E_t,)

max. für $E_t \approx E_i \rightarrow R_{max} \cong \frac{1}{2}\sigma v_{th} N_t n_i exp(\frac{qV}{2kT})$; Faktor 2 in Exponent aus Tatsache, daß R = max, wenn $E_F = E_i$

Rekomb.-Strom
$$j_{rek} = \int_{0}^{W} qRdx = \underbrace{\frac{sv_{ih}N_{i}}{2}n_{i}qWe^{\frac{qV}{2kT}}}_{\text{Rekomb.-Strom}}$$
; wie bei Rückwärtsstrom ~ $n_{i}(T)$
 $j_{ges,v} = \underbrace{q\sqrt{\frac{D_{p}}{t_{p}}n_{i}^{2}}e^{\frac{qV}{kT}}}_{\text{Diff.-Strom}} + \underbrace{\frac{sv_{ih}N_{i}}{2}n_{i}qWe^{\frac{qV}{2kT}}}_{\text{Rekomb.-Strom}}$

 \rightarrow 2 Anteile mit verschiedener Spannungs-Abhängigkeit!

2. Hohe Injektion

Minoritäten werden vergleichbar mit Majoritäten: $n_p \approx < p_{p,}, \ p_n \ \approx < n_n$

$$p_n|_{x_n} = \frac{n_i^2}{n_n} e^{\frac{qV}{kT}}$$
 mit $p_n \approx n_n \rightarrow p_n n_n \approx n_i^2 e^{\frac{qV}{kT}} \rightarrow p_n \approx n_n \approx n_i e^{\frac{qV}{2kT}}$

 $p_n(T)$ wie bei Rekomb.-Strom.

3. Serienwiderstand

Endliche Dotierung resultiert in ohmschem Widerstand der Neutralgebiete \rightarrow Spannungsabfall!

z.B. Hintergrunddotierung $N_D = 10^{15}$ cm⁻³ ($\rho \approx 1 \ \Omega$ cm), Fläche = 1 mm², I = 1A, Länge 1=0,1 mm

 $\rightarrow j = 10 \; \text{A/cm}^2 \; \rightarrow U = \rho \, lj = 1 V \; \text{Spannungsabfall}$

4. Durchbruch

3 Mechanismen: a) thermische Instabilität

- b) Tunnelstrom
- c) Lawinendurchbruch

5. Oberflächeneffekte

z.B. Oberflächenleckströme



I(V)-Charakteristik einer realen Si-Diode. (a)Gebiet der Generation/Rekombination, (b) Diffusionsstrom-Gebiet, (c) Hochinjektion, (d) Serienwiderstands-Effekt, (e) Rückwärts-Leckstrom durch Generation/Rekombination und Oberflächeneffekte (Aus Sze, "Physics of semiconductor devices")

2.3 Die Kapazitäts-Spannungs-Kennlinie C(V)



Der pn-Übergang kann als Kondensator betrachtet werden: zwischen leitenden Bereichen (= Neutralgebieten) befindet sich ein nichtleitendes Dielektrikum

 \rightarrow Kondensator, Plattenabstand W variiert mit V_{ext}

Die Kapazität pro Fläche ist:
$$C^* = \frac{dQ^*}{dU_{ext}}$$
 (12)

$$Q^{*} = qN_{A}x_{p} = qN_{D}x_{n} = q\sqrt{\frac{2ee_{0}}{q} \cdot \frac{N_{A}N_{D}}{(N_{D} + N_{A})} \cdot (U_{D} + U_{ext})}$$
(13)

mit den Raumladungszonen-Weiten xn und xp auf der n- bzw. p-Seite.

Im Spezialfall des einseitig abrupt Übergangs mit $N_D << N_A$ ist der größte Teil der Raumladungszone auf der n-Seite:

$$Q^{*} = \sqrt{2ee_{0} \cdot q \cdot N_{D}(U_{D} + U_{ext})}$$

$$C^{*} = \sqrt{\frac{ee_{0} \cdot q \cdot N_{D}}{2(U_{D} + U_{ext} - 2kT/q)}}$$

$$C^{*} = \frac{ee_{0}}{\sqrt{2L_{Debye}}} \cdot \frac{1}{\sqrt{\frac{q}{kT}(U_{D} + U_{ext}) - 2}}$$
(14)

Zahlenbeispiel (Si)

$$N_D = 10^{15} \text{ cm}^{-3}, N_A = 10^{20} \text{ cm}^{-3}$$

 $C^* = 9.7 \text{ nF/cm}^2 \text{ bei } V = 0$
 $C \approx 100 \text{pF}$ für Diodenfläche von 1 mm²

3 Die Kennlinien von Schottky-Übergängen

3.1 Bandstruktur und Ladungsträgerdichten beim Metall-Halbleiter-Kontakt im thermischen Gleichgewicht



Bringt man einen Halbleiter und eine Metallschicht in Kontakt, werden solange Elektronen ausgetauscht, bis sich im thermischen Gleichgewicht ein konstantes elektrochemisches Potential (= Fermienergie) auf beiden Seiten eingestellt hat. Im Metall gibt es Oberflächenladungen, die durch eine entsprechende, in den Halbleiter je nach Dotierung verschieden weit hineinreichende Raumladungs schicht kompensiert werden. Die Höhe des Bandkantensprungs and er Grenzfläche ist durch die Differenz der Austrittsarbeiten in den beiden Materialien vorgegeben.

Betrachtet wird ein Schottky-Kontakt mit einem homogen dotierten n-Halbleiter, links abgebildet für den Fall der Verarmungsrandschicht. Die Konzentration der vollständig ionisierten Donatoratome sei N_D. Bei Annahme eines abrupten Metall-Halbleiter-Übergangs und Vernachlässigung freier Ladungsträger in der Raumladungs zone gilt für die Raumladungs dichte im Halbleiter⁶:

$$\mathbf{r}(x) = \begin{cases} q N_D & \text{für } 0 \le x < W \\ 0 & \text{sonst} \end{cases}$$

Durch zweimalige Integration der eindimensionalen Poisson-Gleichung

$$\frac{d^2 \mathbf{j}_{-}}{d x^2} = -\frac{\mathbf{r}}{\mathbf{e}\mathbf{e}_0} \tag{16}$$

folgt für den Potentialverlauf:

$$\boldsymbol{j}(x) = \frac{q N_D}{\boldsymbol{e}\boldsymbol{e}_0} \frac{1}{2} (x - W)^2 \quad \boldsymbol{j}(\infty) := 0$$
(17)

Die Diffusionsspannung ergibt sich aus der Potentialdifferenz zwischen dem linken und rechten Rand der RLZ:

$$U_D = \boldsymbol{j}(0) - \boldsymbol{j}(W) = \frac{1}{2} \frac{q N_D}{\boldsymbol{e} \boldsymbol{e}_0} W^2$$
(18)

Daraus berechnet sich die Raumladungszonen-Weite W zu:

$$W = \sqrt{\frac{2 e e_0}{q N_D}} U_D \tag{19}$$

Bei Anlegen einer äußeren Spannung Uext geht dieser Ausdruck über in:

$$W = \sqrt{\frac{2\boldsymbol{e}\boldsymbol{e}_0}{q N_D} (\boldsymbol{U}_D - \boldsymbol{U}_{ext})}$$
(20)

3.2 Die Strom-Spannungs-Kennlinie I(V)

Beim Stromtransport über einen n-typ Schottky-Kontakt in Vorwärtsrichtung spielen mehrere Prozesse eine Rolle (nach [2]):

- 1. Transport von Elektronen über die Barriere (dominanter Prozeß für nicht zu hoch dotierte Halbleiter)
- 2. Quantenmechanische Tunneln der Elektronen durch die Barriere (dominant für hoch dotierte HL; wird ausgenutzt bei den meisten ohmschen Kontakten)
- 3. Rekombination in der RLZ
- 4. Löcherinjektion aus dem Metall in den HL



Für die Strukturen, die gängigerweise in DLTS untersucht werden, dominiert Porzeß (1). Eine der Theorien, mit der die ideale I(V)- Kennlinie des Schottky-Kontakts beschrieben werden kann, ist die sog. "Thermionische Emissions-Theorie" von Bethe (siehe z.B. in [2]). Hier wird über die Elektronen integriert, deren Energie groß genug ist, um über die Barrieren-Oberkante zu kommen. Als Ergebnis erhält man die Gleichung:

$$J_{s \to m} = A^* T^2 \exp\left(-\frac{q\Phi_B}{kT}\right) \exp\left(-\frac{qU_{ext}}{kT}\right)$$
(21)

mit der Barrierenhöhe Φ_B und der effektiven Richardson-Konstanten

$$A^* = \frac{4pqm^*k^2}{h^3} . (22)$$

In Rückwärtsrichtung sollte idealerweise der Sperrstrom konstant sein; er ergibt sich aus obiger Gleichung für $U_{ext} = 0$.

3.3 Die Kapazitäts-Spannungs-Kennlinie C(V)

Die Bereiche beiderseits der Raumladungszone sind elektrisch leitend, die Raumladungszone selbst aber an freien Ladungsträgern verarmt und damit nichtleitend. Diese Struktur wirkt wie ein paralleler Plattenkondensator, der ein Dielektrikum der Dicke W mit der Dielektrizitätskonstante e des Halbleitermaterials einschließt.

Für die Raumladungskapazität ergibt sich damit:

$$C = \frac{\boldsymbol{e}\boldsymbol{e}_0 A}{W} = A^2 \sqrt{\frac{\boldsymbol{e}\boldsymbol{e}_0 q N_D}{2(U_D - U_{ext})}}$$
(23)

4 Das C(V)-Meßverfahren

Durch Variation der äußeren Spannung kann man den Rand der Raumladungszone verschieben, d.h. die Dotieratome umladen. Mißt man eine kmplette Abhängigkeit der Kapazität von der äußeren Spannung, erhält man damit Informationen über die Konzentration der flachen Störstellen in dem Bereich, über den man die RLZ verschieben kann.

Für den Fall der Schottky-Diode ergibt sich aus Glg. (23):

$$\frac{1}{C^2} = \frac{2}{e e_0 A^2 q N_D} (U_D - U_{ext})$$
(24)

Für die pn-Diode folgt aus Glg. (14):

$$\frac{1}{C^{*2}} = \frac{2}{\boldsymbol{e}\boldsymbol{e}_0 q N_D} \left\{ \boldsymbol{U}_D - \boldsymbol{U}_{ext} - 2\frac{q}{kT} \right\}$$
(25)

Bei <u>homogener</u> Dotierung ergibt sich also bei einer Auftragung von $1/C^2$ über der äußeren Spannung U_{ext} eine Gerade:

Aus der Steigung kann man die Dotierung der niedriger dotierten Seite bestimmen, und aus dem x-Achsenabschnitt die Diffusionsspannung U_D.

Gleichung (24) kann man nun auch in differentieller Form anschreiben:



$$\frac{d(1/C^2)}{dU_{ext}} \approx \frac{-2}{\boldsymbol{e}\boldsymbol{e}_0 A^2 q N_D(U_{ext})} \quad \text{bzw.} \quad N_D(U_{ext}) \approx \frac{-2}{\boldsymbol{e}\boldsymbol{e}_0 A^2 q \frac{d(1/C^2)}{dU_{ext}}}$$
(26)

Bei einer <u>inhomogen</u> dotierten Probe ändert sich in einer Auftragung von $1/C^2$ über der äußeren Spannung V die Steigung (~ $1/N_D$) als Funktion der angelegten Spannung. Kennt man die Fläche der Kontakte und die Dielektrizitätskonstante des Halbleitermaterials, so kann man mit der Gleichung für Kondensatoren den aktuellen "Plattenabstand" = Weite W der Raumladungszone berechnen:

$$C(U_{ext}) = \frac{ee_0A}{W(U_{ext})} \rightarrow W(U_{ext}) = \frac{ee_0A}{C(U_{ext})}$$
(27)

Mit der Kenntnis von $N_D(U_{ext})$ und $W(U_{ext})$ kann man nun ein Tiefenprofil der Dotierung aufzeichnen.

Dies ist die sog. C(V)-Messung. Je nach Dotierung und Durchbruchspannung der Kontakte kann man einen engeren oder weiteren Teil der Halbleiterschicht im Bereich des pn-Übergangs messtechnisch erfassen. In einer Abwandlung dieser Methode wird ein Kontakt durch eine Elektrolytlösung ersetzt, die die Halbleiterschicht von der Oberfläche her allmählich definiert abätzt. Hiermit kann man (allerdings nur zerstörend) auch Tiefenprofile der Majoritäts-Dotierung über weitere Bereiche aufnehmen.

5 Wie wird die Kapazität gemessen?

Die Kapazität wird im vorliegenden experimentellen Aufbau mit einer Brückenschaltung gemessen. Ein Zweig der Brücke besteht aus einer Spule mit Mittelanzapfung, der andere Zweig aus einer Vergleichskapazität C_{diff} und der zu messenden Kapazität C_{test} . Ein Generator erzeugt ein Testsignal mit einer Frequenz von 1 MHz. Dieses wird über eine Hilfswicklung auf der Spule in die Brückenschaltung mit ca. 50 mV Amplitude am Meßobjekt eingekoppelt.

In der Diagonale der Brücke tritt nun ein Meßsignal von 1MHz auf, das – abhängig von Real- und Imaginärteil der komplexen Impedanz des Testobjektes - in seiner Phase verschoben ist gegen das eingekoppelte Testsignal. Dieses Signal wird über einen schmalbandigen Zwischenverstärker verstärkt und zwei phasenempfindlichen Detektoren (engl.: PSD = phase sensitive detector) zugeführt. Einer der PSDs erhält direkt das Oszillator-Referenzsignal und zeigt an seinem Ausgang den Anteil des Teststroms an, der in Phase ist mit dem Testsignal und daher vom Realteil des Widerstandes herrührt. Der andere PSD erhält ein um 90° phasenverschobenes Referenzsignal und zeigt an seinem Ausgang daher an, welccher Anteil des Stromes vom Imaginärteil der zu messenden Impedanz herrührt (also hier: ~ zur Kapazität).

Bedingt durch die Bandbreite des Zwischenverstärkers und den Tiefpaß am Ausgang des PSD können Kapazitätsänderungen mit Zeitkonstanten von ca. 50 µsec erfaßt werden.



Prinzipschaltbild der Kapazitäts-Meßbrücke Modell Boonton 7200

6 Rekombination über tiefe Störstellen

Als "tiefe Störstellen" werden gängigerweise solche bezeichnet, deren Bindungsenergie für Elektronen oder Löcher deutlich größer sind als die von Effektive-Massen-ähnlichen Donatoren und Akzeptoren. Typischerweise bedeutet das Bindungsenergien von mehr als 100 meV, weshalb diese Störstellen bei Raumtemperatur normalerweise nicht ionisiert wind. Tiefe Störstellen werden z.B. von Übergansmetallen und Versetzungen gebildet, können aber auch durch – teilweise metastabilen – Einbau von Donator- und Akzeptoratomen in Zwischengitterplätzen entstehen.

Tiefe Störstellen spielen eine große Rolle bei der Rekombination von angeregten Ladungsträgern. Häufig sind von diesen tiefen Störstellen aus effiziente Übergänge in beide Bänder möglich, weshalb die Rekombination über diese (insbesondere in indirekten Halbleitern) schneller geht als direkt von Band zu Band oder über flache Störstellen.

Teilweise werden tiefe Störstellen bewusst in Bauelemente eingebracht, um die Rekombination zu beschleunigen. Dies ist z.B. der Fall bei schnellen Schalterdioden aus Silizium, die in der Raumladungszone zusätzlich mit Gold dotiert werden. In InP oder GaAs bringt man Fe bzw. Cr als tiefe Akzeptoren ein, um die immer unbeabsichtigt vorhandenen flachen Donatoren zu kompensieren und das Substratmaterial damit "semiisolierend" zu machen.

Andererseits sind tiefe Störstellen oft eine unbeabsichtigt vorhandene und störende Spezies, die z.B. in optoelektronischen Bauelementen einen störenden Konkurrenzkanal zu den Licht-emittierenden Übergängen bedeuten.

Störniveaus teilt man in verschiedene Gruppen ein: Generations- bzw. Rekombinationszentrum heißen Störstellen, die mit beiden Bändern in Wechselwirkung treten:

für	$c_n \gg e_p$	und	$c_p \gg e_n$	Rekombinationszentrum,
für	$c_n \ll e_p$	und	$c_p \ll e_n$	Generationszentrum.

Bei Wechselwirkung nur mit einem Band werden tiefe Niveaus als Einfangzentrum oder Haftstelle ("Trap") bezeichnet:

für e _n » e _p	Elektronenhaftstelle,
für e _n « e _p	Löcherhaftstelle.

Das DLTS-Meßverfahren dient dazu, Umladevorgänge von tiefen Störstellen zu messen und die Konzentration, die Bindungstiefe und den Einfangquerschnitt (\rightarrow Einfangeffizienz) dieser Störstellen zu ermitteln. Im folgenden wird nun das Standardmodell zur Beschreibung solcher Einfang- und Wiederemissionsvorgänge vorgestellt.

Die Beschreibung dieser Störstellenrekombination erfolgt durch das Shockley-Read-Hall-Modell [3, 4]. Die vier elementaren Prozesse sind jeweils der Einfang und die Emission von Elektronen und Löchern. Sie sind zur Verdeutlichung in hier dargestellt:



Indirekte Rekombination über ein Störniveau

Die Reaktionskinetik wird über die Einfangraten c_n für Elektronen bzw. c_p für Löcher, sowie die Emissionsraten e_n und e_p beschrieben. Sie geben die Anzahl der in einer Volumeneinheit eingefangenen bzw. emittierten Ladungsträger pro Zeiteinheit an. Einfangprozesse werden mittels eines Einfangquerschnitts s beschrieben. Bei Vorhandensein von n freien Elektronen pro Einheitsvolumen, die sich mit einer mittleren thermischen Geschwindigkeit

$$\langle v_n \rangle = \sqrt{\frac{3 k T}{m_n^*}} \tag{28}$$

bewegen, ist ein Einfangzentrum pro Flächen- und Zeiteinheit einem Strom von $n < v_n >$ Elektronen mit der effektiven Masse m_n^* ausgesetzt.

Pro Einheitsvolumen seien insgesamt N_T tiefe Störstellen vorhanden. Diese seien im unbesetzten Zustand einfach positiv geladen, im mit Elektronen besetzten Zustand neutral. Die Anzahl der besetzten Störstellen im Einheitsvolumen werde mit n_T bezeichnet. Die Anzahl der von den $(N_T - n_T)$ unbesetzten Störstellen in einem Zeitintervall dt eingefangenen Elektronen ist damit:

$$dn_T = \mathbf{s}_n < v_n > n \ (N_T - n_T) \ dt \tag{29}$$

Die Einfangrate c_n einer unbesetzten Störstelle wird nun folgendermaßen definiert:

$$c_n := \frac{I}{N_T - n_T} \frac{dn_T}{dt} \tag{30}$$

Einsetzen von Glg. (29) ergibt:

$$c_n = \mathbf{s}_n < v_n > n \tag{31}$$

Analog ergibt sich ein Ausdruck für die Löchereinfangrate bei Anwesenheit von p freien Löchern mit einer thermischen Geschwindigkeit $\langle v_p \rangle$:

$$c_p := \frac{1}{n_T} \frac{d(N_T - n_T)}{dt} = \mathbf{s}_p < v_p > p$$
(32)

Die thermische Emissionsrate läßt sich folgendermaßen darstellen:

$$e_n = A_n \ e^{-\frac{E_C - E_T}{kT}}$$
(33)

mit An als einem von den Eigenschaften der Störstelle abhängigen Faktor.

Als direkte Folge des zweiten Hauptsatzes der Thermodynamik ergibt sich das Prinzip des detaillierten Gleichgewichts: Im thermischen Gleichgewicht muß jeder Prozeß mit seinem eigenen Umkehrprozeß im Gleichgewicht stehen. Das bedeutet in diesem Fall, daß jeweils gleichviele Elektronen und Löcher eingefangen wie emittiert werden:

$$e_n f = c_n (1 - f) \tag{34}$$

worin f die Wahrscheinlichkeit ist, daß die Störstelle besetzt ist. Diese Wahrscheinlichkeit ist nach der Fermi-Dirac-Statistik:

$$f = \frac{l}{l + e^{\frac{ET \cdot EF}{kT}}}$$
(35)

Für einen nichtentarteten Halbleiter (d.h.: nicht zu hoch dotiert, $E_C \gg E_F \gg E_V$) kann die Elektronenkonzentration durch die Boltzmann-Statistik genähert werden und ergibt sich zu:

$$n = N_C e^{\frac{E_C \cdot E_F}{kT}}$$
(36)

mit N_C der effektiven Zustandsdichte des Leitungsbandes:

$$N_{C} = 2 M_{C} \left(\frac{2 p m_{n}^{*} k T}{h^{2}} \right)^{\frac{1}{2}}$$
(37)

worin M_C die Anzahl äquivalenter Leitungsbandminima, m_n^* die effektive Zustandsdichtemasse für Elektronen und h das Planck'sche Wirkungsquantum bezeichnet.

$$e_n = \boldsymbol{s}_n < v_n > N_C \ e^{\frac{EC \cdot ET}{kT}}$$
(38)

Der analoge Ausdruck für die Löcheremission lautet:

$$e_p = \mathbf{s}_p < v_p > N_V e^{-\frac{E_F \cdot E_V}{kT}}$$
(39)

Für die Konzentration n_T der mit Elektronen besetzten Störstellen gilt die Differentialgleichung⁸⁰:

$$\frac{d n_T}{d t} = (c_n + e_p)(N_T - n_T) - (c_p + e_n)n_T$$
(40)

Als Lösung dieser Gleichung ergibt sich für den reinen Emissionsprozeß, bei dem zur Zeit t = 0 alle Störstellen besetzt sind:

$$n_{T}(T) = \begin{cases} N_{T} & \text{für } t \leq 0\\ N_{T} & \mathbf{t} \left[(c_{n} + e_{p}) + (c_{p} + e_{n}) \left(1 + e^{-\frac{t}{t}} \right) \right] & \text{für } t > 0 \end{cases}$$

$$(41)$$

mit der Emissionszeitkonstanten

$$\boldsymbol{t} = \frac{1}{c_n + e_p + c_p + e_n} \tag{42}$$

Für eine Elektronen-Haftstelle in einem n-Halbleiter (und analog für eine Löcher-Falle im p-Halbleiter) können einige Näherungen gemacht werden:

• Die Emissionsrate für Löcher e_p ist (nach Definition der Elektronen-Haftstelle) klein gegen die der Elektronen e_n und darf vernachlässigt werden.

- Die Löcherkonzentration im Valenzband p und damit die Löcher-Einfangrate c_p darf vernachlässigt werden.
- Wegen der experimentellen Bedingungen in der Raumladungszone einer Schottky-Diode kann die Elektronenkonzentration im Leitungsband n_0 und damit auch der Einfang c_n vernachlässigt werden.

Damit gilt für den Emissionsprozeß die vereinfachte Gleichung:

$$n_T = N_T \quad e^{\frac{1}{t}} \tag{43}$$

mit

$$\boldsymbol{t} = \frac{1}{\boldsymbol{e}_n} = \frac{1}{\boldsymbol{s}_n < \boldsymbol{v}_n > N_C \ \boldsymbol{X}_n \ \boldsymbol{e}^{-\frac{\boldsymbol{E}_C \cdot \boldsymbol{E}_T}{kT}}}$$
(44)

Der sog. Entropiefaktor X_n ergibt sich aus einer Präzisierung des Energiebegriffs: Thermodynamisch betrachtet geht es hier um Experimente unter Vorgabe von Temperatur T und Druck p. Die Variablen T und p sind die zur freien Enthalpie G (Gibbs'sches Potential) gehörenden natürlichen Variablen:

$$\Delta G = \Delta H - T \Delta S \tag{45}$$

Die Energie in Gleichung (45) läßt sich so in zwei Anteile aufspalten:

$$n = N_C X_n e^{-\frac{\Delta H_n}{kT}}$$
(46)

mit der Aktivierungsenthalpie

$$\Delta H_n = E_C - E_T \tag{47}$$

und dem Entropiefaktor

$$X_n = e^{\frac{\Delta S_n}{k}} \tag{48}$$

Gleichung (44) kann für eine Arrhenius-Darstellung umgeformt werden:

$$\ln(t \leq v_n > N_C) = \frac{E_C - E_T}{k} \frac{1}{T} - \ln(X_n \boldsymbol{s}_n)$$
(49)

Durch Messung der Emissionszeitkonstanten in Abhängigkeit von der Temperatur kann damit die energetische Lage des Störstellenniveaus und das Produkt aus Einfangquerschnitt und Entropiefaktor bestimmt werden. Einfangquerschnitt und Entropiefaktor lassen sich separieren, falls zusätzlich die Einfangzeitkonstante gemessen werden kann.

Der Einfangquerschnitt s ist nicht in jedem Falle temperaturunabhängig. Durch verschiedene Effekte, wie z.B. Ladungsträgereinfang über mehrere Stufen oder Multiphononenprozesse, kann ein Temperatureinfluß gegeben sein, der dann bei der Interpretation einer Arrheniusgrafik nach Gleichung (49) berücksichtigt werden muß.

Weiterhin ist auch die Emissionsrate nicht ausschließlich von der Temperatur abhängig, sondern kann durch elektrische Felder beeinflußt werden. Es gibt drei zugrundeliegende physikalischen Prozess, die hier nur am Rande erwähnt seien:

1) Der Poole-Frenkel-Effekt kann die Potentialschwelle einer tiefen Störstelle in einem elektrischen Feld erniedrigen.

- 2) Quantenmechanisches Tunneln ermöglicht (allerdings erst ab hohen Feldstärken $> 10^9$ V/m in merklichem Maße) das Durchdringen von Potentialbarrieren.
- Bei phononenunterstütztem Tunneln nehmen die Ladungsträger thermische Energie aus dem Kristallgitter auf und können so mit größerer Wahr scheinlichkeit die Potentialschwelle durchtunneln.

7 Die Kapazitätstransiente

Eine vorübergehende Änderung der Weite der Raumladungszone und damit der Kapazität eines Schottky- oder pn-Kontaktes entsteht dadurch, dass freie Ladungsträger während eines Vorwärtspulses in tiefe Störstellen eingefangen werden, aus denen sie erst verzögert durch thermische Emission wieder entkommen können. Da damit freie Ladungsträger in den an die Raumladungszone angrenzenden Bereichen fehlen, wird die Raumladungszone vorübergehend breiter und damit die Kapazität kleiner. Diese temporäre Kapazitätsänderung ist die sog. "Kapazitätstransiente".

Für eine quantitative Analyse wird ein Schottky-Kontakt aus einem Metall und einem n-Halbleiter betrachtet. Im Halbleiter sei eine Konzentration N_D an vollständig ionisierten Do-

tieratomen vorhanden, sowie eine (kleinere) Konzentration N_T an tiefen Störstellen. Dotieratome und tiefe Störstellen sollen im unbesetzten Zustand einfach positiv, im besetzten Zustand neutral sein ("Elektronen-Fallen").

Im Ruhezustand liege die Sperrspannung U_R an. Die entstehende Raumladungszone erstreckt sich dann, wie in rebenstehender Abbildung illustriert, über den Bereich x = 0 bis x = W_R .

Die tiefen Störstellen sind in diesem Ruhezustand besetzt und damit neutral, wenn sie sich energetisch unterhalb des Ferminiveaus befinden. Der Kreuzungspunkt zwischen Ferminiveau E_F und dem Niveau der tiefen Störstellen E_T werde mit L_R bezeichnet. In der Raumladungs zone befinden sich damit im Bereich 0..W_R geladene Donatorrümpfe und im Bereich 0..L_R unbesetzte, positiv geladene tiefe Störstellen. Damit lautet die Poisson-Gleichung:

Metall n-Halbleiter G_{F} W_{R} E_{C} G_{F} W_{R} E_{C} G_{F} W_{R} E_{V} N_{D} $+ N_{1}$ N_{D} N_{D} L_{R} W_{R} X

Schottky-Kontakt mit tiefen Störstellen unter Rückwärtsspannung U_R im Gleichgewichtszustand

$$\frac{d^{2}\boldsymbol{j}}{d_{x}^{2}} = -\frac{\boldsymbol{r}}{\boldsymbol{e}\boldsymbol{e}_{0}} = -\frac{q}{\boldsymbol{e}\boldsymbol{e}_{0}} \begin{cases} N_{D} + N_{T} & \text{für } 0 \le x < L_{R} \\ N_{D} & \text{für } L_{R} \le x < W_{R} \\ 0 & \text{für } W_{R} \le x \end{cases}$$
(50)

Durch zweifache Integration ergibt sich daraus das elektrische Potential:

$$\mathbf{j} = -\frac{q}{2 \, \mathbf{e} \mathbf{e}_0} \begin{cases} N_D (x - W_R)^2 + N_T (x - L_R)^2 & \text{für } 0 \le x < L_R \\ N_D (x - W_R)^2 & \text{für } L_R \le x < W_R \\ 0 & \text{für } W_R \le x \end{cases}$$
(51)

Die Potentialdifferenz U_R zwischen x = 0 und $x = W_R$ berechnet sich damit zu:

$$U_{R} = \boldsymbol{j} (W_{R}) - \boldsymbol{j} (0) = \frac{e}{2\boldsymbol{e}\boldsymbol{e}_{0}} \left\{ N_{D} W_{R}^{2} + N_{T} L_{R}^{2} \right\}$$
(52)

Nun wird durch einen sog. Füllpuls die Sperrspannung kurzzeitig von U_R auf $U_P > U_R$ erniedrigt. Der Ort des Kreuzungspunktes zwischen Ferminiveau und Störstellenniveau verschiebt sich währenddessen ebenfalls kurzzeitig auf den Ort L_P . Dadurch werden die Störstellen zwischen $x = L_P$ und x = L_R mit Elektronen besetzt und sind damit neutral.

Der Besetzungs grad der Störstellen nimmt anschließend nach einer Funktion $f_n(t)$ mit der Zeit ab. Für einen rein exponentiellen Abfall der Emission ergibt sich diese Funktion $f_n(t)$ nach Gleichung (43) als eine Exponentialfunktion:

$$f_n = e^{-\frac{t}{t}} \tag{53}$$

Zur Raumladungsdichte tragen nur die unbesetzten Störstellen bei:

$$f_{p}(t) = 1 - f_{n}(t)$$
 (54)



Schottky-Kontakt mit tiefen Störstellen unmittelbar nach dem Füllpuls

Die Raumladungszone wird infolge der neuen Ladungsverteilung auf W(t) aufgeweitet, der Kreuzungspunkt zwischen Fermi- und Störstellenniveau verschiebt sich auf L(t). Da die Störstellenkonzentration N_T aber klein gegen die Grunddotie rung N_D sein soll, kann der Wiedereinfang von Ladungsträgern durch Störstellen zwischen L_R und L(t) am Ende der Raumladungszone vernachlässigt und L(t) = L_R gesetzt werden.

Die Poisson-Gleichung lautet in diesem Fall:

$$\frac{d^{2}\boldsymbol{j}}{d_{x}^{2}} = -\frac{\boldsymbol{r}}{\boldsymbol{e}\boldsymbol{e}_{0}} = -\frac{\boldsymbol{q}}{\boldsymbol{e}\boldsymbol{e}_{0}} \begin{cases} N_{D} + N_{T} & \text{für } 0 \leq x < L_{P} \\ N_{D} + f_{n}(t) N_{T} & \text{für } L_{P} \leq x < L_{R} \\ N_{D} & \text{für } L_{R} \leq x < W_{R} \\ 0 & \text{für } W_{R} \leq x \end{cases}$$
(55)

Das elektrische Potential ergibt sich wieder durch zweimalige Integration:

$$\boldsymbol{j} = \frac{-q}{2\boldsymbol{e}\boldsymbol{e}_{0}} \begin{cases} N_{D} (x - W(t))^{2} + f_{p}(t) N_{T} (x - L_{R})^{2} + (l - f_{p}(t)) N_{T} (x - L_{P})^{2} & f \ddot{u} r \quad 0 \leq x < L_{P} \\ N_{D} (x - W(t))^{2} + f_{p}(t) N_{T} (x - L_{R})^{2} & f \ddot{u} r \quad L_{P} \leq x < L_{R} \\ N_{D} (x - W(t))^{2} & f \ddot{u} r \quad L_{R} \leq x < W(t) \\ 0 & f \ddot{u} r \quad W(t) \leq x \end{cases}$$
(56)

Für die Potentialdifferenz U_R' zwischen x = 0 und x = W(t) gilt damit:

$$U_{R'} = -\frac{e}{2ee_0} \left\{ N_D W^2(t) + N_T \left(f_p(t) (L^2(t) - L_P^2) + L_P^2) \right) \right\}$$
(57)

Da die Sperrspannung vor und nach dem Füllpuls gleich ist, können U_R und U_R' gleichge setzt und nach $W^2(t)$ aufgelöst werden:

$$W^{2}(t) = W_{R}^{2} + \frac{N_{T}}{N_{D}} (1 - f_{p}(t)) (L_{R}^{2} - L_{P}^{2})$$

$$= W_{R}^{2} \left(1 + \frac{N_{T}}{N_{D}} f_{n}(t) \frac{L_{R}^{2} - L_{P}^{2}}{W_{R}^{2}} \right)$$
(58)

Mit der Gleichgewichtskapazität C_R bei stationärer Sperrspannung U_R ergibt sich daraus ein Ausdruck für die Raumladungskapazität:

$$C(t) = C_R \left(I + \frac{N_T}{N_D} f_n(t) \frac{L_R^2 - L_P^2}{W_R^2} \right)^{\frac{1}{2}}$$
(59)

Wegen $N_T \ll N_D$ kann der Ausdruck entwickelt werden, und nach Einsetzen von (53) für einen exponentiellen Abfall ergibt sich:

$$C(t) = C_R \left(1 - \frac{1}{2} \frac{N_T}{N_D} \frac{L_R^2 - L_P^2}{W_R^2} e^{-\frac{t}{t}} \right)$$
(60)

Mit der Bezeichnung ?C für die Amplitude der Transienten

$$\Delta C = C_R - C(t=0) = C(t=\infty) - C(t=0) = \frac{C_R}{2} \frac{N_T}{N_D} \frac{L_R^2 - L_P^2}{W_R^2}$$
(61)

ergibt sich für die Kapazitätstransiente die Formel:

$$C(t) = C_R - \Delta C \ e^{-\frac{t}{t}}$$
(62)

Die Differenz ? = W - L ist bei homogener Dotierung spannungsunabhängig und berechnet sich im ungestörten Halbleiter zu⁵

$$\boldsymbol{l} = \boldsymbol{W} - \boldsymbol{L} = \sqrt{\frac{2\boldsymbol{e}\boldsymbol{e}_0 \left(\boldsymbol{E}_F - \boldsymbol{E}_T\right)}{q^2 N_D}}$$
(63)

Werden durch genügend hohe Füllpulsspannung alle Störstellen in der Raumladungszone besetzt, so ist $L_P = 0$. Ist weiterhin die Sperrspannung und damit die Raumladungsweite so groß, daß ? gegen W_R vernachlässigt werden kann, so vereinfacht sich der Ausdruck für die Amplitude (61):

$$\Delta C = C_R \frac{l}{2} \frac{N_T}{N_D} \tag{64}$$

Diese Gleichung wird in der Literatur meistens angegeben, wobei die oben genannten Näherungen nicht immer explizit erwähnt werden.

8 Prinzip der Kapazitätstransientenmessung DLTS

Für die Bestimmung der relevanten elektronischen Parameter von Störstellen kann man im Prinzip sowohl einen pn-Übergang als auch einen Schottky-Kontakt verwenden. Man legt an diesen Kontakt eine Sperrspannung U_R an, wodurch die zu untersuchenden tiefen Defekte im Bereich der Raumladungszone entladen werden. Dann verringert man kurz (z.B. für 1 msec) die Sperrspannung, wodurch Majoritätsladungsträger in die tiefen Störstellen eingefüllt werden können. Misst man nun unmittelbar nach dem Zurückschalten auf die Spannung U_R die Kapazität, hat diese abgenommen und kehrt –abhängig von der Probentemperatur und der Aktivierungsenergie der Störstelle – mehr oder weniger schnell zum Anfangswert zurück.

Legt man kurz eine Spannung in Vorwärtsrichtung an, werden bei pn-Übergängen auch Minoritätsladungsträger in die Raumladungszone geschwemmt, und auch Fallen für diese Minoritäten können gefüllt werden (wird hier nicht verwendet).

Im folgenden wird die Transientenspektroskopie an einem Schottky-Kontakt mit einem n-Halbleiter und einem tiefen Donator erläutert. Die Abbildung zeigt die zeitliche Abfolge der Signale und Besetzungszustände. An den Schottky-Kontakt wird eine konstante Sperrspannung U_R angelegt, und es bildet sich eine Raumladungszone aus, deren Weite unmittelbar durch eine Kapazitätsmessung bestimmt werden kann.

Die Sperrspannung wird nun kurzzeitig durch einen Spannungspuls verringert. Dadurch wird das Ferminiveau angehoben und die Raumladungszone verkleinert. Alle Störstellen unterhalb des Ferminiveaus füllen sich nun mit Elektronen, weshalb der Spannungspuls gewöhnlich *Füllpuls* genannt wird.

Nachdem wieder die ursprüngliche Sperrspannung anliegt, ist auch die vorherige Lage des Ferminiveaus wiederhergestellt. Die neubesetzten Störstellen befinden sich nun oberhalb des Ferminiveaus, sie benötigen aber eine gewisse Zeit, um die aufgenommenen Elektronen wieder zu emittieren. Die Weite der Raumladungszone ist nun größer als vor dem Füllpuls, da sich durch die aufge ladenen Störstellen auch eine größere Ladung in der Verarmungszone befindet. Durch die größere Raumladungsweite ist die gemessene Kapazität direkt nach dem Füllpuls kleiner als vorher. Während die Störstellen die eingefangenen Elektronen wieder emittieren, erreicht die Raumladungsweite und damit auch die Raumladungskapazität wieder den ursprünglichen Wert.

Es kann also nach dem Füllpuls eine Kapazitätstransiente gemessen werden. Deren Amplitude hängt von der Konzentration der umgeladenen Störstellen ab, und ihre Zeitkonstante ist die des Emissionsprozesses. Die Störstelleneigenschaften können daher umgekehrt auch



wieder aus der Kapazitätstransiente berechnet werden. Dies ist das Grund prinzip des DLTS-Verfahrens.

Die Zeitkonstante des Emissionsprozesses ist von der Temperatur abhängig. Durch Aufnahme von Transienten bei verschiedenen Temperaturen kann nach Gleichung (49) mittels verschiedener Auswertungs methoden die Aktivierungsenergie des Emissionsvorgangs und ein Produkt aus Entropiefaktor und Einfangquerschnitt der Störstellen bestimmt werden. Der Einfangquerschnitt kann separat bestimmt werden, wenn auch der Einfangprozeß gemessen werden kann. Dieser ist aber meist zu schnell, um mittels Kapazitätsmeßbrücken direkt beobachtet zu werden und kann daher häufig nur indirekt durch Experimente gemessen werden.

Aus der Amplitude der Transiente kann die Konzentration der tiefen Störstellen nach Gleichung (61) bzw. (64) berechnet werden.

Da die Störstellen nur in einem begrenzten Gebiet umgeladen werden, das von der Ruhesperrspannung und der Füllpulshöhe abhängt, kann durch Differenzmessung bei zwei verschiedenen Füllpulshöhen auch ein räumliches Tiefenprofil erstellt werden.

Die einfachste Auswertemethode zur Ermittlung der Emissionszeitkonstanten ist die sog. "Boxcarmethode". Dem Vorschlag von D. Lang [5] folgend, wird die Differenz D der Kapazität zu zwei verschiedenen Zeitpunkten t_1 und t_2 nach dem Ende des Füllpulses erfaßt:

$$D = C(t_2) - C(t_1)$$
(65)

Dies kann unmittelbar auf elektronischem Wege geschehen mit Hilfe eines sog. "Doppel-Boxcar-Integrators", oder wie hier per Rechner nach Aufnahme der kompletten Kapazitätstransienten. Das folgende Bild zeigt eine Serie von C-Transienten, die mit (von unten nach oben) ansteigender Temperatur T immer schneller werden. Greift man nun die C-Differenz $D = C(t_2) - C(t_1)$ für diese Transienten ab, ergibt sich bei einer Auftragung des DLTS-Signals D über T der sog. "DLTS-peak" wie im rechten Teil des Diagramms dargestellt.



Kapazitätstransienten und resultierendes Boxcar-Signal

Bei niedrigen Temperaturen erfolgt die Emission der Ladungs träger so langsam, daß im Meßintervall die Kapazitätsänderung sehr gering ist. Bei hohen Temperaturen ist die Emission zum Zeitpunkt t1 schon abgeschlossen, so daß sich das Signal bis t2 nicht mehr verändert. Bei einer bestimmten Temperatur Tmax wird die gemessene Kapazitätsdifferenz maximal.

Für das Maximum des Doppel-Boxcar-Signals ergibt sich nach Einsetzen von (62) durch Ableiten nach t und anschließendes Nullsetzen ein Ausdruck für die Emissions zeitkonstante:

$$\boldsymbol{t}(T_{\max}) = \frac{t_2 \cdot t_1}{\ln\left(\frac{t_2}{t_1}\right)}$$
(66)

Wenn die Torbreite B des Boxcar-Integrators nicht vernachlässigbar klein ist, muß diese Formel modifiziert werden[6]:

$$\mathbf{t}(T_{\max}) = \frac{t_2 \cdot t_1}{\ln\left(\frac{t_2 + \frac{B}{2}}{t_1 + \frac{B}{2}}\right)}$$
(67)

Der reziproke Ausdruck 1/t(Tmax) wird als Ratenfenster bezeichnet. Verändert man die beiden Zeitpunkte t1 und t2, liegen die Maxima bei anderen Temperaturen. Damit die Linienform des Boxcarsignals konstant bleibt, ist ein festes Verhältnis t2/t1 zu wählen. Von diesem Verhältnis hängt auch die Empfindlichkeit des Verfahrens ab: Sie steigt mit zunehmendem Verhältnis t2/t1, während die Auflösung wegen der größeren Halbwertsbreite abnimmt. Sinnvolle Zahlenwerte sind t2/t1 = 2...20.

Der Doppel-Boxcar-Integrator ist ein Spezialfall einer Korrelation des Meßsignals C(t) mit einer Wichtungsfunktion h(t,t). Das Korrelatorsignal f(t) ist dann:

$$f(\mathbf{t}) = \int_{M_e} zeitC(t)h(t,\mathbf{t}) dt$$
(68)

Die spezielle Wichtungsfunktion des Doppel-Boxcar-Integrators ist:

$$h(t, \mathbf{t}) = \mathbf{d}(t - t_2) - \mathbf{d}(t - t_1)$$
(69)

worin t nach Glg (66) bzw. (67) mit t₁ und t₂ zusammenhängt.

In der Literatur wurden verschiedene andere Wichtungsfunktionen vorgeschlagen, u.a. Rechteck-, Sinus-, und Exponentialfunktionen [6, 7,]. Diese haben theoretisch ein besseres Signal/Rauschverhältnis als die Boxcarmethode, sind aber technisch meist aufwendiger zu realisieren [8].

Durch Aufnahme mehrerer DLTS-Kurven mit unterschiedlichen Ratenfenstern erhält man jeweils ein Wertepaar (t, Tmax). Werden diese Zahlenwerte nun als - ln(T2 t) gegen 1/T in einer Arrhenius-Grafik dargestellt, so sollte man sie nach Gleichung (49) durch eine Gerade annähern können. Aus der Geradensteigung erhält man die Ionisierungsenergie $E_C - E_T$ bzw. $E_T - E_V$. Aus dem Schnittpunkt der Geraden mit der - ln(T² t)-Achse kann das Produkt aus Einfangquerschnitt und Entropiefaktor berechnet werden.

Nach (62) kann die Amplitude der Kapazitätstransiente berechnet werden. Kennt man aus einer C(V)-Messung die Konzentration der flachen Störstellen, kann man aus der maximalen Kapazitätsanderung ΔC wiederum nach (61) bzw. (64) die Konzentration der tiefen Störstellen bestimmen.

Lang weist in seiner Arbeit [5] mehrmals darauf hin, daß o.g. Formeln nur im Fall exponentieller Transienten gelten. Wie in Kapitel 6 erwähnt, gibt es jedoch verschiedene Gründe für Abweichungen von der reinen Exponentialität: z.B. mehrere tiefe Störstellen, hohe Konzentrationen tiefer Störstellen, feld unterstützte Emission. In diesen Fällen muß die Gültigkeit der Auswertung sorgfältig geprüft werden[9]. Auch die Existenz einer linearen Arrhenius-Auftragung ist nicht unbedingt Merkmal eines exponentiellen Zeitgesetzes[10].

9 Gold in Silizium: Simulation von DLTS-Kurven

Für eine bekannte Störstelle im Silizium sollen hier zur Demonstration der DLTS-Kurven und der Auswertung mit dem Boxcar-Verfahren hier simulierte Kurven mit realistischen Störstellenparametern vorgestellt werden. Es werden folgende Eingabedaten verwendet , die für das Akzeptorniveau von Gold in Silizium zutreffen[11]:

Einfangquerschnitt s $_{n}$ = 8·10⁻¹⁶ cm² Energie E_C - E_T = 550 meV Kapazitätsamplitude: 1 pF Rauschanteil: 5 % Meßtemperatur: 130 K bis 220 K in Schritten zu 1 K Anzahl der erzeugten Meßpunkte: 1024 Zeitabstand der digitalisierten C-Meßwerte: 100 µs

Die aus den simulierten Transienten per Boxcar-Verfahren berechneten DLTS-Kurven sehen so aus:



Trägt man nun die Lage der Maxima in einem Arrhenius-Plot nach Glg. (49) auf, so ergibt sich dieses Bild:



Man erhält also aus der Auswertung die vorher eingegebene Aktivierungsenergie der Störstelle und deren Einfangquerschnitt recht genau wieder.

<u>Literatur</u>

- [1] siehe z.B. R. Sauer, Vorlesungsskript zum Wahlfach "Halbleiterphysik"
- [2] S.M. Sze, *Physics of Semiconductor Devices*", Wiley (New York)
- [3] W. Shockley, W. T. Read jr.; *Statistics of the Recombinations of Holes and Elektrones*; Phys. Rev. **87** (5), 835, (1952).
- [4] R. N. Hall; *Electron-Hole Recombination in Germanium*; Phys. Rev. 87 (5), 387, (1952).
- [5] D. V. Lang; *Deep-level transient spectroscopy: A new method to characterize traps in semiconductors*; J. Appl. Phys. **45**, 3023 (1974).
- [6] A. Le Bloa, P. N. Favennec; Rev. Phys. Appl. 17, 759 (1982).
- [7] C. A. B. Ball, A. B. Conibear; *Simulated lock-in amplifier deep level transient spectroscopy*; Rev. Sci. Instrum. 62 (11), 2831, (1991).
- [8] G. L. Miller, D. V. Lang, L. C. Kimerling; *Capacitance Transient Spectroscopy*; Ann. Rev. Mater. Sci. 7, 377 (1977)
- [9] W. R. Thurber, R. A. Forman, W. E. Phillips; A novel method to detect nonexponential transients in DLTS; J. Appl. Phys. 53(11), 7397, (1982).
- [10] P. Omling, L. Samuelson, H. G. Grimmeis; *Deep level transient spectroscopy evaluation of nonexponential transient in semiconductor alloys*; J. Appl. Phys. 54 (9), 5117, (1983).
- [11] D. V. LANG, H. G. GRIMMEISS, E. MEIJER, M. JAROS; Complex nature of gold-related deep levels in silicon; Phys. Rev. B 22, 3917 (1980)