

# Verfahren zur Berechnung der Lage des FERMI-Niveaus in Halbleitern mit Hilfe eines elektronischen Digitalrechners\*

Von E. RICHTER

Aus dem AEG-Forschungsinstitut, Frankfurt/Main  
(Z. Naturforsch. 18 a, 39–43 [1963]; eingegangen am 25. Oktober 1962)

Für einen elektronischen Digitalrechner wird ein Rechenverfahren beschrieben, das es gestattet, für homogene und feldfreie, aber sonst beliebige Halbleiter aus der Neutralitätsgleichung die Lage des FERMI-Niveaus für einen beliebig weiten Temperaturbereich (auch unter 1 °K) mit großer Genauigkeit zu berechnen. Dabei wird die Temperaturabhängigkeit des Bandabstandes und der Störstellenniveaus berücksichtigt. Als Beispiel wird der Verlauf des FERMI-Niveaus für Silicium verschiedener Dotierung und für Wismut-Tellurid in einigen speziellen Fällen angegeben.

Für einen homogenen, feldfreien Halbleiter läßt sich bei Kenntnis der erforderlichen Materialkonstanten die Lage des FERMI-Niveaus prinzipiell aus der Neutralitätsbedingung (1) ermitteln:

$$n + n_{A^-} = p + n_{D^+} \quad (1)$$

$n$  = Dichte der Elektronen im Leitungsband,  
 $n_{A^-}$  = Dichte der Elektronen in den Akzeptoren,  
 $p$  = Dichte der Defektelektronen im Valenzband,  
 $n_{D^+}$  = Dichte der Defektelektronen in den Donatoren.

Für die einzelnen Summanden von Gl. (1) liefert die FERMI-Statistik<sup>1</sup> bei einem diskreten Donator- bzw. Akzeptorniveau die in Tab. 1 angegebenen Ausdrücke (3) bis (6).

Die Faktoren  $g_A$  bzw.  $g_D$  berücksichtigen dabei die Festlegung des Spins der in die Störstellen einzubauenden Elektronen bzw. Defektelektronen. Die Temperaturabhängigkeit des Bandabstandes wurde in der Lage der Leitungsbandkante ( $E_c$ ) durch die im allgemeinen ausreichende lineare Korrektur  $\alpha T$  berücksichtigt; entsprechend wurde die Temperaturkorrektur der Störstellenniveaus proportional ihrer Lage bei 0 °K zur Bandkante eingesetzt.

Die vollständige Neutralitätsgleichung läßt sich als Bestimmungsgleichung für  $E_F$  [Gl. (2) in Tab. 1 mit (3) bis (6)] nur für nichtentartete Halbleiter in den drei Sonderfällen der reinen Eigenleitung, der reinen Störleitung und der reinen Kompensations-Störleitung soweit vereinfachen, daß sie nach  $E_F$  auflösbar ist. Für gemischte Halbleiter ist auch bei Nichtentartung eine Auflösung nach  $E_F$  nicht mehr möglich.

\* Teilweise vorgetragen auf der Physikertagung Stuttgart 1962.

<sup>1</sup> Siehe z. B. O. MADELUNG, Halbleiter, Handb. Phys. XX, Springer-Verlag, Berlin 1957.

$$f(E_F) = n + n_{A^-} - p - n_{D^+} = 0, \quad (2)$$

$$n = N_c \cdot \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{y^{1/2} dy}{1 + e^{y - (E_F - E_c)/kT}}, \quad (3)$$

$$p = N_v \cdot \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{y^{1/2} dy}{1 + e^{y - (E_v - E_F)/kT}}, \quad (4)$$

$$n_{A^-} = \frac{n_A}{1 + g_A e^{(E_A - E_F)/kT}}, \quad (5)$$

$$n_{D^+} = \frac{n_D}{1 + \frac{1}{g_D} e^{(E_F - E_D)/kT}}. \quad (6)$$

$E_F$  = FERMI-Niveau [eV];

$E_v$  = Oberkante des Valenzbandes [eV];

$E_{c0}$  = Unterkante des Leitungsbandes bei 0 °K [eV];

$E_c = E_{c0} - \alpha T$  = Unterkante des Leitungsbandes bei  $T$  °K;

$N_c = \frac{2}{h^3} (2\pi m_n k)^{3/2} T^{3/2}$  } Entartungskonzentrationen  
(effektive Zustandsdichten)  
für Leitungs- und Valenzband.

$E_{D0}$  = Donatorniveau bei 0 °K [eV];  $E_D = E_{D0} - \alpha T \frac{E_{D0}}{E_{c0}}$ ;

$E_{A0}$  = Akzeptorniveau bei 0 °K [eV];  $E_A = E_{A0} - \alpha T \frac{E_{A0}}{E_{c0}}$ ;

$n_D$  = Donatorkonzentration [ $\text{cm}^{-3}$ ];

$n_A$  = Akzeptorkonzentration [ $\text{cm}^{-3}$ ].

Tab. 1. Neutralitätsgleichung.

Zur Ermittlung von  $E_F$  ist man daher auf graphische Verfahren<sup>2</sup> angewiesen, die jedoch meistens sehr zeitraubend, nicht sehr genau und nur für eine

<sup>2</sup> K. LEHOVEC u. H. KEDESZY, J. Appl. Phys. 22, 65 [1951]. — W. SHOCKLEY, Electrons and Holes, van Nostrand, London 1956. — E. MOOSER, Z. angew. Math. Phys. 4, 433 [1953]. — E. SPENKE, Elektron. Halbleiter, Springer-Verlag, Berlin 1956.



bestimmte Temperatur oder für einen engeren Temperaturbereich die Lage des FERMI-Niveaus liefern.

Mit Hilfe eines elektronischen Digital-Rechners ausreichender Kapazität läßt sich jedoch ein Rechenverfahren aufstellen, mit dem sich für beliebige Halbleiter und beliebige Parametervariationen das FERMI-Niveau sehr schnell und mit großer Genauigkeit berechnen läßt, ohne daß zum Beispiel für die verschiedenen Halbleiter verschiedene Kurvenscharen anzufertigen sind.

Abb. 1. zeigt die wesentlichen Schritte des Rechenprogramms. Nach Eingabe der Materialkonstanten und der Parameter nimmt der Rechner zunächst mit vorgegebenen, zweckmäßig gewählten Anfangswerten eine Abschätzung der maßgeblichen Größen vor und berechnet entsprechend die einzelnen Summanden der Neutralitätsgleichung entweder nach „FERMI“, nach „BOLTZMANN“ oder mit Hilfe einer Entwicklung. Berechnung nach „FERMI“ bedeutet, daß die FERMI-Integrale bzw. die FERMI-Verteilung benutzt werden, Berechnung nach „BOLTZMANN“, daß die FERMI-Statistik durch die BOLTZMANN-Statistik ersetzt werden darf. Für die analytisch nicht erfaßbaren FERMI-Integrale wurde dabei eine TSCHEBYSCHEFF-Approximation durch Polynome 5. Grades eingesetzt<sup>3</sup>.

Durch die Abschätzung in den vier Gliedern der Neutralitätsgleichung wird gleichzeitig der weitere Rechenverlauf für  $f(E_F) = 0$  festgelegt:

a) Liegt  $E_F$  hinreichend weit (z. B.  $10 kT$ ) von sämtlichen Niveaus und Bandkanten entfernt, so besteht die Möglichkeit,  $f(E_F)$  in eine quadratische Gleichung umzuformen. Um die universelle Anwendbarkeit des Verfahrens und die hohe Rechengenauigkeit zu erreichen, muß eine vollständige Fallunterscheidung durchgeführt werden. Für tiefe Temperaturen können z. B. sehr große Exponenten auftreten, die beim Maschinen-Rechnen Schwierigkeiten bereiten und einen Übergang zum logarithmischen Rechnen erforderlich machen.

b) Im allgemeinen ist man jedoch zur Approximation der Nullstelle für  $f(E_F)$  gezwungen. Während der Iteration durch Addition oder Subtraktion einer vorgegebenen Schrittweite  $\Delta E_F$  zum jeweiligen Wert  $E_F$  bis zum Vorzeichenwechsel von  $f(E_F)$ , und

während der nachfolgenden Einschachtelung der Nullstelle durch Intervallhalbierung durchläuft der Rechenvorgang wiederholt die obere Schleife in Abb. 1, bis die Nullstelle mit der vorgegebenen Genauigkeit approximiert ist.

Sowohl mit dem aus der quadratischen Gleichung als auch mit dem aus der Approximation erzielten  $E_{F0}$  wird eine Nachprüfung vorgenommen. Nach Bestätigung des Wertes  $E_{F0}$  wird für Fehlerabschätzungen durch Bildung des Differenzen-Quotienten die Steigung von  $f(E_F)$  bei  $E_{F0}$  näherungsweise bestimmt.

Die Rechenmaschine gibt schließlich neben der Temperatur und dem berechneten FERMI-Niveau ( $E_{F0}$ ) die Werte der Glieder der Neutralitätsgleichung aus und wiederholt den gesamten Rechenvorgang für die nächste Parameterkombination.

Die nachfolgenden Abbildungen bringen einige Beispiele für den mit diesem Verfahren berechneten Verlauf des FERMI-Niveaus in Abhängigkeit von der Temperatur.

Abb. 2 zeigt für Silicium verschiedener Dotierung die Lage des FERMI-Niveaus zwischen  $50^\circ\text{K}$  und  $1700^\circ\text{K}$ . Die angegebenen Materialkonstanten stellen aus der Literatur entnommene mittlere Werte dar. Neben den temperaturabhängigen Lagen der Niveaus und der Leitungsbandkante sind die Grenzen für einen Abstand von  $10 kT$  von den Niveaus eingetragen (für Zweig a in Abb. 1).

Um die Auswirkung der Kompensation auf dem Kurvenverlauf zu zeigen, ist unter der Kurve für  $n_D = 10^{14} \text{ cm}^{-3}$  die zugehörige Kurve mit 50% Kompensation punktiert eingezeichnet.

Die Verhältnisse bei Temperaturen unter  $25^\circ\text{K}$  sind in Abb. 3 und 4 wiedergegeben. Es sei noch vermerkt, daß bei gleichem Kompensationsgrad die Kurven für die verschiedenen Dotierungskonzentrationen unterhalb einer bestimmten Temperatur in eine gemeinsame Gerade einlaufen [z. B. die Kurven  $(10^{12}/10^6)$ ,  $(10^6/10^4)$ ,  $(1/10^{-6})$  in Abb. 3]. Durch Verlängerung der für eine bestimmte Dotierung (z. B.  $n_D = 1$ ) berechneten Geraden verschiedener Kompensation läßt sich auch für andere Konzentrationen der Verlauf des FERMI-Niveaus unterhalb der Verzweigungspunkte interpolieren.

<sup>3</sup> Die Koeffizienten wurden mit einem von WERNER, RAYMANN und KLEIN aufgestellten Programm nach dem sogenannten REMES-Algorithmus ebenfalls mit der elektronischen Rechenmaschine bestimmt. Die Genauigkeit der Approximation

der FERMI-Integrale ist im ganzen Bereich von  $-\infty$  bis  $+\infty$  besser als  $5 \cdot 10^{-4}$ . Vgl. hierzu H. WERNER u. G. RAYMANN, Mathem. of Comput. (Baltimore), z. Veröffentlichung eingereicht; E. REMES, C. R. Acad. Sci., Paris **199**, 337 [1934].

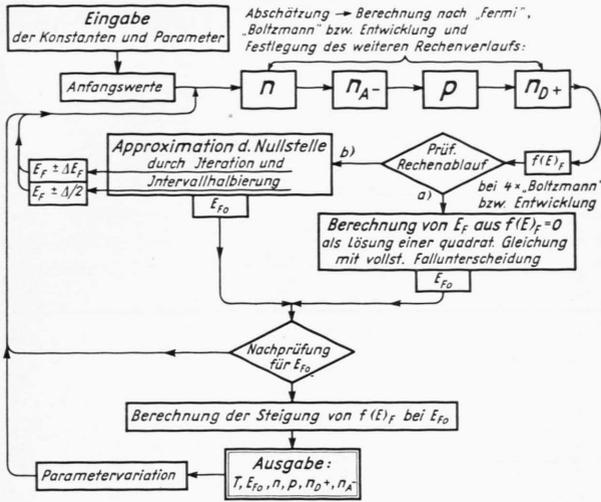


Abb. 1. Programmschema zur Berechnung des FERMI-Niveaus.

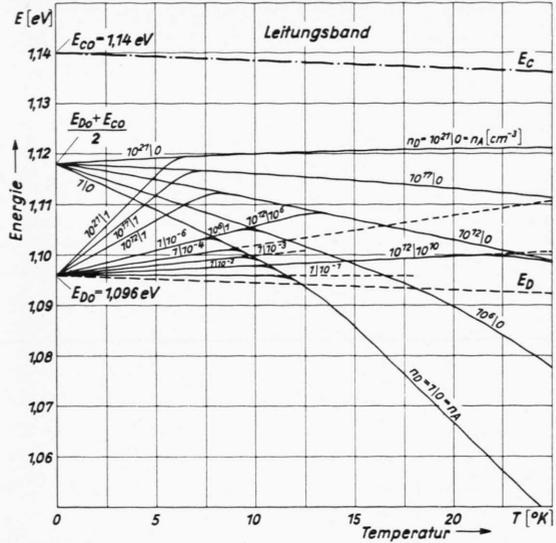


Abb. 3. Lage des FERMI-Niveaus für Silicium. Materialkonstanten wie bei Abb. 2.

Abb. 2. Lage des FERMI-Niveaus für Silicium.

$E_v \equiv 0$ ;  
 $E_{c0} = 1,14$  [eV];  
 $\alpha = 1,56 \cdot 10^{-4}$  [eV/°K];  
 $N_c = 5,57 \cdot 10^{15} T^{3/2}$ ;  
 $N_v = 2,191 \cdot 10^{15} T^{3/2}$ ;  
 $E_{D0} = 1,096$  [eV] (Phosphor);  
 $g_D = \frac{1}{2}$ ;  
 $E_{A0} = 0,045$  [eV] (Bor);  
 $g_A = 2$ ;  
 $\dots \dots n_D = 10^{14}$ ;  
 $n_A = 5 \cdot 10^{13}$ ;  
 --- Abstand  $10 k T$  von den Störniveaus.

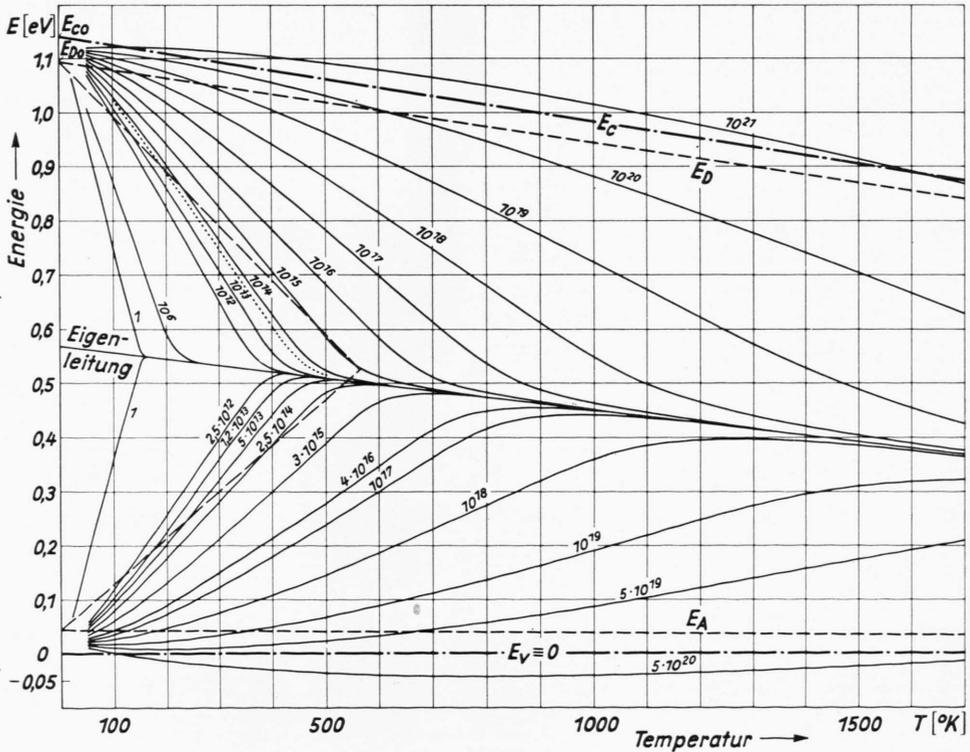


Abb. 4 zeigt die gleichen Verhältnisse für die untere Bandhälfte und den Verlauf des FERMI-Niveaus im Bereich der Bandmitte.

In Abb. 5 ist für den Fall der vertauschten Lage der Störstellenniveaus ein Beispiel gerechnet. Bei sonst gleichen Materialkonstanten wie in Abb. 2 ist

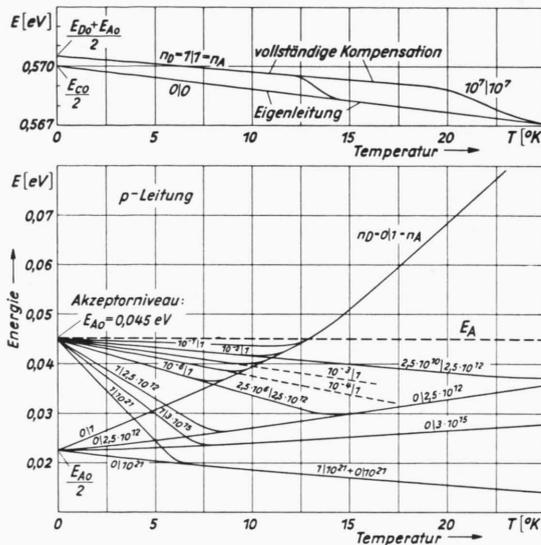


Abb. 4. Lage des FERMI-Niveaus für Silicium. Materialkonstanten wie bei Abb. 2.

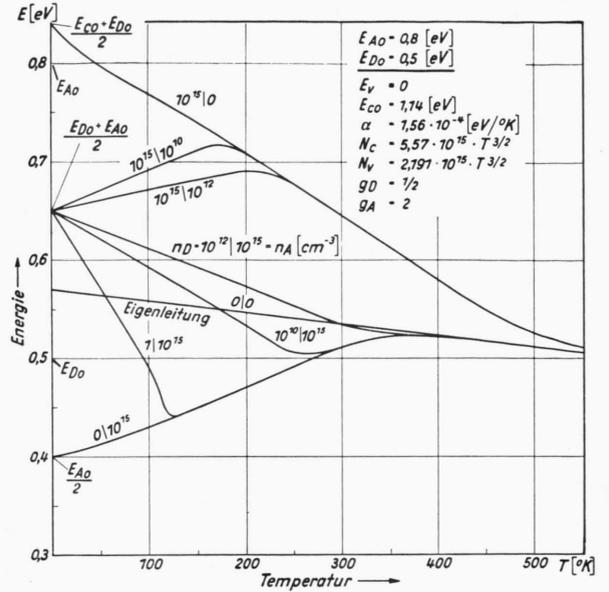


Abb. 5. FERMI-Niveau für Silicium mit vertauschten Störniveaulagen.

das Akzeptorniveau in der oberen Bandhälfte bei 0,8 eV und das Donatorniveau in der unteren Bandhälfte bei 0,5 eV angenommen<sup>4</sup>. Für den Verlauf des FERMI-Niveaus in Abb. 5 ist neben der Vertauschung der Niveaus noch der relativ große Unterschied (25%) in der Aktivierungsenergie der beiden Störstellenarten verantwortlich.

Wie im vorher berechneten Fall mit normaler Lage der Niveaus (s. Abb. 3 und 4) laufen die Kurven bei Anwesenheit nur einer Störstellenart für  $T \rightarrow 0$  in die Mitte zwischen Niveau und Bandkante. Dagegen treffen sich bei vertauschten Niveaus sämtliche Kurven bei vorliegender Kompensation in einem Punkt in der Mitte zwischen beiden Niveaus, sowohl bei Donatoren- als auch bei Akzeptorenüberschuß. Da bei sehr tiefen Temperaturen die Donatoren fast alle besetzt sind und bei steigender Temperatur zunächst Elektronen an die hochliegenden Akzeptoren abgeben, wird der Gleichgewichtszustand in diesem Falle vorwiegend durch die Besetzung in den Störniveaus geregelt, d. h. das FERMI-Niveau läuft für  $T \rightarrow 0$  genau in die Mitte der beiden Niveaus. Liegt dieser Punkt auf Grund unterschiedlicher Aktivierungsenergien wesentlich über oder unter der Eigenleitungsgeraden, so erfolgt notwendig

für einen Leitungstyp ein unter Umständen steiler Übergang von der einen in die andere Bandhälfte, und zwar bei relativ hohen Temperaturen und insbesondere bei geringer Kompensation<sup>5</sup>.

Den Einfluß stark unterschiedlicher Aktivierungsenergien bei normaler Lage der Störniveaus und bei hohen Dotierungskonzentrationen zeigt in Abb. 6 der als Beispiel für Wismut-Tellurid berechnete Verlauf des FERMI-Niveaus. Als Lage der Störstellen-niveaus wurden zwei Kombinationen eingesetzt:

$$E_D^1 = 0,13; E_A^1 = 0,01 \text{ und } E_D^2 = 0,14; E_A^2 = 0,02 \text{ (eV).}$$

Für die erste Kombination erfolgt nur bei Donatorenüberschuß und gleichzeitig sehr hohem Kompensationsgrad ein Wechsel des FERMI-Niveaus von der oberen in die untere Bandhälfte. Dieser Effekt ist um so ausgeprägter, je größer der Kompensationsgrad ist – im Gegensatz zu der Auswirkung unterschiedlicher Aktivierungsenergien bei vertauschten Niveaulagen (vgl. Abb. 5). Für einen Kompensationsgrad  $K \leq 90\%$  (Kurven I und V) tritt der Wechsel nicht mehr auf.

Ein solcher Verlauf des FERMI-Niveaus läßt sich damit erklären, daß mit steigender Temperatur bevorzugt das Niveau mit der kleineren Aktivierungs-

<sup>4</sup> Für einige Metalle sind bisher ähnlich liegende, dotierende Niveaus in Silicium festgestellt worden; siehe z. B. E. M. CONWELL, Proc. Inst. Radio Engrs 45, 1281 [1958].

<sup>5</sup> Hier z. B. für Kurve ( $n_D=10^{10}, n_A=10^{15}$ ) bei ca. 175 °K.

energie ionisiert wird, d. h. in diesem Falle  $E_A^1$ . Daher geht die n-Leitung bei genügend hoher Akzeptorenkonzentration schließlich in p-Leitung über. Mit höherer Temperatur macht sich die zunehmende Eigenleitung bemerkbar und führt das FERMI-Niveau zur Bandmitte zurück.

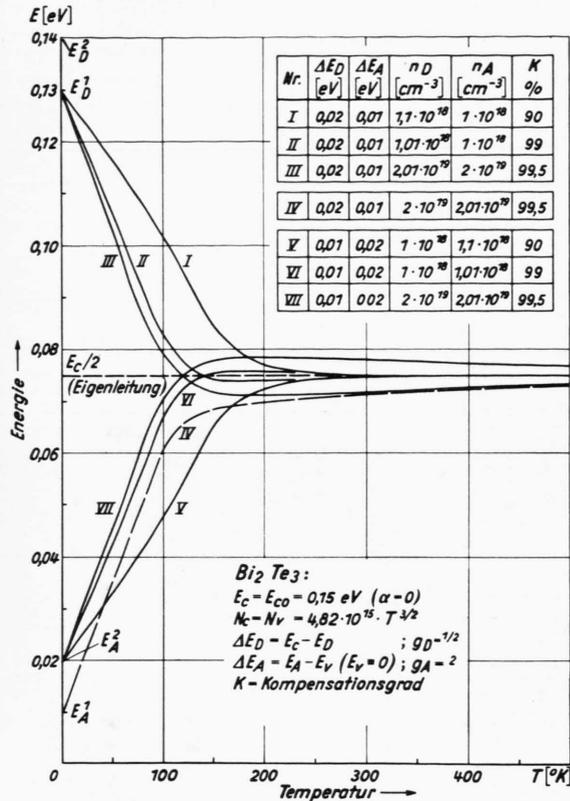


Abb. 6. FERMI-Niveau für Wismut-Tellurid, Bandabstand nach I. G. AUSTIN u. A. SHEARD, J. El. Contr. 3, 236 [1957].

Für die untere Bandhälfte gilt das Entsprechende bei Umkehrung der Konzentrationen und des Verhältnisses der Aktivierungsenergien (zweite Kombination der Niveaulagen, Kurven V – VII). Die ge-

strichelte Kurve in der unteren Bandhälfte zeigt ein normales Verhalten, da nur die Störstellenkonzentrationen vertauscht wurden.

Die Größe des möglichen Fehlers bei der Berechnung des FERMI-Niveaus mit dem vorliegenden Verfahren ist ausschließlich bedingt durch die Genauigkeit der Approximation der FERMI-Integrale, da sich die Abschätzungen und Entwicklungen so abfassen lassen, daß die übrigen Größen der Neutralitätsgleichung hinreichend genau berechnet werden.

Eine eingehende Fehlerabschätzung liefert mit Hilfe der gleichzeitig zu  $E_{F0}$  berechneten Steigung der Funktion  $f(E_F)$  einen maximalen relativen Fehler von etwa  $1 \cdot 10^{-4}$  für die angegebenen Werte des FERMI-Niveaus. Für sehr hohe Dotierungskonzentrationen (d. h. bei Entartung) und gleichzeitig tiefe Temperaturen kann der relative Fehler auf ca.  $5 \cdot 10^{-4}$  ansteigen.

Man ist daher bei der Anwendung des Rechenverfahrens<sup>6</sup> sicher, daß außer den Ungenauigkeiten bei der Angabe der Materialkonstanten keine zusätzlichen Fehler in die Berechnung des FERMI-Niveaus eingehen.

Zu dem Verfahren selbst sei abschließend noch erwähnt, daß es sehr leicht auf mehrere Donator- und Akzeptorniveaus erweitert werden kann. Außerdem läßt sich die Störbandleitung bei sehr tiefen Temperaturen gegebenenfalls durch einen Übergang von diskreten Störniveaus zu Störbändern berücksichtigen. Auch Effekte zweiter Ordnung können ohne weiteres in die Rechnung eingebaut werden, so z. B. die Temperaturabhängigkeit der effektiven Massen oder die Änderung der Aktivierungsenergien der Störstellen mit der Störstellenkonzentration<sup>7</sup>.

Herrn Direktor Dr. Dr. E. h. STEIMEL danke ich sehr für sein förderndes Interesse. Den Herren des Rechenzentrums der AEG in Frankfurt möchte ich für ihre Unterstützung, den Herren Dr. SCHULTZ und Dr. BÖTTGER für Diskussionen und Hinweise und Frau DYMKE für ihre unermüdliche Zeichenhilfe recht herzlich danken.

<sup>6</sup> Das vorliegende Rechenprogramm steht Interessenten zur Verfügung.

<sup>7</sup> J. F. BARNES u. R. H. TREGOLD, Proc. Phys. Soc., Lond. 78, 716 [1961].