

Dissertation von

Herbert Krömer
Nobelpreisträger für Physik 2000

Zur Theorie des Germaniumgleichrichters
und des Transistors

Göttingen 1952

Zur Theorie des Germaniumgleichrichters und des Transistors*.

Von

HERBERT KRÖMER.

Mit 10 Figuren im Text.

(Eingegangen am 1. Dezember 1952.)

Bei den in den Randschichten von Halbleiter-Metall-Kontakten herrschenden hohen Feldstärken sind für den Transport der Elektronen und Defektelektronen („Löcher“) nicht die normalen Beweglichkeiten maßgebend, die gültig sind, wenn der Potentialabfall längs einer mittleren freien Weglänge klein ist gegen die thermische Energie kT .

In stärkeren Feldern gelangen die Ladungsträger mit merklicher Wahrscheinlichkeit ins Bandinnere, im Grenzfall sehr starker Felder oszilliert ein Teilchen zwischen zwei Stößen mehrere Male im Band hin und her und kommt dadurch langsamer vorwärts („Stauereffekt“). Die Beweglichkeit nimmt dann so ab, daß die Teilchendichte proportional zu Feldstärke und Stromdichte ansteigt. Der Wert der Proportionalitätskonstanten („Staukonstante“) wird für die Löcher zu $2 \cdot 10^3 \text{ Watt}^{-1}$ abgeschätzt; für die Elektronen dürfte er erheblich kleiner sein.

Die Dichte der gestauten Teilchen kann von gleicher Größenordnung wie die der Störstellen werden. Das Randschichtpotential weicht dann erheblich von der einfachen SCHOTTKYSchen Parabelgestalt ab. Mit dem abgeänderten Potential werden Kennlinien von Ge-Transistoren und -Dioden berechnet. Bei geeigneter Wahl der eingehenden Parameter ergeben sich bei den Transistoren hohe Werte für den Stromverstärkungsfaktor, bei den Dioden das experimentell beobachtete Umbiegen der Sperrkennlinien.

I. Einleitung.

Nach SCHOTTKY [13] bäumen sich an der Grenzfläche zwischen Germanium und einem Metall die Bänder des Halbleiters (künftig abgekürzt HL) ohne angelegte Spannung um das sog. Diffusionspotential V_D , bei angelegter Spannung U um $eU + V_D$ auf. Es entsteht eine die Elektronenbewegung hemmende, von U abhängige Potentialschwelle, deren Höhe über der FERMI-Kante (künftig abgekürzt FK) des Metalls

$$e\Phi_0 = V_D + E_L \quad (1)$$

ist, unabhängig von U (Fig. 1).

Infolge der Bildkraftanziehung des Metalls auf die Elektronen wird die Spitze dieser Schwelle um

$$\Delta V = \sqrt{\frac{e^3 F}{\eta}} \quad (2)$$

* Auszug aus einer Göttinger Dissertation



math.-nat. 1953

abgeflacht (Fig. 2), wobei F die Feldstärke am Ende der Randschicht (Randfeldstärke) bedeutet. Anstatt der statischen Dielektrizitätskonstante ϵ des HL wurde dabei eine „dynamische“ DK $\eta < \epsilon$ eingeführt. Denn da die Elektronen eine Geschwindigkeit von 10^7 bis 10^8 cm/sec haben, und da der Abstand des Maximums der abgerundeten Schwelle von der Grenzfläche

$$x_s = \frac{1}{2} \sqrt{\frac{e}{\eta F}}$$

zwischen einigen 10^{-7} und 10^{-6} cm liegt, kann man bei einer Vorbeiflugzeit von 10^{-13} bis 10^{-15} sec wohl kaum mit dem statischen Wert der DK von $\epsilon = 16$ [5] rechnen. Auch hinkt die Gitterpolarisation in der Phase nach, was die wirksame DK abermals herabsetzt.

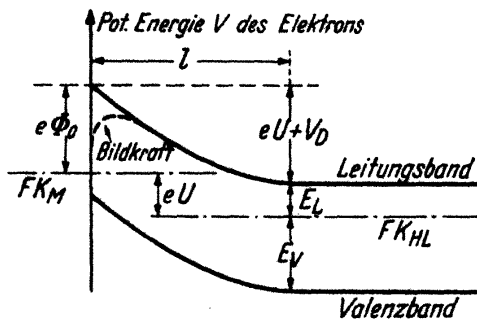


Fig. 1. Randschichtpotential. E_L und E_V sind die Abstände des Leitungs- und Valenzbandes des Halbleiters von der FERMI-Kante.

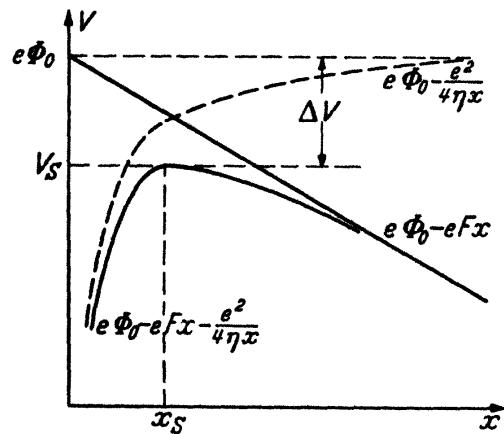


Fig. 2. Abbau der Potentialschwelle durch die Bildkraft.

Unter diesen Umständen muß die Gl. (2) als eine erste Näherung und η als ein Parameter angesehen werden, über den sich zur Zeit theoretisch noch nichts sagen läßt, und der daher vorerst experimentell bestimmt werden muß. Messungen von S. BENZER [3] an Ge-Gleichrichtern lassen sich am besten mit $\eta \approx 1$ beschreiben (s. unten). Ähnliches fand auch W. OLDEKOP [10] an Selengleichrichtern. Wir rechnen daher im folgenden mit $\eta = 1$.

Mit Berücksichtigung der Bildkraft wird die Gesamthöhe der von den Elektronen zu überwindenden Potentialschwelle

$$V_s = e\Phi_0 - \Delta V = V_D + E_L - \sqrt{\frac{e^3 F}{\eta}}. \quad (3)$$

Nach der Diodentheorie des Kristallgleichrichters [4], [18] beträgt dann die Dichte des in Sperrichtung für $eU \gg kT$ fließenden Elektronenstromes

$$j_n = AT^2 e^{-\frac{V_s}{kT}} \quad \text{mit} \quad A = \frac{4\pi e m k^2}{h^3}. \quad (4)$$

Im Fall der einfachen Erschöpfungsrandschicht mit der Störstellenkonzentration N_0 gilt nach SCHOTTKY [13]

$$F = \sqrt{\frac{8\pi N_0}{\epsilon} (eU + V_D)}. \quad (5)$$

Daraus folgt mit (3) und (4) für $eU \gg V_D$

$$j_n = AT^2 \exp \frac{1}{kT} \left[-e\Phi_0 + \sqrt[4]{\frac{8\pi e^7 N_0}{\epsilon \eta^2}} \sqrt[4]{U} \right].$$

Solche im $(\lg j/j_0, U)$ -Diagramm geradlinige Kennlinien wurden z.B. von SEILER [12] an Si-Detektoren gefunden. Aus der Steigung der Geraden läßt sich rückwärts N_0/η^2 bestimmen. SEILER, der mit $\eta = \epsilon$ rechnete, fand so Werte für N_0 , die bis zu einer Zehnerpotenz über anderweitig bestimmten lagen. Wir sehen darin eine Bestätigung unserer Vermutung $\eta < \epsilon$.

Bei der Herleitung von (5) wird die Annahme gemacht, daß die Dichte der für den Stromtransport verantwortlichen Ladungsträger, die im HL-Innern mit der der ionisierten Störstellen übereinstimmt, gegen die Randschicht hin auf einer kurzen Strecke praktisch völlig abklingt, so daß in der Randschicht nur mit der Ladungsdichte der Störstellenionen zu rechnen ist. Wie im folgenden gezeigt wird, ist diese Annahme nur bei nicht zu großen angelegten Spannungen richtig. Wird jedoch die Randfeldstärke hinreichend groß, so wird die effektive Trägerbeweglichkeit in der Randschicht so stark herabgesetzt, daß die Trägerdichte stark ansteigt und schließlich von der gleichen Größenordnung wie die der Störstellen werden kann.

Mit der hierdurch bedingten abgeänderten Raumladungsdichte wird die Berechnung der Erschöpfungsrandschicht wiederholt und anschließend auf die Berechnung von Dioden- und Transistorkennlinien angewandt werden.

II. Der Staueffekt.

II, 1. Transportgesetze in schwachen und starken Feldern.

Die Transportgesetze für die Ladungsträger hängen davon ab, wie stark sich die potentielle Energie V der Elektronen längs einer mittleren freien Weglänge λ ändert. Letztere ergibt sich aus der Beweglichkeit b gemäß

$$\lambda = \frac{3}{4} \frac{\sqrt{2\pi m kT}}{e} \cdot b \quad \dagger.$$

Bei Zimmertemperatur folgt daraus mit

$$b_n = 3600 \text{ cm}^2 \text{ V}^{-1} \text{ sec}^{-1}, \quad b_p = 1700 \text{ cm}^2 \text{ V}^{-1} \text{ sec}^{-1} \quad [11]:$$

$$\lambda_n = 2,5 \cdot 10^{-5} \text{ cm}, \quad \lambda_p = 1,2 \cdot 10^{-5} \text{ cm}.$$

† Siehe z.B. W. SHOCKLEY [15], S. 277.

Für $|\text{grad } V| = eF \ll \frac{kT}{\lambda}$ gilt:

$$\text{Elektronenstromdichte } \vec{j}_n = b_n [n \cdot \text{grad } V + kT \cdot \text{grad } n], \quad (6a)$$

$$\text{Löcherstromdichte } \vec{j}_p = b_p [p \cdot \text{grad } V - kT \cdot \text{grad } p] \quad (6b)$$

(n bzw. p = Elektronen- bzw. Löcherdichten).

Daraus folgen die Ladungsdichten

$$\varrho_n = -\frac{1}{b_n F} \cdot |\vec{j}_n - b_n kT \text{ grad } n|; \quad \varrho_p = +\frac{1}{b_p F} |\vec{j}_p + b_p kT \text{ grad } p|. \quad (7)$$

Bei festgehaltenem Strom nehmen also die Ladungsdichten mit zunehmendem Feld ab.

Für $eF \gtrsim \frac{kT}{\lambda}$ werden die Verhältnisse recht unübersichtlich und ergeben erst im Grenzfall sehr starker Felder wieder einfache Gesetze. Wird

$$|\text{grad } V| \geq eF_s = \frac{2B}{\lambda}, \quad (8)$$

wo B die Breite des Valenz- bzw. Leitungsbandes ist, so laufen die Ladungsträger im Feld nicht mehr davon, sondern oszillieren vor dem nächsten Stoß zwischen den im Feld gekippten Bandkanten hin und her (Fig. 3), halten sich also länger am gleichen Ort auf, so daß die Teilchendichte wieder ansteigt.

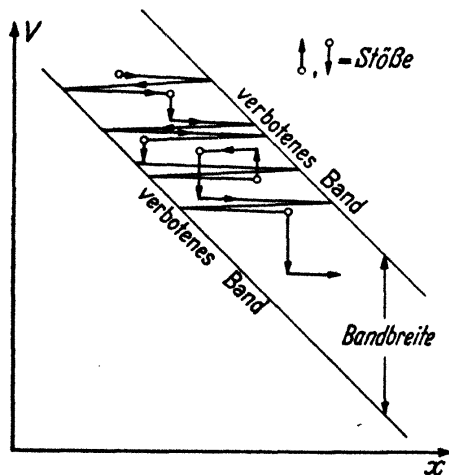


Fig. 3. Elektronenbewegung in sehr starken Feldern.

Über die Breiten B_V und B_L des Valenz- und Leitungsbandes bei Ge sind uns keine experimentellen Angaben bekannt, an theoretischen Untersuchungen nur eine Näherungsrechnung von H. MÜLLER [9] nach einem vereinfachten WIGNER-SEITZ-Verfahren, deren quantitative Brauchbarkeit zweifelhaft ist.

Es scheint jedoch sicher zu sein, daß das Leitungsband erheblich breiter als das verbotene Band ist und mit mehreren höheren Bändern überlappt, während das Valenzband eine Breite von gleicher Größenordnung wie das verbotene Band hat. MÜLLER findet $B_V < E_B$, mit einem allerdings zu großen E_B . Wir möchten annehmen, daß B_V in der Größenordnung von etwa 1 eV liegt, B_L aber $\gg 1$ eV ist. Mit $B_V = 1$ eV und $\lambda_p = 1,2 \cdot 10^{-5}$ cm geht (8) über in $F \geq 1,7 \cdot 10^5$ cm $^{-1}$, ein Wert, der bequem erreicht werden kann. Wir untersuchen diesen Fall näher.

II, 2. Berechnung der Staukonstanten.

Die Raumladungsdichte der gestauten Teilchen beträgt offenbar

$$\varrho_s = \pm \frac{j \cdot e F}{(\Delta E / \tau)} \quad \text{für } \left\{ \begin{array}{l} \text{Löcher} \\ \text{Elektronen} \end{array} \right\}.$$

Dabei ist ΔE der Energieverlust pro Stoß und τ die Zeit zwischen zwei Stößen. Da ΔE und τ statistisch unabhängig sind, gilt

$$\varrho_s = \pm \frac{j \cdot e \cdot F}{\Delta E \cdot (1/\tau)}. \quad (9)$$

Nach der Theorie der Elektronenstreuung an Gitterwellen lassen sich $\overline{\Delta E}$ und $\overline{(1/\tau)}$ abschätzen.

II, 2a. Danach ändert ein Elektron oder ein Loch bei Streuung an einer Gitterwelle der Frequenz ν seine Energie um $\pm h\nu$. Die Wahrscheinlichkeit für Energieverlust verhält sich zu der für Energiegewinn wie $\frac{N+1}{N}$, wo N die mittlere Quantenzahl ist, bis zu der diese Gitterwelle angeregt ist. Auf $2N+1$ Stöße kommt also ein Energieverlust der Größe $h\nu$, also durchschnittlicher Verlust pro ν -Stoß

$$\Delta E_\nu = \frac{h\nu}{2N+1}. \quad (10)$$

Die Wahrscheinlichkeiten für Streuung an zwei Wellen mit verschiedenen Frequenzen ν_1 und ν_2 verhalten sich angenähert wie

$$\frac{W_1}{W_2} = \frac{h\nu_1(2N_1+1)}{h\nu_2(2N_2+1)} \dagger. \quad (11)$$

Führt man noch die Zahl $z(\nu)$ der Gitterschwingungen pro Frequenzintervall 1 ein, so ist wegen (10) und (11)

$$\overline{\Delta E} = \frac{\int \Delta E_\nu \cdot h\nu(2N_\nu+1) z(\nu) d\nu}{\int h\nu(2N_\nu+1) z(\nu) d\nu} = \frac{\int (h\nu)^2 z(\nu) d\nu}{\int h\nu \cdot (2N_\nu+1) z(\nu) d\nu}. \quad (12)$$

Für $z(\nu)$ setzen wir das DEBYESche Spektrum

$$z(\nu) = \begin{cases} \text{const} \cdot \nu^2 & \text{für } \nu \leq \nu_g = \frac{k\Theta}{h} \\ 0 & \text{für } \nu > \nu_g \end{cases}$$

ein (Θ = DEBYE-Temperatur). Beschränkungen auf Temperaturen $T > \Theta$ gibt

$$2N_\nu + 1 = 1 + \frac{2}{e^{\frac{h\nu}{kT}} - 1} \approx 2 \frac{kT}{h\nu}. \quad (13)$$

† Gl. (10) und (11) folgen aus Gl. (34,33) und (34,36) bei SOMMERFELD und BETHE [17] mit $C_1 = C_2$ und $q_1 : q_2 = \nu_1 : \nu_2$.

Damit wird

$$\overline{\Delta E} = \frac{3}{10} \frac{(k\Theta)^2}{kT}. \quad (14)$$

II, 2b. Wenn Gl. (8) erfüllt ist, bewegt sich das Elektron bzw. Loch im Mittel durch das ganze Band hin und her, ehe es das nächste Mal stößt. In den Impulsraum übertragen heißt dies, daß es sich *gleichförmig* quer durch die gesamte, zu diesem Band gehörige BRILLOUIN-Zone bewegt (Fig. 4). Da die Quantenzustände im Impulsraum mit konstanter Dichte verteilt sind, hält es sich in jedem Zustand gleich lange auf, und die Stoßwahrscheinlichkeit $(1/\tau)$ ergibt sich einfach als Mittelwert über die ganze BRILLOUIN-Zone. Unter Einführung der freien Weglänge λ und der Gruppengeschwindigkeit $v = \text{grad}_P E(P)$ heißt das:

$$\left(\frac{1}{\tau}\right) = \frac{1}{V_P} \iiint \frac{|\text{grad}_P E(P)|}{\lambda(P)} dP_x dP_y dP_z. \quad (15)$$

Für die weitere Auswertung müssen wir die schwerwiegende Annahme machen, daß λ nicht nur am Bandrande, sondern auch im Bandinneren von der Geschwindigkeit und damit von P unabhängig ist. Dann ist

$$\left(\frac{1}{\tau}\right) = \frac{1}{\lambda} |\overline{\text{grad}_P E}|. \quad (16)$$

Da uns die genaue Gestalt der Flächen konstanter Energie im P -Raum unbekannt ist, können wir diesen Mittelwert nicht exakt ausrechnen. Wir erhalten aber mindestens die richtige Größenordnung, wenn wir den Gradienten in (16) durch die Größe

$$\begin{aligned} & \frac{\text{Energieunterschied innerhalb der Zelle}}{\text{Halbe Kantenlänge der Zelle}} \\ &= \frac{B}{h/2a} \end{aligned}$$

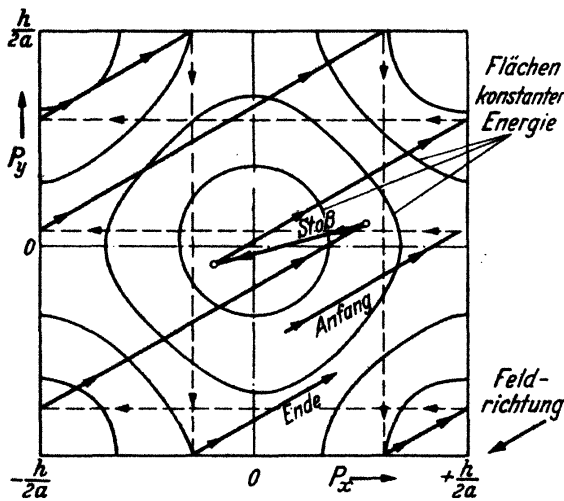


Fig. 4. Schnitt durch den reduzierten Impulsraum eines kubischen Kristalls (a Gitterkonstante). Bewegung eines Elektrons im äußeren Feld.

ersetzen (a = Gitterkonstante). Damit wird

$$\left(\frac{1}{\tau}\right) = \frac{a}{h} \cdot \frac{2B}{\lambda} \cdot g, \quad (17)$$

wo g ein von der Zellengeometrie und der Gestalt der Energieflächen abhängiger Zahlenfaktor der Größenordnung eins ist. Wir rechnen mit $g=1$ weiter; dieser Wert gilt exakt z.B. für eindimensionale Gitter.

II, 2c. Einsetzen von (14) und (17) in (9) gibt

$$\varrho = j \cdot e \cdot F \cdot \frac{10}{3} \cdot \frac{h}{a} \cdot \frac{kT}{(k\Theta)^2} \cdot \frac{\lambda}{2B} = j \cdot e F \cdot s, \quad (18)$$

wobei die „Staukonstante“ s durch

$$s = \frac{1}{\Delta E \cdot (1/\tau)} = \frac{10}{3} \cdot \frac{h}{a} \cdot \frac{kT}{(k\Theta)^2} \cdot \frac{1}{eF_s} \quad (19)$$

definiert ist.

Die bei Ge für Defektelektronen einzusetzenden Werte ($B_V = 1$ eV, $\lambda_p = 1,2 \cdot 10^{-5}$ cm; $\Theta = 400^\circ \text{ K}^1$, $a = 5,6 \cdot 10^{-8}$ cm) geben bei Zimmertemperatur

$$F_{s,p} = 1,7 \cdot 10^5 \text{ V cm}^{-1}$$

$$s_p = 2 \cdot 10^8 \frac{\text{cm}^{-3}}{\text{V cm}^{-1} \text{ Amp cm}^{-2}} = 2 \cdot 10^8 \text{ W}^{-1}. \quad (20)$$

Die üblichen Punktkontakte haben Durchmesser von etwa 10^{-3} cm, die im Transistor fließenden Löcherströme sind von der Größenordnung 1 mA. Das gibt Stromdichten von einigen hundert $\text{Amp} \cdot \text{cm}^{-2}$. Da die auftretenden Feldstärken bei $10^5 \text{ V} \cdot \text{cm}^{-1}$ liegen, entstehen somit Staukonzentrationen von derselben Größenordnung wie die Störstellenkonzentrationen.

II, 3. Diskussion der gemachten Vernachlässigungen.

a) Gl. (11) gilt nicht streng, da die BLOCHSche Wechselwirkungskonstante $C^\dagger$, die in die Stoßwahrscheinlichkeit eingeht, vom Anfangs- und Endzustand nicht ganz unabhängig ist. Da sie jedoch in (12) im Zähler und Nenner auftritt, ist ihr Einfluß gering. Wäre z.B. W proportional zu $(h\nu)^{1+\alpha} \cdot (2N+1)$ mit $\alpha \neq 0$, so träte in (14) nur ein Faktor $\left(1 + \frac{\alpha}{3}\right) / \left(1 + \frac{\alpha}{5}\right)$ vor die rechte Seite, der von der Größenordnung eins ist.

b) Aus demselben Grunde ist die Benutzung des einfachen DEBYE'schen Spektrums unerheblich.

c) Die Staukonstante s ist nicht temperaturabhängig! Denn das Integral im Nenner von (12) ist bis auf Konstanten gerade die gesamte Stoßwahrscheinlichkeit, also proportional zu $\overline{(1/\tau)}$, so daß die T -Abhängigkeit herausfällt. Die Näherung (13) ermöglichte nur die einfache Berechnung (λ proportional zu $1/T$ für $T > \Theta$), der Zahlenwert (20) gilt, wenn überhaupt, so für alle Temperaturen.

d) Die Problematik von (18) liegt in der Annahme einer geschwindigkeits- und richtungsunabhängigen mittleren freien Weglänge. Diese Annahme ist am Bandrande richtig², im Inneren zweifelhaft. Eine

¹ Die angegebenen Werte schwanken stark. Wir schließen uns an HILL und PARKINSON [7] an.

[†] Vgl. Fußnote S. 439.

² Vgl. dazu SHOCKLEY [15], Kap. 17.

genaue Theorie dieses Falles liegt noch nicht vor und kann an dieser Stelle auch nicht gegeben werden. Hier bleibt daher eine Möglichkeit für spätere Korrekturen.

II, 4. Mittelstarke Felder.

In mittelstarken Feldern ($kT/\lambda < |\text{grad } V| < eF_s$) ist nicht mehr jede Stelle des Impulsraumes gleichwahrscheinlich. Die Teilchen ziehen sich mit abnehmendem Feld mehr und mehr auf die Bereiche geringster Energie, d.h. auf die Mitte des P -Raumes zurück, die bei Löchern dem oberen, bei Elektronen dem unteren Bandrand entsprechen. In (15) tritt dann ein gegen die Ränder abnehmender Gewichtungsfaktor mit unter das Integral, dessen genaue Form zur Zeit noch nicht angegeben werden kann.

Auch die Berechnung von $\overline{\Delta E}$ ist dann nicht mehr so einfach durchführbar, da unmittelbar am Bandrande kein Energieverlust mehr möglich ist.

Für schwache Felder ($eF \ll kT/\lambda$) muß wieder (7) herauskommen. Im mittleren Feldstärkenbereich sind wir auf eine plausible Interpolation zwischen (7) und (18) angewiesen. Als einfachste Möglichkeit bietet sich die Summe beider Ausdrücke, da jeder Teil im „falschen“ Bereich gegen Null geht. Berücksichtigt man beide Trägersorten, so ist dann die gesamte Raumladung

$$\varrho = eN_0 + \left\{ \frac{j_p + b_p \cdot kT \cdot \text{grad } p}{b_p} - \frac{j_n - b_n \cdot kT \cdot \text{grad } n}{b_n} \right\} \cdot \frac{1}{F} + \left\{ + e(s_p j_p - s_n j_n) \cdot F \right\} \quad (21)$$

Wir hatten nur s_p abgeschätzt; s_n dürfte wegen der größeren Bandbreite erheblich größer sein. In späteren Beispielen setzen wir willkürlich $s_n = 0,2 \cdot s_p$.

II, 5. Berechnung der Randfeldstärken.

II, 5a. Vernachlässigung des Staugliedes in (21) führt auf die normale Erschöpfungsrandschicht, von der SCHOTTKY [13] gezeigt hat, daß man den wahren Raumladungsverlauf durch eine konstante Ladungsdichte $\varrho = e \cdot N_0$ von endlicher Dicke l ersetzen kann.

Wir rechnen daher ebenfalls mit einer endlichen Dicke der Randschicht, setzen aber für die Ladungsdichte

$$\varrho = e \cdot N_0 + e s_p \cdot j_s \cdot F, \quad (21a)$$

wo

$$j_s = j_p - \frac{s_n}{s_p} \cdot j_n \quad (22)$$

ist.

Bei Flächenkontakten lautet die Poisson-Gleichung dann:

$$\frac{d^2 \varphi}{dx^2} = -\frac{4\pi e}{\varepsilon} \left[N_0 + s_p \cdot j_s \cdot \frac{d\varphi}{dx} \right]$$

mit den Randbedingungen

$$e[\varphi(l) - \varphi(0)] = eU + V_D; \quad \varphi'(l) = 0.$$

Man findet durch Elimination von l das implizite Gesetz für die Randfeldstärke F_R :

$$eU + V_D = \frac{\varepsilon}{4\pi} \cdot \frac{N_0}{(s_p \cdot j_s)^2} \left[\frac{s_p \cdot j_s \cdot F_R}{N_0} - \ln \left(1 + \frac{s_p \cdot j_s \cdot F_R}{N_0} \right) \right]. \quad (23)$$

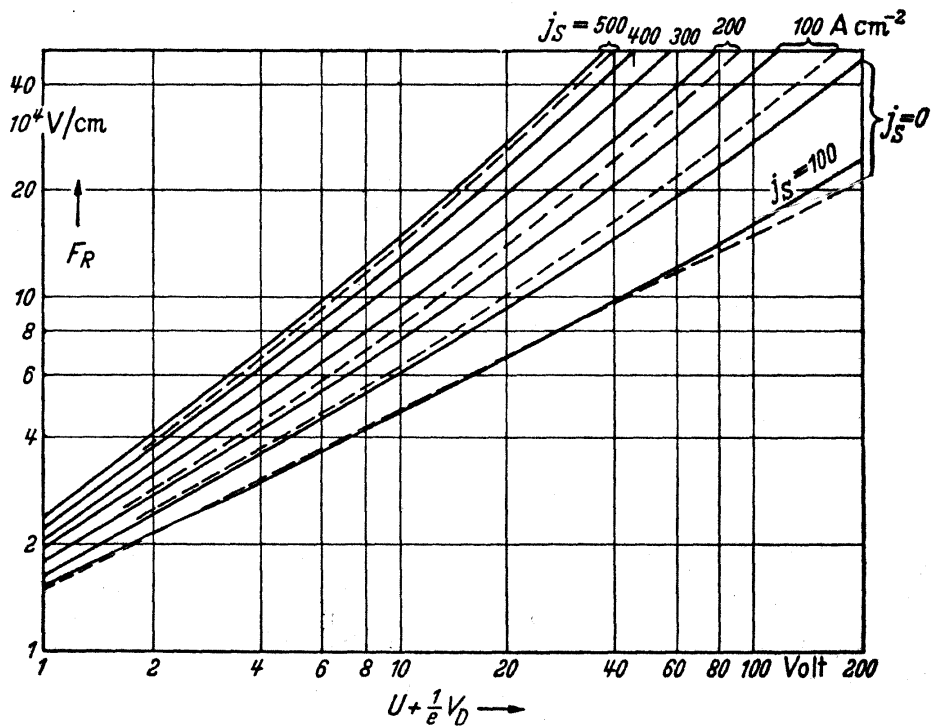


Fig. 5. Randfeldstärke für $N_0 = 10^{15} \text{ cm}^{-3}$ für Spitzenkontakte mit $r_0 = 10^{-3} \text{ cm}$ (—) und für Flächenkontakte (---).

II, 5b. Die uns mehr interessierenden Kontakte sind Spitzenkontakte von etwa 10^{-3} cm Durchmesser. Wir rechnen daher noch F_R für kugelförmige Kontakte (Radius r_0) mit kugelförmiger Randschicht (Radius R) aus.

Weil $j(r)$ nicht mehr räumlich konstant ist, sondern

$$j(r) = j \cdot \frac{r_0^2}{r^2} \quad \dagger$$

ist, lautet die Poisson-Gleichung jetzt

$$\frac{d^2 \varphi}{dr^2} + \frac{2}{r} \cdot \frac{d\varphi}{dr} = -\frac{4\pi e}{\varepsilon} \left[N_0 + s_p \cdot j_s \cdot r_0^2 \cdot \frac{1}{r^2} \frac{d\varphi}{dr} \right]$$

† Wenn j ohne Argument angegeben ist, bedeutet es künftig stets die Stromdichte in der Kontaktfläche.

mit den Randbedingungen

$$e[\varphi(R) - \varphi(r_0)] = eU + V_D; \quad \varphi'(R) = 0.$$

Dann wird der Zusammenhang zwischen U , F_R und R durch das Gleichungspaar

$$F_R = \frac{3\alpha}{r_0^2} \cdot e^{\frac{\beta}{r_0}} \cdot \int_{r_0}^R e^{-\frac{\beta}{s}} \cdot s^2 ds, \quad (24a)$$

$$eU + V_D = \frac{r_0^2}{\beta} \cdot e F_R - \frac{e\alpha}{\beta} (R^3 - r_0^3) \quad (24b)$$

gegeben, wobei

$$\alpha = \frac{4\pi e}{\varepsilon} \cdot \frac{N_0}{3} \quad \text{und} \quad \beta = \frac{4\pi e}{\varepsilon} \cdot s_p \cdot j_s \cdot r_0^2$$

ist.

Fig. 5 zeigt $F_R(U, j_s)$, wobei $s_p = 2 \cdot 10^8 \text{ W}^{-1}$ und $N_0 = 10^{15} \text{ cm}^{-3}$ gesetzt wurde, für einen Spitzenkontakt mit $r_0 = 10^{-3} \text{ cm}$. Zum Vergleich sind (gestrichelt) einige Kurven für Flächenkontakte ebenfalls eingetragen.

III. Anwendung auf die Kennlinien.

Durch Kombination von Gl. (3) und (4) und Fig. 5 läßt sich die Schar der Kennlinien mit j_s als Scharparameter bestimmen.

III, 1. Transistorenkennlinien.

III, 1a. Beim Transistor kommt praktisch der gesamte Löcheranteil von j aus dem Emittorstrom, und es interessieren die Kennlinien mit bestimmtem j_p . Diese ergeben sich aus denen mit j_s als Scharparameter, indem man noch die Schar der waagerechten Geraden

$$j_n = \frac{s_p}{s_n} (j_p - j_s)$$

in das $(j_n - U)$ -Diagramm einzeichnet und die Schnittpunkte zusammengehöriger Kurven verbindet (Fig. 6). Die Kennlinien für den Gesamtstrom $j_n + j_p$ ergeben sich daraus einfach durch senkrechtes Verschieben um j_p . Wenn man annimmt, daß der Emittorstrom reiner Löcherstrom ist, der vollständig zum Kollektor gelangt, sind das die fertigen Transistorkennlinien.

Der Löcherstrom erhöht den Elektronenstrom in mit steigender Spannung zunehmendem Maße, jedoch höchstens solange, bis die Zunahme der Stauladung des letzteren die des Löcherstromes wieder kompensiert hat. Dann werden die zu verschiedenen j_p gehörigen Kurven nahezu parallel. Die „innere Stromverstärkung“

$$\alpha_i = \left(\frac{\partial(j_n + j_p)}{\partial j_p} \right)_{U=\text{const}}$$

des Kontakts nimmt also zunächst mit der Spannung zu, bleibt aber schließlich nahezu konstant, und zwar bleibt stets

$$\alpha_i < 1 + \frac{s_p}{s_n}.$$

Die Schnelligkeit des Anstiegs von α_i hängt stark von Φ_0 und T ab, und zwar gilt allgemein: Je höher der Leerstrom des Kontaktes ist, desto rascher steigt α_i an.

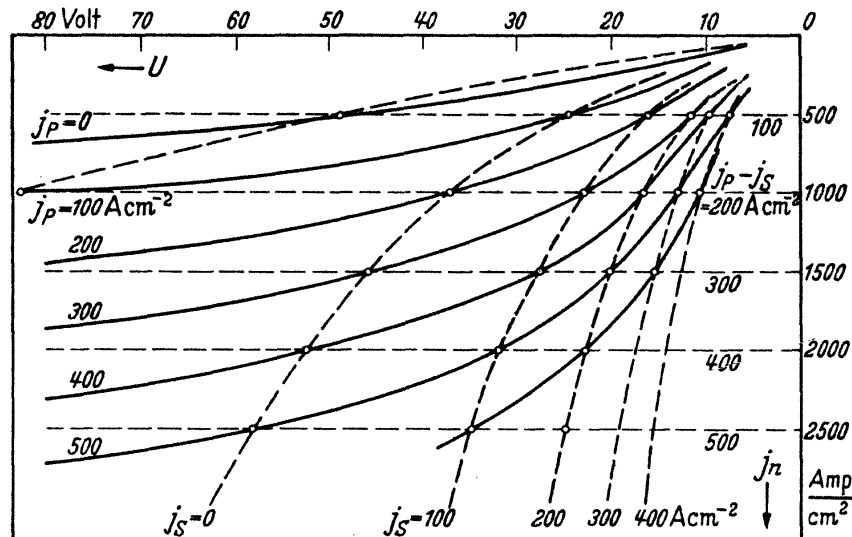


Fig. 6. Konstruktion der Elektronenstromkennlinien für konstanten Löcherstrom. Daten: $r_0 = 10^{-3}$ cm; $e\Phi_0 = 0,38$ eV; $s_n:s_p = 1:5$; $T = 300^\circ$ K. Sonst wie Fig. 5.

Fig. 7a u. b zeigen zusätzlich zu Fig. 6 zwei Transistorkennlinien mit verschiedenem Φ_0 . Die Daten sind: $N_0 = 10^{15}$ cm $^{-3}$, $r_0 = 10^{-3}$ cm, $s_p:s_n = 5$, $T = 300^\circ$ K und $e\Phi_0 = 0,38$ und $0,50$ eV.

III, 1b. Während im Gültigkeitsbereich der Diffusionstheorie nur α_i -Werte mit $\alpha_i < 1 + \frac{b_n}{b_p} \approx 3$ verständlich sind [14], liefert der Stau-effekt die Möglichkeit höherer Werte, maximal $1 + \frac{s_p}{s_n}$, vorausgesetzt, daß $\frac{s_p}{s_n} > \frac{b_n}{b_p}$ ist, was wir als sehr wahrscheinlich ansehen möchten. Auch die Allgemeingestalt der Kennlinien wird durch Fig. 7 befriedigend wiedergegeben¹, jedenfalls solange es sich nicht um zu stark durchfor-mierte Kollektoren handelt.

Beim Formieren ändern sich die Kennlinien sehr stark: Die α -Werte steigen und werden dabei gleichzeitig bis zu weit niedrigeren Spannungen als vorher praktisch spannungsunabhängig, so daß die Kennlinienschar oberhalb größenordnungsmäßig 1 V praktisch parallel wird. Was dabei im einzelnen vorgeht, ist noch nicht genau bekannt. Vermutlich tritt sowohl eine Erhöhung von F_R durch eingeschwemmte Donatorionen als

¹ Vgl. etwa BARDEEN und BRATTAIN [1], Fig. 3, oder SHOCKLEY [15], Fig. 2 und 6.

auch eine Erniedrigung von Φ_0 durch irgendwelche Grenzflächeneffekte auf [2]. Es ist jedoch fraglich, ob man damit bereits auskommt. Uns scheint es sehr wahrscheinlich, daß noch irgendwelche Veränderungen in der Randschicht eintreten, welche die durch die Diodentheorie gegebene Feld-Stromabhängigkeit [Gl. (3) und (4)] abändern¹.

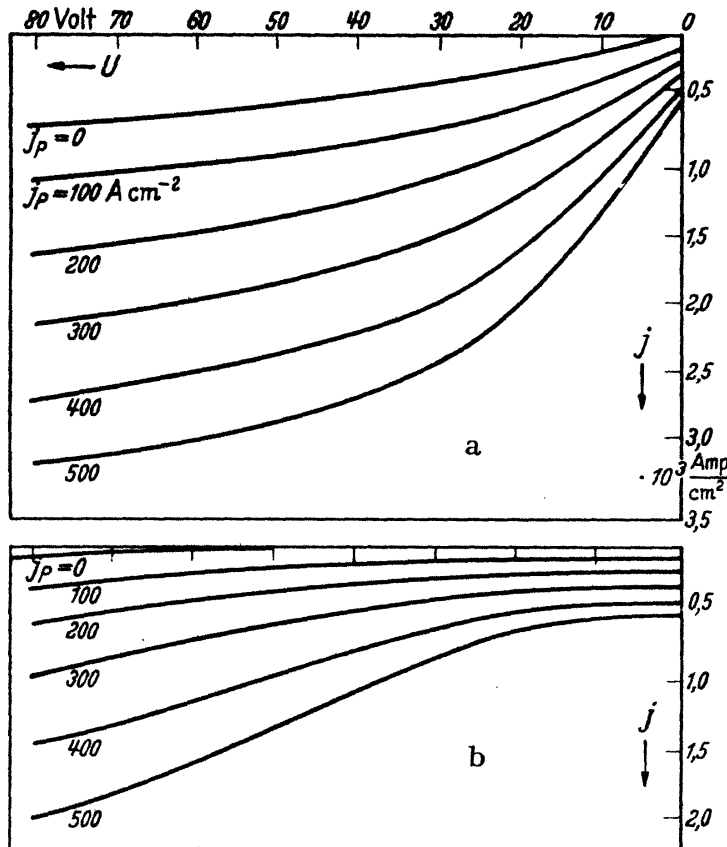


Fig. 7 a u. b. Transistorkennlinien für a) $e\Phi_0 = 0,38$ eV; b) $e\Phi_0 = 0,50$ eV. Sonstige Daten wie Fig. 6.

III, 2. Diodenkennlinien.

III, 2a. Bei Dioden ist keine fremde Löcherquelle vorhanden, und der Löcherstrom besteht aus der geringen Anzahl von Löchern die im Halbleiter bereits vorhanden sind, oder die von der Oberfläche oder von der Basiselektrode her einströmen. Diese Löcher gelangen durch Diffusion und durch das Potentialgefälle, das der vom Kollektor ausgehende Elektronenstrom im HL erzeugt, zum Kollektor. Bei hinreichend hohem Elektronenstrom (worauf wir uns hier be-

schränken) ist die Diffusion vernachlässigbar und der Löcherstrom wird proportional zum Elektronenstrom

$$j_p = G \cdot j_n. \quad (25)$$

Bei homogener Temperaturverteilung und vernachlässigbarer Oberflächenleitung ist

$$G = G_0 = \frac{b_p \cdot p_0}{b_n \cdot n_0} = \frac{b_p}{b_n} \cdot e^{-\frac{E_V - E_L}{kT}}.$$

Im allgemeinen weicht G von diesem Wert aus drei Gründen ab: 1. Die Temperatur ist infolge der Kontakterwärmung inhomogen. 2. Es existiert eine Oberflächenleitfähigkeit. 3. Bei nicht völlig sperrfreier Basiselektrode kann an der Basis eventuell $j_p/j_n > G_0$ sein; dann wirkt die Basis wie ein Emitter.

¹ Vgl. hierzu das auf S. 450 Gesagte. Eine andere Abweichung von der Diodentheorie stellt z. B. die p-n-Hook-Theorie (SHOCKLEY [14], [15], [16]) dar, die uns jedoch nicht frei von Schwierigkeiten scheint.

Wir verzichten auf eine allgemeine Abschätzung von G und wollen die Kennliniengestalt in Abhängigkeit von G untersuchen.

Die Kennlinien ergeben sich aus der Schar der Kennlinien für festes j_s , indem man diese mit der Schar der Geraden

$$j_n = \frac{j_s}{G - \frac{s_n}{s_p}}$$

schneidet und wieder die zusammengehörigen Schnittpunkte verbindet.

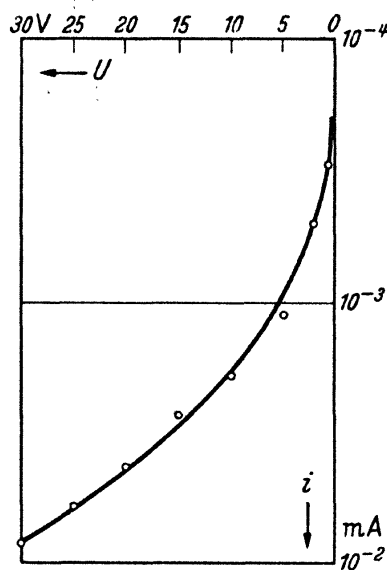


Fig. 8. Vergleich mit einer von BENZER gemessenen Kennlinie.

III, 2b. Bei sehr hochsperrenden Kontakten ist der Strom so niedrig, daß der Staueffekt sich erst

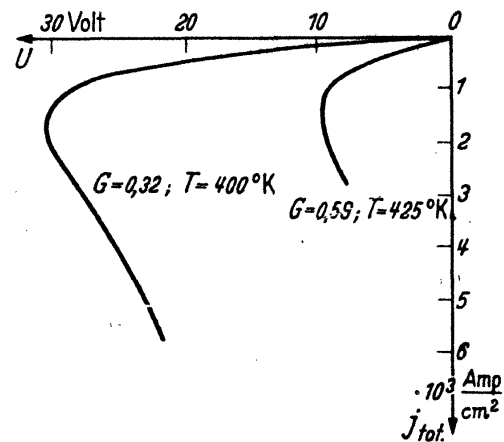


Fig. 9. Umbiegende Kennlinien. Daten außer der Temperatur wie Fig. 7a.

bei hohen Spannungen bemerkbar macht. Man kann dann im Bereich mittlerer Spannungen einfach mit der zu $j_s = 0$ gehörigen Kennlinie rechnen. Das ist in Fig. 8 für $e\Phi_0 = 0,54 \text{ eV}$, $V_0 = 10^{-3} \text{ cm}$ und $T = 298^\circ \text{ K}$ geschehen und mit der zur gleichen Temperatur gehörigen Kurve bei BENZER¹ verglichen. Dabei wurde mit $N_0 = 10^{15} \text{ cm}^{-3}$ und $\eta = 1$ gerechnet. Die Übereinstimmung ist befriedigend. Für ein höheres η wäre sie wesentlich schlechter; ein höheres N_0 , was auch ein höheres η zulassen würde, ist unwahrscheinlich. Sinn dieses Vergleichs ist es, zu zeigen, in welcher Größenordnung Φ_0 bei hochsperrenden Gleichrichtern liegt, und daß man tatsächlich mit $\eta \approx 1$ rechnen muß, um die Experimente erklären zu können.

III, 2c. Für $G < s_n/s_p$ ist der Staueffekt der Löcher schwächer als der der Elektronen. Es kommen nur Schnittpunkte mit negativem j_s in Frage; der Strom ist geringer als für $j_s = 0$. Irgendwelche Besonderheiten treten nicht auf.

Es sei jetzt $G > s_n/s_p$. Aus $j_n = j_n(U, j_s)$ folgt wegen (22)

$$dj_n = \left(\frac{\partial j_n}{\partial U} \right)_{j_s} \cdot dU + \left(\frac{\partial j_n}{\partial j_s} \right)_U \cdot \left(dj_p - \frac{s_n}{s_p} \cdot dj_n \right).$$

¹ Literatur [3], Fig. 10.

Mit $dj = dj_n + dj_p$ und $dj_n = \frac{dj}{1+G}$; $dj_p = \frac{G \cdot dj}{1+G}$ folgt dann

$$\frac{dj}{dU} = \left(\frac{\partial j_n}{\partial U} \right)_{j_s} \cdot \frac{1+G}{1 - \left(G - \frac{s_n}{s_p} \right) \cdot \left(\frac{\partial j_n}{\partial j_s} \right)_U}. \quad (26)$$

Wenn der Nenner verschwindet, erhält die Kennlinie eine senkrechte Tangente und biegt in einen Bereich negativen differentiellen Widerstandes um. Da $\partial j_n / \partial j_s$ mit zunehmendem U und j_s unbegrenzt steigt, ist das für alle Kennlinien mit $G > s_n/s_p$ schließlich der Fall. Fig. 9 zeigt ein Beispiel.

Es erscheint uns nicht ausgeschlossen, daß infolge der Kontakt-erwärmung und einer Oberflächenleitfähigkeit G hinreichend groß wird, und daß dieser Mechanismus in vielen Fällen für das beobachtete Umbiegen der Kennlinien verantwortlich ist.

III, 3. Thermische Effekte.

III, 3a. Die bisher berechneten Kennlinien waren „isotherme“, d.h. zu einer festen Kontakttemperatur gehörige Kennlinien. Tatsächlich erwärmen sich aber die Kontakte ungefähr proportional zur umgesetzten JOULEschen Wärme:

$$T_K - T_0 = \delta \cdot (j \cdot U). \quad (27)$$

Für halbkugelförmige Kontakte in einem nicht zu kleinen HL-Stück gilt

$$\delta = \frac{r_0}{\kappa} \approx 1,7 \cdot 10^{-3} \text{ } ^\circ\text{C} \cdot \text{W}^{-1} \text{ cm}^2$$

bzw.

$$\frac{\delta}{2\pi r_0^2} = \frac{1}{2\pi \kappa r_0} \approx 270^\circ \text{C W}^{-1}.$$

Dabei ist $r_0 = 10^{-3} \text{ cm}$ und die Wärmeleitfähigkeit κ zu $0,14 \text{ cal sec}^{-1} \text{ cm}^{-1} \text{ } ^\circ\text{C}^{-1}$ [6] angenommen worden.

Wenn der Strom reiner Elektronenstrom ist, ist bei seiner Berechnung T_K statt T_0 einzusetzen, da aber T_K seinerseits von j und U abhängt, werden die Gleichstromkennlinien graphisch ermittelt, indem man die Schar der zu verschiedenen T_K gehörigen isothermen Kennlinien mit der der Hyperbeln

$$j = \frac{T_K - T_0}{\delta \cdot U}$$

zum Schnitt bringt und zusammengehörige Schnittpunkte verbindet¹. Hierbei tritt, auch wenn die Isothermen weiterlaufen, meist ein Umbiegen auf (Fig. 10).

¹ HUNTER [8], Die praktische Durchführung geschieht am besten im doppelt logarithmischen Maßstab.

Wenn ein merklicher Löcherstrom auftritt, ist das graphische Verfahren in dieser einfachen Form nicht mehr durchführbar, was wir hier nicht weiter verfolgen wollen.

III, 3b. Die Idee, das beobachtete Umbiegen der Kennlinien in dieser Weise *rein* thermisch zu erklären, findet sich bereits bei HUNTER [8]. Dieser einfachen Deutung steht die Tatsache gegenüber, daß die Hochfrequenzkennlinien selbst bei 30 MHz noch das Umbiegen mit einer Grenzspannung von derselben Größenordnung wie bei Gleichstrom

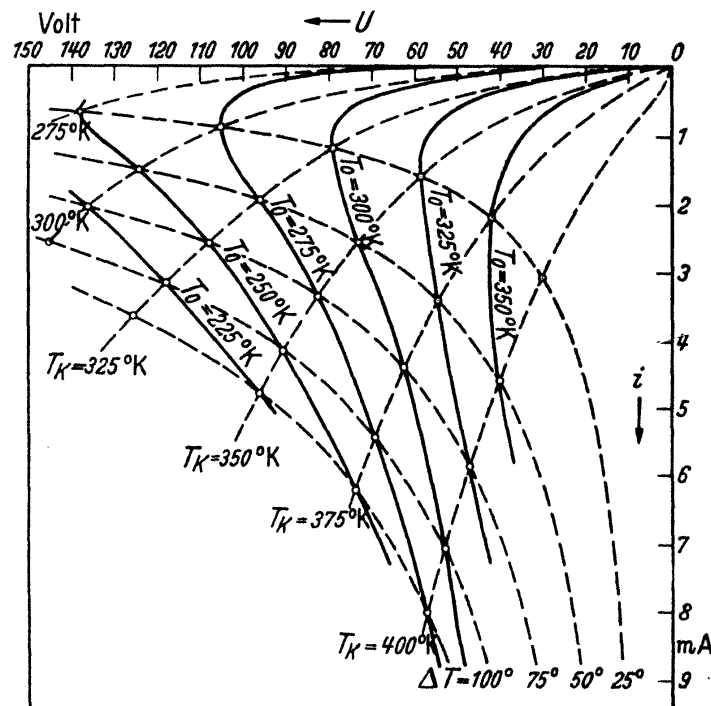


Fig. 10. Konstruktion der Gleichstromkennlinien. Daten: $e \Phi_0 = 0,50 \text{ eV}$; $r_0 = 10^{-3} \text{ cm}$; $j_s = 0$; $\delta = 1,7 \cdot 10^{-3} \text{ } ^\circ\text{C cm}^2 \text{ W}^{-1}$.

zeigen¹, obwohl bei diesen Frequenzen die Temperaturänderung nicht mehr mitkommen kann. Es muß also ein Mechanismus vorhanden sein, durch den bereits die isothermen Kennlinien umbiegen. Einen solchen lieferte oben der Stauereffekt.

III, 3c. Die Temperaturabhängigkeit des Umklappgebietes wird durch die hier vorgelegte, eine Erweiterung der Diodentheorie durch den Stauereffekt darstellende Theorie noch nicht richtig dargestellt. Während (nach BENZER) das Produkt Strom und Spannung am Umbiegepunkt mit steigender Basistemperatur über einen weiten Bereich linear abfällt und bei Extrapolation des linearen Teils etwa beim Einsetzen der Eigenleitung verschwinden würde, liefert die Theorie in der bisherigen Form nur eine sehr schwache T_0 -Abhängigkeit dieses Produktes, es sei denn, man macht die unplausible Annahme, daß G mit steigender Temperatur stark abfällt. Eine nähere Untersuchung zeigt, daß zur Erklärung der

¹ S. TORREY-WHITMER [18], S. 381.

Experimente im Bereich höherer Spannungen eine weit schwächere T -Abhängigkeit der isothermen Kennlinien gefordert werden muß, als die Diodentheorie (und auch die Diffusionstheorie) sie liefert.

Ob man zur Erfüllung dieser Forderung die Annahme $\frac{d\Phi_0}{dT} > 0$ machen darf, müßte erst untersucht werden. Wir möchten eher annehmen, daß diese Diskrepanz neben dem auf S. 446 Gesagten ein zweiter Hinweis darauf ist, daß das Randschichtmodell der Diodentheorie wesentlich verfeinert werden muß.

Für die Anregung zu dieser Arbeit und für zahlreiche Diskussionen habe ich Herrn Prof. SAUTER zu danken.

Literatur.

- [1] BARDEEN, J., and W. H. BRATTAIN: Phys. Rev. **75**, 1208 (1949). — [2] BARDEEN, J., and W. G. PFANN: Phys. Rev. **77**, 401 (1950). — [3] BENZER, S.: J. Appl. Phys. **20**, 804 (1949). — [4] BETHE, H. A.: NDRC-Rep. 14, 23. Nov. 1942. — BETHE, H. A., u. R. G. SACHS: NDRC-Rep. 14, 10. Sept. 1942 u. 15. Juni 1943. — [5] BRIGGS, H. B.: Phys. Rev. **77**, 284 (1950). — [6] GRIECO, A., and H. C. MONTGOMERY: Phys. Rev. **86**, 574 (1952). — [7] HILL, R. W., and D. H. PARKINSON: Phil. Mag. **43**, 309 (1952). — [8] HUNTER, L. P.: Phys. Rev. **81**, 151 (1951). — [9] MÜLLER, H.: Ann. Phys. (6) **9**, 141 (1951). — [10] OLDEKOP, W.: Diss. Göttingen 1952. — [11] PEARSON, G. L., J. R. HAYNES and W. SHOCKLEY: Phys. Rev. **78**, 295 (1950). — [12] SEILER, K.: Z. Naturforsch. **5a**, 393 (1950). — [13] SCHOTTKY, W.: Z. Physik **118**, 539 (1942). — [14] SHOCKLEY, W.: Phys. Rev. **78**, 294 (1950). — [15] SHOCKLEY, W.: Electrons and Holes in Semiconductors. New York 1950. — [16] SHOCKLEY, W., M. SPARKS and G. K. TEAL: Phys. Rev. **83**, 151 (1951). — [17] SOMMERFELD, A., u. H. A. BETHE: Handbuch der Physik, Bd. 24/II, Kap. 3. — [18] TORREY, H. C., and C. A. WHITMER: Crystal Rectifiers. New York 1948.

Göttingen, Institut für Theoretischen Physik¹.

¹ Jetzt am Fernmeldetechnischen Zentralamt in Darmstadt.