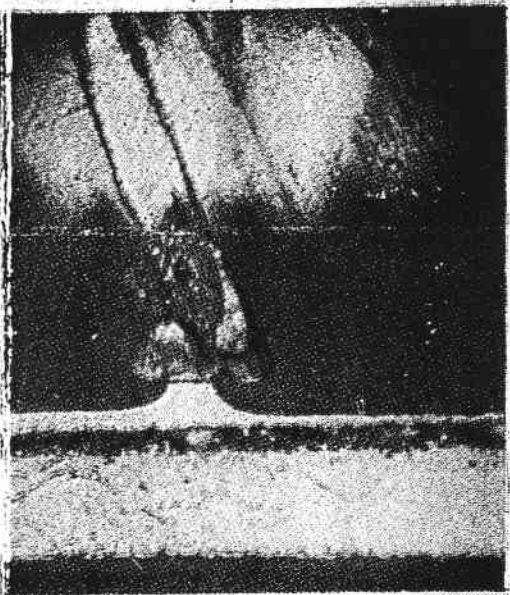
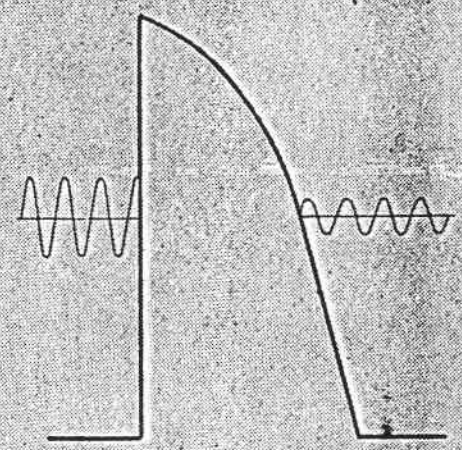
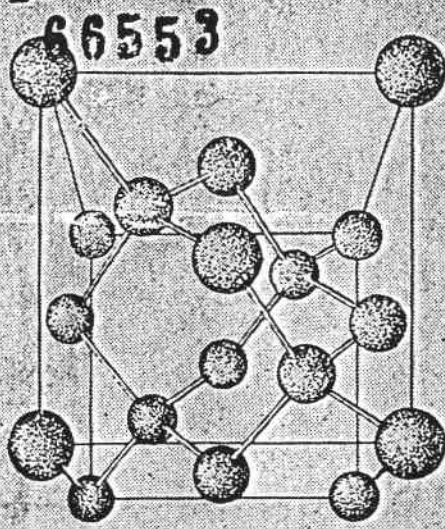


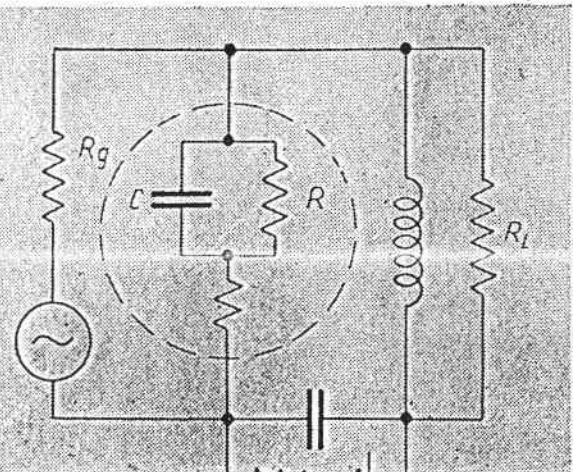
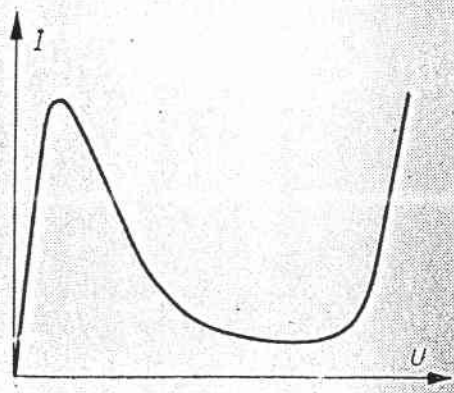
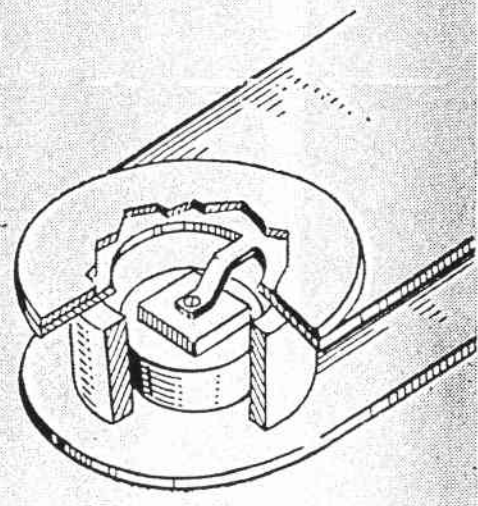
nowa
technika



Miłosz Chmielewski

DIODY TUNELOWE

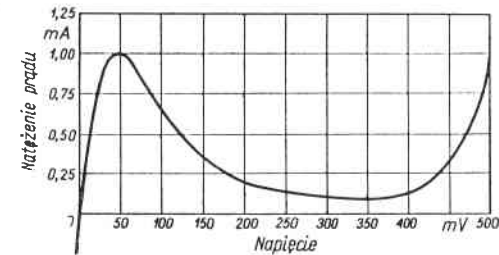
WŁASNOŚCI – TECHNOLOGIA – ZASTOSOWANIE



NT
Zeszyt 39

WSTĘP

W styczniu 1958 r., w znanym czasopiśmie naukowym „Physical Review”, opublikowany został list do redakcji pt. „Nowe zjawisko w germanowych wąskich złączach P-N”. Autor listu, młody japoński fizyk dr Leo Esaki, współpracownik firmy Sony Corporation¹⁾, doniósł o zaobserwowanej anomalii charakterystyki napięciowo-prądowej wykonanego złącza P-N na germanie o bardzo małej oporności właściwej rzędu



Rys. 1. Typowa charakterystyka germanowej diody tunelowej

$10^{-3} \Omega \text{ cm}$, a więc w przybliżeniu o trzy rzędy wielkości mniejszej od stosowanej w produkcji tranzystorów. Charakterystyka statyczna $I-U$ nowej diody znacznie odbiega od kształtu klasycznej charakterystyki zwykłej diody półprzewodnikowej. Najważniejszą jej cechą jest występowanie opadającej części — inaczej mówiąc oporności ujemnej (rys. 1). Zwykle diody odznaczają się dość znacznym napięciem przebicia w kierunku zaporowym (wstecznym). W przypadku diody Esakiego napięcie to jest niezwykle małe, rzędu ułamka wolta.

Olbrzymią zasługą Esakiego było podanie elementarnej i zarazem poprawnej teorii nowej diody, przy czym podstawowym zjawiskiem warunkującym jej działanie okazało się od dawna znane w fizyce kwantowej zjawisko tunelowe. Stąd też nazwa „dioda tunelowa”, może zresztą niezbyt poprawna, bowiem zjawisko tunelowe występuje również w złączach nie wykazujących ujemnej oporności, na przykład w niskonapięciowych diodach Zenera.

¹⁾ Obecnie dr Esaki jest konsultantem firmy International Business Machines Corporation (USA).

Zamieszczony w wyż. wym. czasopiśmie list Esakiego, podobnie jak i referat pod tym samym tytułem, wygłoszony w Brukseli na międzynarodowej konferencji „Fizyka ciała stałego w elektronice i telekomunikacji” (czerwiec 1958), nie zwróciły początkowo uwagi, na jaką bez wątpienia zasługiwało nowe to odkrycie. Kilka jednak firm¹⁾ podjęło doświadczalną produkcję diod tunelowych i wkrótce zainteresowanie nowymi diodami przybrało rozmiary dość dużej sensacji naukowej i technicznej, gdyż okazało się, że diody tunelowe mogą pracować przy bardzo wielkich częstotliwościach, nawet w zakresie fal milimetrowych. Jednocześnie stwierdzono, że diody tunelowe wykazują stosunkowo dużą niewrażliwość na działanie temperatury, światła, promieniowania korpuskularnego, oraz zanieczyszczeń powierzchniowych. Równocześnie uzyskano diody tunelowe wykonane z innych półprzewodników niż german, a mianowicie z krzemu (Si), antymonku indu (InSb), arsenku indu (InAs), fosforu indu (InP), antymonku galu (GaSb), węgla krzemu (SiC). Specyficzne właściwości tych materiałów umożliwiają wytwarzanie diod o nieco odmiennych parametrach, niż w przypadku najczęściej spotykanych diod germanowych.

Zastosowania diod tunelowych mogą być bardzo różnorodne. Przykładowo można tu wymienić:

- wzmacniacze niskosumowe,
- generatory,
- mieszacze i konwertery (przebiegniki) częstotliwości,
- bardzo szybkie przełączniki,
- bistabilne elementy pamięciowe,
- oporniki ujemne dla odtłumiania filtrów,
- detektory małych sygnałów,
- obwody wielofunkcyjne.

Z powyższego zestawienia widać, że diody tunelowe mogą spełniać szereg funkcji zarówno tranzystora, jak i zwykłej diody. Pozornie mogłyby się wydawać, że dioda tunelowa eliminuje tranzystor. W rzeczywistości jednak należy pamiętać, że dioda tunelowa jest elementem o bardzo małej mocy. Jako urządzenie dwójnikowe pozwala znakomicie uprościć szereg układów, lecz trudnym tu problemem jest kaskadowe łączenie diod tunelowych, np. we wzmacniaczach. Równocześnie jednak diody tunelowe mogą pracować przy częstotliwościach wielokrotnie przewyższających maksymalne graniczne częstotliwości tranzystorów.

Odkrycie diody tunelowej, niezależnie od jego doniosłości natury praktycznej, ma również kolosalne znaczenie dla fizyki ciała stałego. Należy tu wymienić przede wszystkim półprzewodniki o bardzo małej oporności właściwej. Są to tzw. półprzewodniki zdegenerowane. Ich własnościami poświęcono dotychczas znikomo mało uwagi, obecnie jednak

¹⁾ General Electric, RCA, Sony, Philco, Hoffmann Electronics, Texas Instruments i inne.

półprzewodniki te znalazły się w centrum zainteresowania elektroniki. Badania mechanizmu tunelowego w złączach P-N w niskich temperaturach doprowadziły do stwierdzenia szeregu interesujących faktów w zakresie wzajemnego oddziaływania elektronów i drgań sieci krystalicznej półprzewodnika oraz stworzyły nowe narzędzie poznawania zjawisk elektronowych w ciałach stałych. Jako przykład można przytoczyć odkrycie zjawiska tunelowego w cienkiej warstwie dielektrycznej pomiędzy nadprzewodzącymi metalami, prowadzące zresztą do powstania nowego elementu, tzw. tunelotronu.

O dużym zainteresowaniu diodami tunelowymi świadczyć może bardzo pokaźna liczba zagranicznych publikacji, poświęconych ich własnościom, a w szczególności praktycznym zastosowaniom¹⁾.

W Polsce pierwsze prace nad diodami tunelowymi zostały zapoczątkowane jesienią 1959 r. w Zakładzie Elektroniki Instytutu Podstawowych Problemów Techniki PAN w Warszawie. Pierwsze serie diod tunelowych wyprodukował Zakład Konstrukcji Prototypów IPPT PAN. Niewątpliwie w niedalekiej przyszłości przemysł nasz podejmie masową ich produkcję, tak że tego rodzaju diody staną się ogólnie dostępne.

1. ZJAWISKO TUNELOWE

Dioda tunelowa stanowi złącze P-N wykonane w kryształach półprzewodnikowych o bardzo małej oporności właściwej. Na granicy obszarów P i N istnieje bariera energii potencjalnej — nawet w stanie równowagi, tzn. bez przyłożenia zewnętrznego napięcia. Szerokość tej bariery jest znikomo mała i dlatego elektrony mogą ją przenikać dzięki zjawisku tunelowemu. Zagadnienie to będzie opisane dokładniej w rozdziale 4. W niniejszym natomiast rozdziale zapoznamy się z istotą zjawiska tunelowego, przy czym szczególnie będzie nas interesowało otrzymanie wzoru na przezroczystość bariery energetycznej o danej wysokości i szerokości.

1.1. Fale de Broglie'a i równanie Schrodingera

Francuski fizyk i laureat nagrody Nobla L. de Broglie wprowadził w 1924 r. hipotezę fal materii, która zakładała, że cząstki materialne (np. elektrony) niezależnie od swych cech korpuskularnych wykazują własności falowe. Długość fali de Broglie'a wyraża wzór:

$$\lambda = \frac{h}{p} = \frac{h}{mv} \quad (1.1)$$

w którym: h — stała Plancka

p — pęd (iloczyn masy m i prędkości v).

¹⁾ Możliwie pełny wykaz publikacji (obejmujący okres do pierwszych miesięcy 1961 r.) podany jest na końcu książki (s. 127).

Wprowadzając pojęcie wektora falowego \vec{k} o wartości bezwzględnej $|\vec{k}| = k$:

$$k = \frac{2\pi}{\lambda} \quad (1.2)$$

mamy

$$\vec{p} = \frac{\hbar \vec{k}}{2\pi} \quad (1.3)$$

Równanie fali płaskiej opisującej ruch swobodnych cząstek jest następujące:

$$\psi = A e^{j(kr - \omega t)} \quad (1.4)$$

przy czym

$$\omega = 2\pi f$$

Kwadrat modułu $|\psi|^2 = A^2$ określa natężenie fali. Prędkość grupowa fali zdefiniowana jako:

$$v_g = \frac{d\omega}{dk} \quad (1.5)$$

jest równa prędkości cząstki. Korzystając ze znanego wzoru na energię:

$$W = hf = \frac{\hbar \omega}{2\pi} \quad (1.6)$$

mamy

$$v_g = \frac{2\pi}{\hbar} \frac{dW}{dk} \quad (1.7)$$

Hipoteza fal materii została potwierdzona przez Davissona i Germera w słynnym doświadczeniu z dyfrakcją wiązki elektronowej na kryształach niklu, spełniającym rolę siatki dyfrakcyjnej, a także w szeregu innych doświadczeniach.

Idea de Broglie'a zapoczątkowała powstanie mechaniki kwantowej, w ramach której staje się możliwe wyjaśnienie interesującego nas zjawiska tunelowego, niezrozumiałego na gruncie mechaniki klasycznej. W dalszym ciągu skorzystamy z najbardziej chyba zasadniczego równania mechaniki kwantowej, mianowicie równania Schroedingera:

$$\nabla^2 \psi + \frac{8\pi^2 m}{\hbar^2} (W - W_p) \psi = 0 \quad (1.8)$$

w którym:

$$\nabla^2 \psi \equiv \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2}$$

przy czym W — energia całkowita,
 W_p — energia potencjalna.

Fizyka teoretyczna równania tego nie wyprowadza, ale słuszność jego potwierdza zgodność wyników otrzymanych z równania Schroedingera z doświadczeniem.

Szczególnie interesujący jest sens fizyczny kwadratu amplitudy modułu funkcji ψ . Kwadrat ten daje gęstość prawdopodobieństwa znalezienia się cząstki w objętości $dx dy dz$, a więc w obszarze zamkniętym między $x + dx$, $y + dy$, $z + dz$. Samo prawdopodobieństwo jest proporcjonalne do iloczynu:

$$|\psi|^2 dx dy dz \quad (1.9)$$

Prawdopodobieństwo znalezienia się cząstki gdziekolwiek w przestrzeni jest oczywiście równe jedności. Tak więc:

$$\int_{-\infty}^{+\infty} |\psi|^2 dx dy dz = 1 \quad (1.10)$$

Równanie Schroedingera można przepisać w nieco innej postaci. Zauważmy, że $W_k = W - W_p$ jest energią kinetyczną.

Ponieważ

$$W_k = \frac{mv^2}{2} = \frac{m^2 v^2}{2m} = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{8\pi^2 m} \quad (1.11)$$

więc

$$\frac{8\pi^2 m}{\hbar^2} (W - W_p) = k^2 \quad (1.12)$$

czyli

$$k = \frac{2\pi}{\hbar} \sqrt{2m(W - W_p)} \quad (1.13)$$

Wzór (1.8) przyjmie postać

$$\nabla^2 \psi + k^2 \psi = 0 \quad (1.14)$$

1.2. Bariera progowa. Współczynniki przeźroczystości i odbicia

Rozpatrzmy barierę energii potencjalnej o wysokości W_p , na którą pada elektron o energii W , mniejszej od W_p (rys. 1-1). Według mechaniki klasycznej elektron dozna odbicia od bariery. Mechanika kwantowa przyjmuje możliwość przejścia „tunelowego” do obszaru większej energii potencjalnej.

W obszarze $x \leq 0$ równanie Schroedingera będzie miało postać:

$$\frac{d^2 \psi_1}{dx^2} + k_1^2 \psi_1 = 0 \quad (1.15)$$

przy czym

$$k_1 = \frac{2\pi}{\hbar} \sqrt{2mW_k}$$

W obszarze $x \geq 0$:

$$\frac{d^2 \psi_2}{dx^2} + k_2^2 \psi_2 = 0 \quad (1.16)$$

przy czym

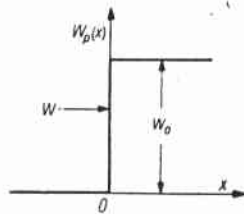
$$k_2 = \frac{2\pi}{h} \sqrt{2m(W_0 - W)}$$

Rozwiązaniem równań (1.15) i (1.16) będzie:

$$\psi_1 = a_1 e^{ik_1 x} + b_1 e^{-ik_1 x} \quad (1.17)$$

$$\psi_2 = a_2 e^{ik_2 x} + b_2 e^{-ik_2 x} \quad (1.18)$$

Rozwiązania te przedstawiają fale płaskie de Broglie'a, przy czym znak dodatni odnosi się do fali poruszającej się w kierunku dodatnim osi x , a znak ujemny — do fali w kierunku przeciwnym.



Rys. 1-1. Bariera progowa

Kwadraty współczynników a_1 , a_2 , b_1 oznaczają odpowiednio natężenia fali padającej na barierę, fali przepuszczonej w obszar bariery oraz fali odbitej. Współczynnik b_2 jest oczywiście równy zero, gdyż nie można oczekiwać fali odbitej w drugim obszarze ($x > 0$). W dalszym ciągu będziemy zakładać dla wygody, że natężenie fali padającej jest równe jedności, tzn. że:

$$a_1^2 = 1 \quad a_1 = 1$$

Analogicznie jak w optyce, można wprowadzić pojęcie współczynnika odbicia R , zdefiniowanego jako:

$$R = \frac{\text{natężenie fali odbitej od bariery}}{\text{natężenie fali padającej na barierę}} = \frac{b_1^2}{a_1^2} = b_1^2 \quad (1.19)$$

Oznaczmy symbolem ρ gęstość przestrzenną elektronów o prędkości v . Iloczyn ρv jest strumieniem elektronów przechodzących w ciągu 1 s przez powierzchnię 1 cm^2 .

Przez współczynnik przeźroczystości P bariery rozumiemy stosunek:

$$P = \frac{\text{strumień cząstek przechodzących przez barierę}}{\text{strumień cząstek padających na barierę}} = \frac{\rho_2 v_2}{\rho_1 v_1} \quad (1.20)$$

Gęstość przestrzenna ρ jest proporcjonalna do kwadratu amplitudy fali de Broglie'a. Prędkość v jest proporcjonalna do pędu:

$$v \sim p$$

Z uwagi na wzór (1.3) $v \sim k$, tak że:

$$\frac{\rho_2}{\rho_1} = \frac{a_2^2}{a_1^2} \quad \frac{v_2}{v_1} = \frac{k_2}{k_1}$$

i tym samym

$$P = a_2^2 \frac{k_2}{k_1} \quad (1.21)$$

Skorzystajmy teraz z warunków ciągłości w punkcie $x = 0$ funkcji ψ_1 i ψ_2 , jak i ich pochodnych:

$$\begin{aligned} \psi_1 &= \psi_2 & 1 + b_1 &= a_2 \\ \frac{d\psi_1}{dx} &= \frac{d\psi_2}{dx} & 1 - b_1 &= a_2 \frac{k_2}{k_1} \end{aligned}$$

Rozwiązanie powyższych równań daje:

$$b_1 = \frac{k_1 - k_2}{k_1 + k_2} \quad a_2 = \frac{2k_1}{k_1 + k_2} \quad (1.22)$$

Podstawiając wzór (1.22) do wzoru (1.21) oraz (1.19), otrzymujemy:

$$R = \left(\frac{k_1 - k_2}{k_1 + k_2} \right)^2 \quad (1.23)$$

$$P = \frac{4k_1 k_2}{(k_1 + k_2)^2} \quad (1.24)$$

Łatwo można się przekonać, że suma obu współczynników jest równa jedności:

$$R + P = 1 \quad (1.25)$$

1.3. Bariera prostokątna

Założmy, że mamy barierę prostokątną o wysokości W_0 i szerokości d (rys. 1-2), dzielącą jednowymiarową przestrzeń na trzy obszary:

$$\begin{aligned} x &\leq 0 \\ 0 &\leq x \leq d \\ x &\geq d \end{aligned}$$

Równania Schroedingera dla poszczególnych obszarów przyjmą postać:

$$\frac{d^2 \psi_1}{dx^2} + k_1^2 \psi = 0 \quad \text{gdzie } k_1 = \frac{2\pi}{h} \sqrt{2mW} \quad (1.26)$$

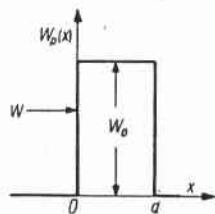
$$\frac{d^2 \psi_2}{dx^2} + k_2^2 \psi = 0 \quad \text{gdzie } k_2 = \frac{2\pi}{h} \sqrt{2m(W - W_0)} \quad (1.27)$$

$$\frac{d^2 \psi_3}{dx^2} + k_3^2 \psi = 0 \quad \text{gdzie } k_3 = \frac{2\pi}{h} \sqrt{2mW} \quad (1.28)$$

W wyniku rozwiązania mamy:

$$\begin{aligned} \psi_1 &= a_1 e^{ik_1 x} + b_1 e^{-ik_1 x} \\ \psi_2 &= a_2 e^{ik_2 x} + b_2 e^{-ik_2 x} \\ \psi_3 &= a_3 e^{ik_3 x} + b_3 e^{-ik_3 x} \end{aligned} \quad (1.29)$$

Podobnie jak poprzednio, przyjmijmy $a_1 = 1$. Jest oczywiste, że współczynnik $b_3 = 0$, gdyż nie można oczekiwać fali odbitej w trzecim obszarze.



Rys. 1-2. Bariera prostokątna

Napiszmy analogicznie jak poprzednio warunki ciągłości funkcji falowych oraz ich pochodnych w punktach $x = 0$ oraz $x = d$.

$$\left. \begin{aligned} \psi_1 = \psi_2 \left\{ \begin{array}{l} x=0 \\ x=d \end{array} \right. \left. \begin{array}{l} 1 + b_1 = a_2 + b_2 \\ a_2 e^{ik_2 d} + b_2 e^{-ik_2 d} = a_3 e^{ik_3 d} \end{array} \right\} \\ \frac{d\psi_1}{dx} = \frac{d\psi_2}{dx} \left\{ \begin{array}{l} x=0 \\ x=d \end{array} \right. \left. \begin{array}{l} k_1 - k_1 b_1 = k_2 a_2 - k_2 b_2 \\ a_2 e^{ik_2 d} - b_2 e^{-ik_2 d} = a_3 \frac{k_1}{k_2} e^{ik_3 d} \end{array} \right\} \end{aligned} \quad (1.30)$$

Współczynnikiem przeźroczystości będzie analogicznie jak we wzorze (1.21):

$$P = \frac{|a_3|^2}{|a_1|^2} \cdot \frac{k_3}{k_1} = |a_3|^2 \quad (1.31)$$

Z równań (1.30) znajdujemy:

$$a_3 = \frac{4k_1 k_2 e^{ik_1 d}}{(k_1 + k_2)^2 e^{-ik_2 d} - (k_1 - k_2)^2 e^{ik_2 d}} \quad (1.32)$$

Ze wzoru $k_2 = \frac{2\pi}{h} \sqrt{2m(W - W_0)}$ wynika, że k_2 jest liczbą urojoną. Przez podstawienie $k_2 = jk$ otrzymujemy:

$$k = \frac{2\pi}{h} \sqrt{2m(W_0 - W)} \quad (1.33)$$

Wzór (1.32) przepisujemy w postaci:

$$a_3 = \frac{+j4k_1 k e^{-jk_1 d}}{(k_1 - jk)^2 e^{kd} - (k_1 + jk)^2 e^{-kd}}$$

Wartość sprzężoną a_3^* otrzymujemy zmieniając znaki przy j :

$$a_3^* = \frac{-j4k_1 k e^{-jk_1 d}}{(k_1 - jk)^2 e^{kd} - (k_1 + jk)^2 e^{-kd}}$$

Ponieważ $|a_3|^2 = a_3 a_3^*$, to

$$P = \frac{4k_1^2 k}{(k_1^2 - k^2) \sin h^2 kd + 4k_1^2 k^2 \cosh^2 kd}$$

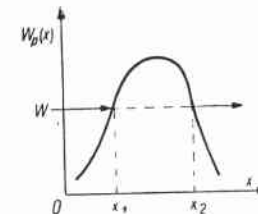
W przybliżeniu $\sin h^2 kd \approx \frac{e^{2kd}}{2}$ oraz $k_1 \approx k$, tak że

$$P \approx \frac{4}{\left(\frac{k_1}{k} + \frac{k}{k_1}\right)^2 e^{2kd} + 4} \approx \frac{1}{e^{2kd} + 1}$$

Pomijając w mianowniku 1 w porównaniu z wyrazem e^{2kd} , otrzymujemy przybliżony wzór:

$$P \approx e^{-2kd} = e^{-\frac{4\pi}{h} \sqrt{2m(W_0 - W)} d} \quad (1.34)$$

Przenikalność bariery wzrasta zatem bardzo szybko przy zmniejszaniu jej szerokości. Wzór (1.34) jest słuszny dla bariery prostokątnej. Dla



Rys. 1-3. Bariera o dowolnym kształcie

bariery o szerokości $d = x_2 - x_1$, lecz o kształcie uwarunkowanym przebiegiem energii potencjalnej $W_p(x)$ — (rys. 1-3) mamy:

$$P \approx e^{-\frac{4\pi\sqrt{2m}}{h} \int_{x_1}^{x_2} \sqrt{W_p(x) - W} dx} \quad (1.35)$$



Przenikanie elektronu przez barierę w zjawisku tunelowym ¹⁾ nie jest związane ze stratą lub zyskiem energii. Elektron opuszczający barierę ma tę samą energię, z jaką uderzył w barierę. Ze wzoru (1.35) skorzystamy w rozdziale 2 przy omawianiu zjawiska emisji w silnym polu elektrycznym. Wielkość P traktujemy jako prawdopodobieństwo.

2. PODSTAWOWE WŁASNOŚCI PÓLPRZEWODNIKÓW

2.1. Ruch elektronu w polu periodycznym kryształu

W rozdziale 1 podane zostały wzory na energię elektronu swobodnego

$$W = \frac{h^2}{8\pi^2 m} k^2 \quad (2.1)$$

prędkość

$$v = \frac{2\pi}{h} \cdot \frac{dW}{dk} \quad (2.2)$$

funkcję falową

$$\psi = Ae^{ikx} \quad (2.3)$$

W kryształach elektron porusza się w polu periodycznym o potencjale:

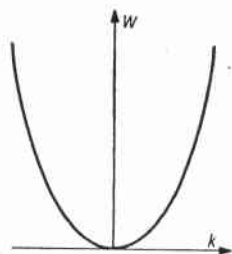
$$V(x) = V(x+a) \quad (2.4)$$

gdzie a — stała sieci krystalicznej.

Rozwiązując równanie Schrodingera zawierające potencjał periodyczny, otrzymuje się tzw. funkcję falową Blocha:

$$\psi = b(k; x) e^{ikx} \quad (2.5)$$

gdzie $b(k; x)$ jest funkcją periodyczną o okresie stałej sieci. Wykres zależności energii W od wektora falowego k stanowi parabolę (rys. 2-1).



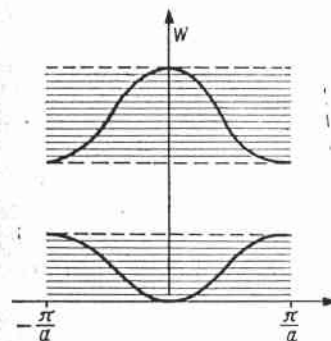
Rys. 2-1. Zależność energii od wektora falowego dla elektronu swobodnego

Podobny wykres w przypadku ruchu elektronu w kryształach daje nieciągłości wykresu $W(k)$ — rys. 2-2, prowadzące do powstania pasm energii

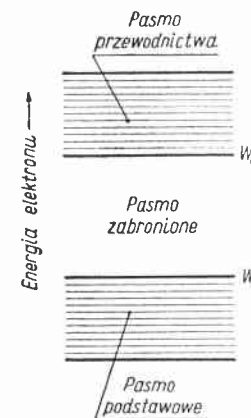
¹⁾ Bardziej szczegółowe omówienie podstaw fizycznych zjawiska tunelowego znajdzie Czytelnik w książkach: Szpolski — „Fizyka atomowa”, D. Błochincew — „Podstawy mechaniki kwantowej”.

w kryształach. Pasma dozwolonej wartości energii przedzielone są pasmami zabronionymi (rys. 2-2). Nieciągłości występują dla wartości wektora falowego:

$$k = \pm n \frac{\pi}{a} \quad (n = 1, 2, \dots) \quad (2.6)$$



Rys. 2-2. Zależność energii od wektora falowego dla elektronu w kryształach



Rys. 2-3. Model pasmowy półprzewodnika

Obszar w przestrzeni k zawierający się między $-\frac{\pi}{a}$ oraz $+\frac{\pi}{a}$ nazywa się strefą Brillouina.

W badaniach elektrycznych własności półprzewodników interesują nas tylko trzy pasma: podstawowe (walencyjne), zabronione i przewodnictwa.

Układ tych trzech pasm przedstawiony jest schematycznie na rys. 2-3.

2.2. Masa efektywna elektronu

W mechanice klasycznej istnieje następująca zależność pomiędzy siłą F , masą m i przyspieszeniem a :

$$F = ma \quad (2.7)$$

którą można napisać również w postaci:

$$a = \frac{dv}{dt} = \frac{F}{m} \quad (2.8)$$

Korzystając ze wzoru (2.2) otrzymujemy:

$$a = \frac{2\pi}{h} \frac{d}{dt} \left(\frac{dW}{dk} \right) = \frac{2\pi}{h} \cdot \frac{d}{dk} \left(\frac{dW}{dt} \right) \quad (2.9)$$

Praca wykonana przez siłę F w ciągu czasu dt równa jest zmianie energii dW . Siła F działająca na drodze dx wykonuje pracę:

$$Fdx = Fvdt$$

równą zmianie energii dW w tym czasie:

$$Fvdt = dW \quad (2.10)$$

Podstawiając wzór (2.10) do wzoru (2.9) otrzymujemy wyrażenie:

$$a = \frac{dv}{dt} = F \frac{2\pi}{h} \cdot \frac{d^2W}{dk^2} \quad (2.11)$$

tak że rolę masy przejmuje wielkość:

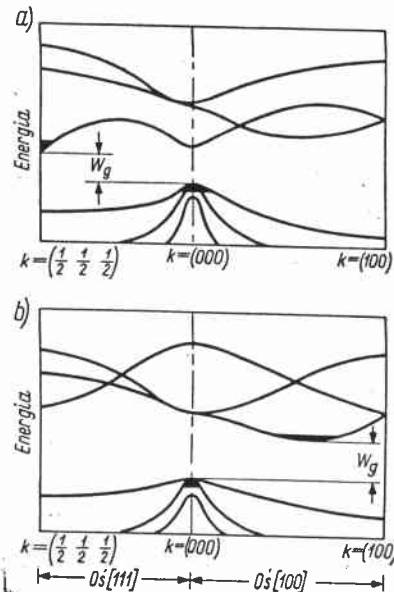
$$m^* = \frac{1}{\frac{4\pi^2}{h^2} \cdot \frac{d^2W}{dk^2}} \quad (2.12)$$

nazwana masą efektywną. Masa efektywna m^* jest wielkością anizotropową, zależną od kierunku w kryształach. Pod wpływem zewnętrznego pola elektrycznego K elektron w kryształach nabiera przyspieszenia:

$$a = \frac{qK}{m^*} \quad (2.13)$$

2.3. Struktura pasmowa germanu i krzemu

Nowsze badania, zarówno teoretyczne jak i doświadczalne wskazują na złożoność struktury pasmowej półprzewodników. Na rys. 2-4 uwi-



Rys. 2-4. Struktura pasmowa: a) germanu, b) krzemu

doczniono zależność $W(k)$ w przypadku germanu i krzemu. Lewa strona rysunku odnosi się do kierunku $[111]$, prawa zaś do kierunku $[100]$. Wiadać, że w pasmie przewodnictwa w germanie istnieją cztery minima energii. Badania wykazały, że w pobliżu każdego z tych minimum powierzchnia stałej energii ma kształt elipsoidy obrotowej; w tym przypadku masę efektywną cechują dwie składowe: podłużna m_l^* i poprzeczna m_t^* . Stosunek tych składowych nazywany jest współczynnikiem anizotropii. Wartości mas m_l^* i m_t^* dla germanu są następujące:

$$\frac{m_l}{m_0} = 1,64 \pm 0,003 \quad \frac{m_t}{m_0} = 0,0819 \pm 0,003$$

gdzie m_0 — masa elektronu swobodnego.

2.4. Gęstość stanów w pasmie

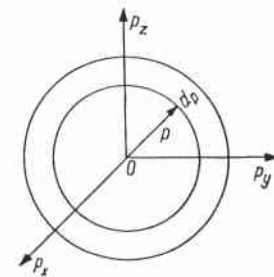
Stan elektronu w kryształach określa trzy współrzędne x, y, z oraz trzy składowe pędu p_x, p_y, p_z . Zależność:

$$W = \frac{h^2 k^2}{8\pi^2 m^*} = \frac{p^2}{2m^*}$$

pozwala stwierdzić, że powierzchnią stałej energii jest kula o promieniu:

$$p = \sqrt{p_x^2 + p_y^2 + p_z^2} = \sqrt{2m^*W}$$

Całą przestrzeń pędów i współrzędnych można rozłożyć na komórki o objętości $dp_x dp_y dp_z dx dy dz = h^3$. Liczba tych komórek określa liczbę



Rys. 2-5. Warstwa kulista w przestrzeni pędów

bę możliwych stanów. W objętości jednostkowej kryształu można wydzielić wszystkie stany o energii w przedziale $W, W + dW$. W przestrzeni pędów p_x, p_y, p_z (rys. 2-5) będzie to warstwa kulista o promieniu p i grubości dp . Objętość tej warstwy jest równa $4\pi p^2 dp$.

Gęstością stanów $D(W)$ jest

$$D(W) = \frac{4\pi p^2 dp}{h^3} \quad (2.14)$$

Ponieważ $p^2 = 2 m^* W$ oraz $dp = \frac{1}{2} \sqrt{2m^* W}^{-\frac{1}{2}} dW$

to

$$D(W) dW = \frac{2\pi(2m^*)^{\frac{3}{2}} W^{\frac{1}{2}} dW}{h^3} \quad (2.15)$$

lub

$$D(W) dW = 4\pi \left(\frac{2m^*}{h^2}\right)^{\frac{3}{2}} (W - W_c)^{\frac{1}{2}} dW \quad (2.16)$$

gdzie W_c jest energią dna pasma przewodnictwa.

Podobnie dla pasma podstawowego:

$$D(W) dW = 4\pi \left(\frac{2m^*}{h^2}\right)^{\frac{3}{2}} (W_v - W)^{\frac{1}{2}} dW \quad (2.17)$$

gdzie W_v jest energią wierzchołka pasma podstawowego.

2.5. Półprzewodnik samoistny

Najmniejsza odległość pomiędzy wierzchołkiem pasma podstawowego W_v i dnem pasma przewodnictwa W_c (rys. 2-4) nazywa się przerwą energetyczną W_g . Wartości przerwy energetycznej dla kilku przykładowo podanych tu półprzewodników są następujące: Ge - 0,665¹⁾, Si - 1,1, GaAs - 1,35, GaSb - 0,8, InSb - 0,17, SiC - 3,5, InAc - 0,47.

Wskutek pobudzenia cieplnego elektrony z pasma podstawowego przenoszą się do pasma przewodnictwa. Brak elektronu w pasmie podstawowym nazywa się dziurą. Jest oczywiste, że koncentracja dziur p w pasmie podstawowym będzie równa koncentracji elektronów n w pasmie przewodnictwa:

$$n = p = n_i$$

Półprzewodnik zawdzięczający swe przewodnictwo tego rodzaju mechanizmowi nazywa się półprzewodnikiem samoistnym i oznaczany jest symbolem I²⁾.

Przyłożenie pola elektrycznego do kryształu powoduje ruch elektronów w pasmie przewodnictwa i dziur w pasmie podstawowym. Przewodnictwo w dużym stopniu zależy od temperatury, co powoduje, że współczynnik temperaturowy oporności właściwej jest ujemny. Jest to charakterystyczną cechą półprzewodników.

2.6. Półprzewodniki domieszkowane

W praktyce mamy do czynienia z półprzewodnikami niemal wyłącznie zawierającymi wprowadzone w sposób kontrolowany domieszki obcych atomów. Domieszki mogą mieć charakter donorowy lub akceptorowy.

¹⁾ W temperaturze 300 °K.

²⁾ Ang. intrinsic (samoistny).

Typowymi domieszkami donorowymi w germanie i krzemie są pierwiastki z piątej kolumny układu periodycznego: fosfor (P), antymon (Sb), arsen (As). Typowe domieszki akceptorowe stanowią pierwiastki z trzeciej kolumny: gal (Ga), ind (In), bor (B), aluminium (Al).

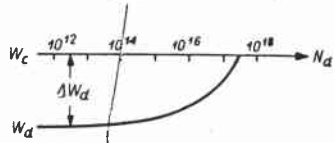
Donory wprowadzają lokalne poziomy energetyczne leżące w pobliżu i poniżej pasma przewodnictwa, natomiast akceptory wprowadzają poziomy w pobliżu i powyżej pasma podstawowego. Niezależnie od tego rodzaju domieszek płytkich istnieją domieszki wprowadzające poziomy głębiej usytuowane w pasmie zabronionym. Różnicę $W_c - W_d = \Delta W_d$ nazywamy energią jonizacji donorów, natomiast różnicę $W_a - W_v = \Delta W_a$ - energią jonizacji akceptorów. Wartości ΔW_a oraz ΔW_d dla różnych domieszek w germanie i krzemie ujęte są w tablicy 1.

Tablica 1

Energie jonizacji donorów i akceptorów w germanie i krzemie

Kolumna układu okresowego pierwiastków	Domieszka w germanie lub krzemie	Donor (D) Akceptor (A)		Energia jonizacji [eV] mierzona od pasma przewodnictwa do poziomu domieszek		Energia jonizacji [eV] mierzona od pasma podstawowego do poziomu domieszki	
		w germanie	w krzemie	w germanie	w krzemie	w germanie	w krzemie
I	Li	D	D	0,01	0,033	—	0,33
	Cu	—	—	—	—	0,040; 0,31	—
	Au	A; (D)	D	0,2	0,30	0,15; 0,53	0,39
II	Zn	A	A	—	—	0,029	0,092 0,30
III	B	A	A	—	—	0,0104	0,045
	Al	A	A	—	—	0,0102	0,057
	Ga	A	A	—	—	0,0108	0,065
	In	A	A	—	—	0,0112	0,16
	Tl	A	—	—	—	0,014	—
V	P	D	D	0,0120	0,039	—	—
	As	D	D	0,0127	0,049	—	—
	Sb	D	D	0,0097	0,039	—	—
	Bi	D	—	0,012	—	—	—
VII	Mn	A	—	—	0,35	0,16	
VIII	Fe	A	—	—	—	0,27	0,34
	Co	A	—	—	—	0,31	0,25
	Ni	A	—	—	—	0,30	0,22
	Pt	A	—	—	—	0,20	0,04

Płytkie donory oddają elektrony do pasma przewodnictwa, natomiast płytkie akceptory odbierają elektrony z pasma podstawowego. Ze względu na małe wartości energii jonizacji ΔW_d oraz ΔW_a w porównaniu z energią kT (0,025 eV w temperaturze $T = 300$ °K), płytkie donory i akceptory są całkowicie zjonizowane i tym samym stanowią odpowiednio dodatnio i ujemnie naładowane jony w sieci krystalicznej, których ładunek jest skompensowany przez elektrony i dziury, tak że kryształ półprzewodnikowy jako całość jest elektrycznie obojętny.



Rys. 2-6. Zależność energii jonizacji donorów od koncentracji

Przy dużych koncentracjach donorów $N_d > 10^{17}$ cm⁻³ oraz akceptorów N_a energie jonizacji maleją do zera (rys. 2-6).

Półprzewodnik, którego przewodnictwo elektryczne określają domieszki donorowe, nazywa się półprzewodnikiem typu N. W przypadku domieszek akceptorowych mówimy o półprzewodniku typu P.

2.7. Położenie poziomu Fermiego w półprzewodniku

Prawdopodobieństwo obsadzenia stanu o energii W przez elektron jest zgodnie ze statystyką Fermiego-Diraca równe:

$$f(W) = \frac{1}{e^{\frac{W-F}{kT}} + 1} \quad (2.18)$$

Parametr F nazywa się poziomem Fermiego. Dla $W = F$, $f(W) = \frac{1}{2}$. Wzór (2.18) dla $W - F \gg kT$ przyjmuje postać:

$$f(W) = e^{-\frac{W-F}{kT}} \quad (2.19)$$

odpowiadającą rozkładowi statystyki Boltzmanna.

Jest rzeczą istotną, że dla kilku stykających się półprzewodników poziom Fermiego w stanie równowagi zajmuje jednakowe położenie.

Wyznamy teraz położenie poziomu Fermiego w półprzewodniku samoistnym. Ze wzoru (2.16) i (2.19) można wyprowadzić następujące wyrażenie na koncentrację elektronów w pasmie przewodnictwa:

$$n = \int_{W_c=0}^{\infty} D(W) f(W) dW = 4\pi \left(\frac{2m_n^*}{h^2} \right)^{\frac{3}{2}} \int_0^{\infty} W^{\frac{1}{2}} e^{-\frac{W-F}{kT}} dW$$

Podstawiając $\eta = \frac{W}{kT}$ otrzymujemy:

$$n = 4\pi \left(\frac{2m_n^* kT}{h^2} \right)^{\frac{3}{2}} e^{\frac{F}{kT}} \int_0^{\infty} \eta^{\frac{1}{2}} e^{-\eta} d\eta$$

a stąd

$$n = 2 \left(\frac{2\pi m_n^* kT}{h^2} \right)^{\frac{3}{2}} e^{\frac{F}{kT}} \quad (2.20)$$

Podobnie ze wzoru (2.17) oraz (2.19) otrzymuje się:

$$p = \int_{-\infty}^{W_v=W_g} D(W) f(W) dW = 2 \left(\frac{2\pi m_p^* kT}{h^2} \right)^{\frac{3}{2}} e^{\frac{F}{kT}} \quad (2.21)$$

W półprzewodniku samoistnym $n = p$, tak że:

$$\left(m_n^* \right)^{\frac{3}{2}} e^{\frac{F}{kT}} = \left(m_p^* \right)^{\frac{3}{2}} e^{-\frac{F+W_g}{kT}}$$

stąd

$$F = -\frac{W_g}{2} + \frac{3kT}{4} \ln \left(\frac{m_p^*}{m_n^*} \right) \quad (2.22)$$

dla $m_n^* = m_p^*$

$$F = -\frac{W_g}{2} \quad (2.23)$$

co oznacza, że w półprzewodniku samoistnym poziom Fermiego leży w przybliżeniu w środku przerwy energetycznej (pasma zabronionego).

W półprzewodniku domieszkowym typu N, przy założeniu całkowitej jonizacji donorów, $n = N_d$.

$$\text{Stąd } n = N_d = \frac{2}{\sqrt{\pi}} N_c e^{\varphi N} \int_0^{\infty} \eta^{\frac{1}{2}} e^{-\eta} d\eta = \frac{2}{\sqrt{\pi}} N_c e^{\varphi N} \cdot \Gamma\left(\frac{3}{2}\right) \quad (2.24)$$

przy czym

$$\eta = \frac{W}{kT} \quad \varphi = \frac{F}{kT}$$

Wiadomo, że

$$\Gamma\left(\frac{3}{2}\right) = \int_0^{\infty} \eta^{\frac{3}{2}-1} e^{-\eta} d\eta = \frac{\sqrt{\pi}}{2}$$

przy czym symbol Γ oznacza funkcję gamma.

Wzór (2.24) sprowadza się więc do postaci:

$$N_d = N_c e^{\varphi N}$$

albo inaczej

$$\varphi_N = \ln \left(\frac{N_d}{N_c} \right) \quad (2.25)$$

$$F_N = \varphi_N kT = kT \ln \left(\frac{N_d}{N_c} \right)$$

Dla $N_d \rightarrow N_c$, $\varphi_N \rightarrow 0$, czyli dla dużych koncentracji domieszek donorowych, poziom Fermiego zbliża się do dna pasma przewodnictwa.

W zupełnie analogiczny sposób można wykazać, że dla półprzewodnika domieszkowanego typu P:

$$F_P = -W_g - kT \ln \left(\frac{N_a}{N_v} \right) \quad (2.26)$$

co oznacza, że poziom Fermiego w miarę zwiększania koncentracji akceptorów zbliża się do wierzchołka pasma podstawowego.

2.8. Efektywne gęstości stanów

Gęstości efektywne stanów wyrażone są wzorami:

$$N_c = 2 \left(\frac{2 \pi m_n^* kT}{h^2} \right)^{\frac{3}{2}} \quad (2.27)$$

$$N_v = 2 \left(\frac{2 \pi m_p kT}{h^2} \right)^{\frac{3}{2}} \quad (2.28)$$

W powyższych wzorach wartości m_n^* oraz m_p^* reprezentują tzw. masy efektywne dla gęstości stanów, które wyraża następujący wzór:

$$m_d = \sqrt[3]{m_i m_i^2}$$

Jak już wiemy, w pasmie przewodnictwa germanu istnieją cztery minima energetyczne. Liczba tych minimów powiększa czterokrotnie efektywną gęstość stanów, tak że ostatecznie:

$$N_c = 2 \cdot 4 \left(\frac{2 \pi m_d kT}{h} \right)^{\frac{3}{2}} \text{ dla germanu typu N} \quad (2.29)$$

$$N_c = 2 \cdot 6 \left(\frac{2 \pi m_d kT}{h} \right)^{\frac{3}{2}} \text{ dla krzemu typu N} \quad (2.30)$$

Wartości m_d dla elektronów i dziur w germanie i krzemie podane są w tablicy 2.

Tablica 2

Wartości m_d dla elektronów i dziur w germanie i krzemie

Półprzewodnik	Masa efektywna $\frac{m_d}{m_0}$	
	dla elektronów	dla dziur
German	0,22	0,39
Krzem	0,33	0,55

Podstawienie wartości m_d do wzorów (2.29) i (2.30) daje:

$$\text{dla germanu} \begin{cases} N_c = 2,00 \cdot 10^{15} T^{\frac{3}{2}} = 1,04 \cdot 10^{19} \left(\frac{T}{300} \right)^{\frac{3}{2}} \text{ cm}^{-3} & (2.31) \\ N_v = 1,17 \cdot 10^{15} T^{\frac{3}{2}} = 6,10 \cdot 10^{19} \left(\frac{T}{300} \right)^{\frac{3}{2}} \text{ cm}^{-3} & (2.32) \end{cases}$$

$$\text{dla krzemu} \begin{cases} N_c = 5,4 \cdot 10^{15} T^{\frac{3}{2}} = 2,8 \cdot 10^{19} \left(\frac{T}{300} \right)^{\frac{3}{2}} \text{ cm}^{-3} & (2.33) \\ N_v = 1,95 \cdot 10^{15} T^{\frac{3}{2}} = 1,01 \cdot 10^{19} \left(\frac{T}{300} \right)^{\frac{3}{2}} \text{ cm}^{-3} & (2.34) \end{cases}$$

2.9. Półprzewodniki zdegenerowane

Dla bardzo dużych koncentracji N_a i N_d poziom Fermiego może zająć położenie w pasmie podstawowym lub w pasmie przewodnictwa. Tego rodzaju półprzewodniki nazywają się zdegenerowanymi.

Dla wyznaczania położenia poziomu Fermiego w półprzewodniku zdegenerowanym typu N korzystamy ze wzoru:

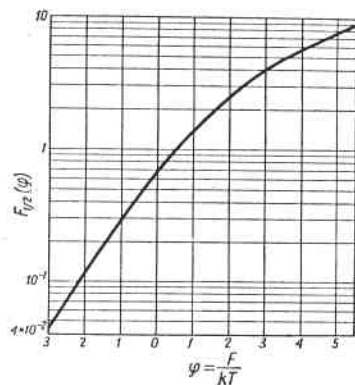
$$n = N_d = \frac{2}{\sqrt{\pi}} N_c \int_0^{\infty} \frac{\eta^{\frac{1}{2}} d\eta}{e^{\eta - \varphi} + 1}$$

$$\text{Całka } F_{\frac{1}{2}}(\varphi) = \int_0^{\infty} \frac{\eta^{\frac{1}{2}} d\eta}{e^{\eta - \varphi} + 1} \text{ nazywa się całką Fermiego-Diraca.}$$

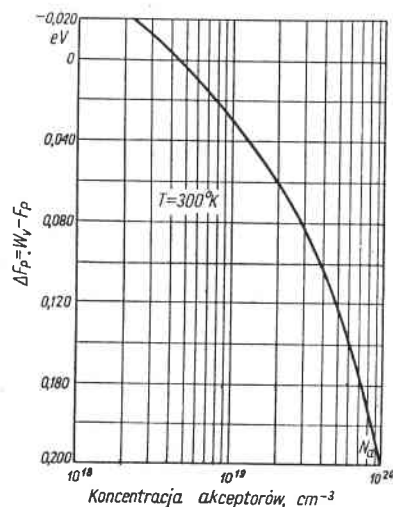
Wykres $F_{\frac{1}{2}}(\varphi)$ przedstawiony jest na rys. 2-7, a wartości liczbowe podane są w Uzupełnieniu I. Zależność poziomu Fermiego w germanie typu N od koncentracji domieszek przedstawiona jest na rys. 2-8, zaś w germanie typu P — na rys. 2-9.

Krytyczne koncentracje domieszek, dla których poziom Fermiego zajmuje położenie na granicy pasma przewodnictwa lub podstawowego są następujące (dla $T = 300 \text{ }^\circ\text{K}$):

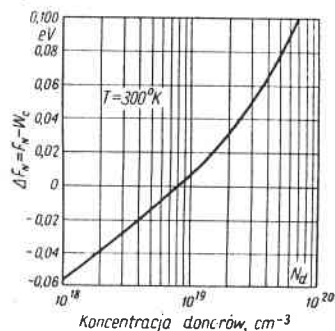
$$\begin{aligned} \text{German} & \begin{cases} N_d = 8 \cdot 10^{18} \text{ cm}^{-3} \\ N_a = 4,7 \cdot 10^{18} \text{ cm}^{-3} \end{cases} \\ \text{Krzem} & \begin{cases} N_d = 2,2 \cdot 10^{19} \text{ cm}^{-3} \\ N_a = 8 \cdot 10^{18} \text{ cm}^{-3} \end{cases} \end{aligned} \quad (2.35)$$



Rys. 2-7. Wykres całki Fermiego-Diraca



Rys. 2-9. Położenie poziomu Fermiego w germanie typu P



Rys. 2-8. Położenie poziomu Fermiego w germanie typu N

2.10. Oporność właściwa półprzewodników

Oporność właściwa półprzewodnika typu N zależy od koncentracji elektronów n oraz ich ruchliwości μ_n

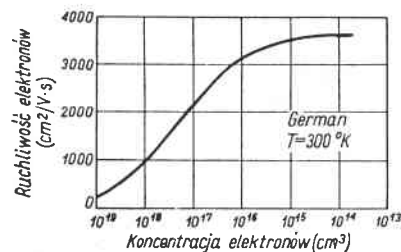
$$\rho_n = qn \mu_n \quad (2.36)$$

Podobnie dla półprzewodnika typu P:

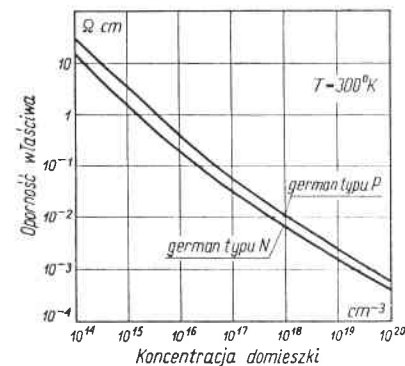
$$\rho_p = qp \mu_p \quad (2.37)$$

gdzie μ_p — ruchliwość dziur.

Ruchliwość nośników uzależniona jest od dwóch zasadniczych mechanizmów rozpraszania: na domieszkach zjonizowanych oraz na drganiach cieplnych sieci. W półprzewodnikach zdegenerowanych decydują-



Rys. 2-10. Zależność ruchliwości elektronów od koncentracji domieszek w germanie



Rys. 2-11. Zależność oporności właściwej germanu od koncentracji domieszek

cym mechanizmem jest — jak można sądzić — rozpraszanie na zjonizowanych donorach i akceptorach. Świadczyć może o tym znaczny spadek ruchliwości przy wzroście koncentracji domieszek (rys. 2-10). W czystym, niemal samoistnym germanie:

$$\mu_n = 3700 \text{ cm}^2/\text{V s}$$

$$\mu_p = 1800 \text{ cm}^2/\text{V s}$$

Dla koncentracji $N_d = 10^{19} \text{ cm}^{-3}$ w germanie ruchliwość elektronów spada do wartości poniżej $200 \text{ cm}^2/\text{V s}$.

Normalnie oporność właściwa półprzewodników zależy w znacznym stopniu od temperatury; dla półprzewodników zdegenerowanych oporność tę można uważać praktycznie za stałą.

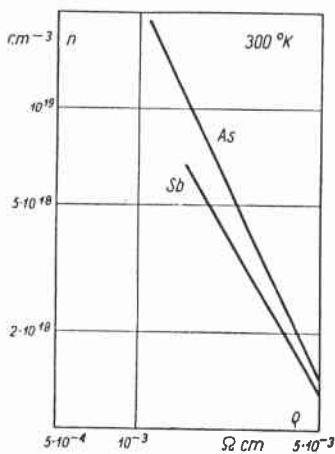
Zależność oporności właściwej germanu z płytkami domieszkami (rys. 2-11) ujęta jest empirycznym wzorem [A10]:

$$\rho = \frac{1}{B} N^{-\alpha} \quad (2.38)$$

Wartości współczynników B oraz α dla różnych zakresów koncentracji domieszek podane są w tablicy 3.

Wartość współczynników B oraz α dla różnych zakresów koncentracji domieszek

Zakres N	Współczynniki	
	B	α
German typu N		
$10^{14} - 10^{15}$	$2,74 \cdot 10^{-15}$	0,957
$10^{15} - 10^{16}$	$1,22 \cdot 10^{-14}$	0,914
$10^{16} - 10^{17}$	$5,05 \cdot 10^{-13}$	0,813
$10^{17} - 10^{18}$	$1,70 \cdot 10^{-10}$	0,664
$10^{18} - 10^{20}$	$1,74 \cdot 10^{-9}$	0,608
German typu P		
$10^{14} - 10^{16}$	$1,61 \cdot 10^{-15}$	0,950
$10^{16} - 10^{17}$	$6,28 \cdot 10^{-14}$	0,851
$10^{17} - 10^{18}$	$1,74 \cdot 10^{-11}$	0,707
$10^{18} - 10^{19}$	$1,61 \cdot 10^{-10}$	0,653
$10^{19} - 10^{20}$	$1,11 \cdot 10^{-9}$	0,609

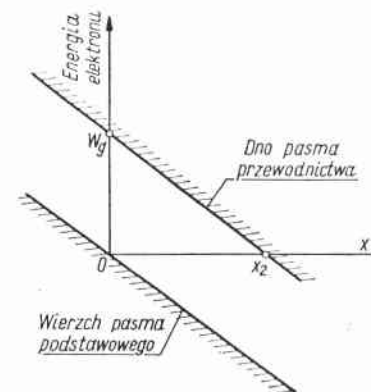


Rys. 2-12. Zależność oporności właściwej germanu typu N od koncentracji arsenu oraz atymonu (wg Furukawy)

Furukawa [A12] wykazał jednak, że domieszkowanie germanu różnymi domieszkami donorowymi o tej samej koncentracji daje nieco różne wartości oporności właściwej (rys. 2-12).

2.11. Zjawisko Zenera

Pod wpływem doprowadzonego silnego pola elektrycznego w półprzewodniku może wystąpić zjawisko wewnętrznej emisji polowej, polegające na przenikaniu tunelowym elektronów z pasma podstawowego do pasma przewodnictwa. Szczególnie prostym przykładem występowania zjawiska wewnętrznej emisji polowej jest przebicie izolatora. W izolatorach pasmo podstawowe jest całkowicie wypełnione, zaś pasmo



Rys. 2-13. Pochylenie pasm energetycznych w silnym polu elektrycznym

przewodnictwa puste. Zewnętrzne pole elektryczne powoduje pochylenie pasm energetycznych (rys. 2-13), umożliwiając tym samym emisję elektronów z pasma podstawowego do pasma przewodnictwa. Zjawisko wewnętrznej emisji polowej rozpatrzył po raz pierwszy angielski fizyk Clarence Zener [D1] w 1934 roku. W literaturze naukowo-technicznej mówi się wprost o zjawisku Zenera. Ma ono zasadnicze znaczenie w działaniu diody tunelowej.

Prawdopodobieństwo międzypasmowego przenikania elektronów można łatwo obliczyć wykorzystując wzór (1.35):

$$P \approx e^{-\frac{4\pi\sqrt{2m^*}}{h} \int_{x_1}^{x_2} \sqrt{W_p(x) - W} dx}$$

Podstawiając

$$W_p(x) = W_g - qKx \quad (2.39)$$

oraz dobierając układ współrzędnych tak, że $W = 0$, znajdujemy, że w granicach

$$x_1 = 0 \quad x_2 = \frac{W_g}{qK}$$

$$P = e^{-\frac{8\pi}{3} \sqrt{2m^*} \frac{W_g^2}{qK}} \quad (2.40)$$

Z powyższego wzoru widzimy, że przenikanie tunelowe jest szczególnie korzystne dla małych wartości przerwy energetycznej i małej masy efektywnej elektronu. Prawdopodobieństwo przenikania tunelowego zwiększa się bardzo szybko ze wzrostem natężenia pola elektrycznego.

Gęstość prądu w zjawisku Zenera można obliczyć, wychodząc z rozpatrzenia zmiany wektora falowego elektronu w czasie, uwarunkowanej przyłożeniem zewnętrznego pola elektrycznego:

$$\frac{dk}{dt} = \frac{2\pi qK}{h} \quad (2.41)$$

Zmiana wektora falowego następuje w granicach od $-\frac{\pi}{a}$ do $+\frac{\pi}{a}$ (rys. 2-2), a więc na drodze $\frac{2\pi}{a}$.

Okresem oscylacji jest

$$T = \frac{\text{droga}}{\text{prędkość}} = \frac{\frac{2\pi}{a}}{\frac{2\pi qK}{h}} = \frac{h}{aqK} \quad (2.42)$$

W dostatecznie silnym polu elektrycznym częstotliwość oscylacji jest bardzo duża, rzędu 10^{13} s^{-1} .

Gęstość prądu tunelujących elektronów wyraża wzór:

$$J = \frac{zq^2}{a^2h} UP \quad (2.43)$$

gdzie: z — liczba elektronów w komórce jednostkowej kryształu,

U — spadek napięcia, a — stała siatki krystalicznej,

P — prawdopodobieństwo przenikania tunelowego.

Jak już podkreślono, w izolatorze pasmo przewodnictwa jest puste. W przypadku półprzewodnika elektrony z pasma przewodnictwa mogą tunelować do pustych stanów w pasmie podstawowym. W tym przypadku mówimy o zjawisku Esakiego.

3. PÓLPRZEWODNIKOWE ZŁĄCZE P-N

3.1. Rodzaje złącz

Granica dzieląca obszary P i N w półprzewodniku nazywa się złączem $P-N$ lub przejściem $P-N$. W przypadku gdy oba obszary P i N mają zbliżone wartości koncentracji domieszek, mówimy o złączu symetrycznym. W przeciwnym przypadku, najczęściej zresztą spotykanym w praktyce, mamy do czynienia ze złączem asymetrycznym, które oznaczamy symbolem P^+N lub PN^+ , przy czym znak $+$ oznacza obszar silniej do-

mieszkowany, a więc o znacznie mniejszej oporności właściwej. Nazwę „złącze” rozszerzamy również na przypadek granicy dzielącej obszary PP^+ względnie NN^+ , w którym oba obszary złącza są tego samego typu, lecz różnią się wyraźnie koncentracjami akceptorów albo donorów. Szczególny rodzaj złącza przedstawia tzw. heterozłącze, w którym oba obszary P i N stanowią różne półprzewodniki (np. german i arsenek galowy) różniące się wartością przerwy energetycznej.

Istnieje wiele metod wykonywania złącz $P-N$. Z punktu widzenia techniki diod tunelowych najbardziej jednak istotną jest metoda stopowa, polegająca na wtopieniu domieszki przeciwnego typu do półprzewodnika. Typowym tu przykładem jest wtopienie indu, a więc domieszki akceptorowej do germanu typu N . German rozpuszcza się w ciekłym indu i po rekrystalizacji powstaje obszar P o dużej stosunkowo koncentracji atomów indu. Szczególną cechą metody stopowej jest to, że uzyskane złącze jest bardzo strome. W przypadku stosowania innych metod złącze jest „rozmyte” i charakteryzuje się np. liniowym gradientem koncentracji domieszek.

3.2. Szerokość warstwy ładunku przestrzennego w złączu $P-N$

Utworzenie złącza $P-N$ spowoduje dyfuzję nośników z obu obszarów: dziur z obszaru P do obszaru N , a elektronów z obszaru N do obszaru P . Dalszy przepływ elektronów i dziur zostaje zahamowany przez pole elektryczne. Tym samym nośniki z bezpośredniego sąsiedztwa złącza zostają usunięte i tworzy się podwójny obszar ładunku przestrzennego zjonizowanych donorów i akceptorów. Ze względu na brak nośników prądu oporność właściwa tego obszaru będzie bardzo duża. Wytworzona różnica potencjałów na złączu nazywa się napięciem dyfuzyjnym U_D . W stanie równowagi (brak przyłożonego napięcia zewnętrznego) poziomy Fermiego znajdują się na tej samej wysokości ($F_N = F_P$).

W dalszym ciągu będzie nas interesowało wyznaczenie szerokości warstwy podwójnej ładunku przestrzennego. Przyjmijmy jednowymiarowy model złącza $P-N$. Potencjał $\psi(x)$ i gęstość ładunku przestrzennego ρ związane są ze sobą równaniem Poissona, które w układzie jednostek MKS będzie miało postać:

$$\frac{d^2 \psi(x)}{dx^2} = - \frac{\rho(x)}{\epsilon_0 \epsilon} \quad (3.1)$$

przy czym ϵ_0 — przenikalność dielektryczna próżni,

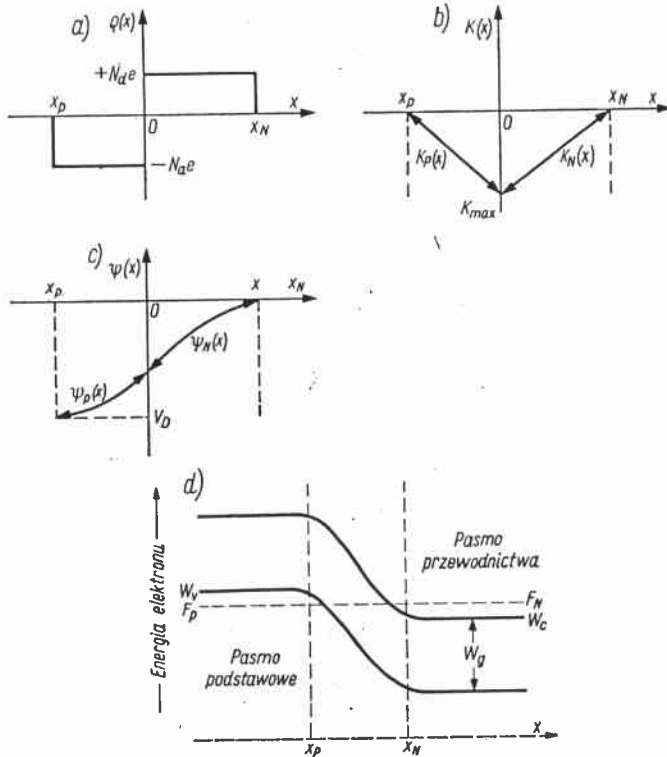
ϵ — stała dielektryczna półprzewodnika.

Ponieważ natężenie pola elektrycznego $\vec{K} = - \text{grad } \psi$, to:

$$\text{div } \vec{K} = \frac{\rho(x)}{\epsilon_0 \epsilon} \quad (3.2)$$

Dla złącza stromeo (rys. 3-1a) można napisać:

$$\begin{aligned} -qN_a &= \rho(x) & \text{dla } x_p \leq x \leq 0 \\ +qN_d &= \rho(x) & \text{dla } 0 \leq x \leq x_N \end{aligned} \quad (3.3)$$



Rys. 3-1. Model złącza P-N: a) gęstość ładunku przestrzennego donorów i akceptorów; b) przebieg natężenia pola elektrycznego; c) przebieg potencjału elektrycznego; d) model energetyczny złącza P-N

Natężenie pola elektrycznego (rys. 3-1b) osobno dla tych dwóch obszarów, wyprowadzamy ze wzoru (3.2):

$$\left. \begin{aligned} K_P &= -\frac{qN_a}{\epsilon_0 \epsilon} (x + x_p) \\ K_N &= \frac{qN_d}{\epsilon_0 \epsilon} (x - x_N) \end{aligned} \right\} \quad (3.4)$$

Potencjał

$$\left. \begin{aligned} \psi_P &= -\int K_P dx \\ \psi_N &= -\int K_N dx \end{aligned} \right\} \quad (3.5)$$

Scalkowanie wzoru (3.5) daje:

$$\psi_P(x) = \frac{N_a q}{2 \epsilon_0 \epsilon} (x - x_p)^2 - U_D + U \quad (3.6)$$

Skorzystalismy tu z warunku brzegowego:

$$\psi = -U_D + U \quad \text{dla } x = x_p$$

gdzie U jest napięciem dodatnim przylozonym do obszaru P, a U_D — napięciem dyfuzyjnym.

Podobnie korzystajac z warunku brzegowego

$$\psi = 0 \quad \text{dla } x = x_N$$

otrzymujemy ze scalkowania wzoru (3.5):

$$\psi_N = -\frac{N_d q}{2 \epsilon_0 \epsilon} (x - x_N)^2 \quad (3.7)$$

Wykres zaleznosci (3.6) i (3.7) przedstawiony jest na rys. 3-1c.

Poniewaz dla $x = 0$, $K_P = K_N$, to warunkiem ciaglosci natężenia pola elektrycznego bedzie:

$$N_a x_p = N_d x_N \quad (3.8)$$

Podobnie otrzymamy warunek ciaglosci potencjalu dla $x = 0$, przyjmujac:

$$\psi_P(0) = \psi_N(0)$$

tak ze

$$\frac{N_a q}{2 \epsilon_0 \epsilon} x_p^2 - U_D + U = -\frac{N_d q}{2 \epsilon_0 \epsilon} x_N^2 \quad (3.9)$$

Podstawiajac wyrazenie (3.8) do (3.9) i rozwiazujac wzgledem x_p , otrzymujemy:

$$\begin{aligned} x_p &= \sqrt{\frac{2 \epsilon_0 \epsilon}{q} \cdot \frac{U - U_D}{\left(1 + \frac{N_a}{N_d}\right) N_a}} \\ x_N &= \sqrt{\frac{2 \epsilon_0 \epsilon}{q} \cdot \frac{U - U_D}{\left(1 + \frac{N_d}{N_a}\right) N_d}} \end{aligned}$$

Szerokosc d warstwy ladunku przestrzennego:

$$d = x_N - x_p, \text{ tak ze}$$

$$d = \sqrt{\frac{2 \epsilon_0 \epsilon}{q} \cdot \frac{N_a + N_d}{N_a \cdot N_d} (U - U_D)} \quad (3.10)$$

Napięcie dyfuzyjne wyrażone jest wzorem:

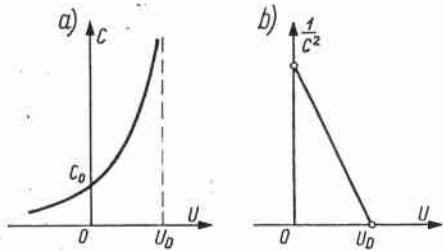
$$U_D = kT \ln \frac{n_n}{n_p} \quad (3.11)$$

w którym $\frac{n_n}{n_p}$ jest stosunkiem koncentracji elektronów w obszarze P i N. Na przykład dla $T = 300 \text{ }^\circ\text{K}$, $n_n = 10^{16}$, $n_p = 10^{10}$ otrzymujemy $U_D = 0,36 \text{ V}$. Wzór (3.11) nie jest słuszny w przypadku półprzewodników zdegenerowanych.

3.3. Pojemność złącza P-N

W warstwie przestrzennego ładunku złącza istnieje ładunek dodatni $+Q$ w obszarze N oraz ładunek $-Q$ w obszarze P:

$$Q = qN_d x_N = 2q \epsilon_0 \epsilon (U - U_D) \frac{N_a N_d}{N_a + N_d}$$



Rys. 3-2. Pojemność stromego złącza P-N: a) zależność pojemności od napięcia; b) zależność $\frac{1}{C^2}$ od napięcia

Pojemność (na jednostkę powierzchni złącza) jest równa:

$$C = \frac{dQ}{dU} = \sqrt{\frac{q \epsilon_0 \epsilon}{2} \cdot \frac{N_a N_d}{N_a + N_d} \cdot \frac{1}{(U - U_D)}} = \frac{\epsilon_0 \epsilon}{d} \quad (3.12)$$

Warstwę ładunku przestrzennego można więc traktować jak płaski kondensator o odległości d między płytkami. Pojemność diody (rys. 3-2a) można wyrazić wzorem:

$$C = \frac{C_0}{\sqrt{1 - \left(\frac{U}{U_D}\right)}} \quad (3.13)$$

gdzie $C = C_0$ dla $U = 0$. Dla $U \rightarrow U_D$, $C \rightarrow \infty$; dla $U \rightarrow \infty$; $C \rightarrow 0$. Ważne jest to, że dla złącza stromego pojemność zależy od napięcia, jak:

$$C \sim U^{-\frac{1}{2}} \quad (3.14)$$

W przypadku złącza nie stromego, lecz np. o liniowej zmianie koncentracji domieszek w funkcji odległości x , otrzymuje się:

$$C \sim U^{-3}$$

Wykreślenie $\frac{1}{C^2}$ w funkcji przyłożonego napięcia U daje linię prostą dla złącza stromego (rys. 3-2b).

3.4. Model energetyczno-pasmowy złącza P-N

W stanie równowagi poziomy Fermiego w obszarach P i N muszą się znaleźć na tej samej wysokości. Przy założeniu, że oba obszary są zdegenerowane, przebieg poziomów energetycznych będzie taki, jak na rys. 3-1d, przy czym qU_D jest wysokością bariery energii potencjalnej.

Elektrony pobudzone cieplnie w pasmie przewodnictwa w obszarze P mogą płynąć „w dół” do obszaru N, tworząc tzw. prąd generacyjny I_g , proporcjonalny do liczby elektronów (nośników mniejszościowych w obszarze P). Równocześnie elektrony w pasmie przewodnictwa z obszaru N płyną „do góry” do obszaru P i tam rekombinują z dziurami. Elektrony te muszą jednak pokonać barierę. Prąd ten jest prądem rekombinacyjnym I_r . W stanie równowagi oba prądy elektronowe znoszą się. Podobnie dzieje się w przypadku dziur.

Przyłożenie do obszaru P potencjału dodatniego powoduje obniżenie bariery. Prąd generacyjny wskutek tego się nie zmieni, jednak prąd rekombinacyjny znacznie wzrośnie o czynnik Boltzmanna $e^{\frac{qU}{kT}}$. W stanie równowagi $I_r = I_g$, obecnie zaś $I_r = I_g e^{\frac{qU}{kT}}$.

Prąd wypadkowy I jest równy różnicy $I_r - I_g$, tak że

$$I = I_r - I_g = I_g e^{\frac{qU}{kT}} - I_g = I_g (e^{\frac{qU}{kT}} - 1) \quad (3.15)$$

Podobnie przyłożenie potencjału w kierunku wstecznym (potencjał ujemny do obszaru P) spowoduje powiększenie wysokości bariery. Prąd rekombinacyjny elektronów zmaleje, prąd generacyjny pozostanie bez zmiany. Podobna sytuacja zaistnieje dla dziur.

3.5. Wzór Shockleya

Z dokładnej analizy przepływu dyfuzyjnego nośników przez złącze wynika następująca zależność gęstości natężenia prądu wypadkowego J w funkcji przyłożonego napięcia U :

$$J = J_s (e^{\frac{qU}{kT}} - 1) \quad (3.16)$$

gdzie:

$$J_s = q \left[\left(\frac{n_p}{\tau_n} \right) L_n + \left(\frac{p_n}{\tau_p} \right) L_p \right] \quad (3.17)$$

- n_p — koncentracja elektronów w obszarze P,
- p_n — koncentracja dziur w obszarze N,
- τ_n — czas życia elektronów w obszarze P,
- τ_p — czas życia dziur w obszarze N,
- $L_n = \sqrt{D_n \tau_n}$ — długość dyfuzyjna elektronów,
- $L_p = \sqrt{D_p \tau_p}$ — długość dyfuzyjna dziur,
- D_n — stała dyfuzyjna elektronów,
- D_p — stała dyfuzyjna dziur.

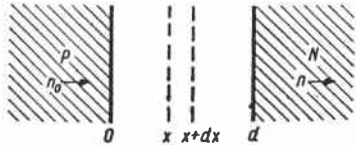
Z powyższego wzoru widać, że prąd zwiększa się bardzo szybko przy napięciach dodatnich przyłożonych do obszaru P (kierunek przedni). W kierunku wstecznym płynie natomiast mała gęstość prądu — J_s . Pełniejszą zgodność z doświadczeniem otrzymuje się niekiedy ze wzoru:

$$J = J_s (e^{\beta \frac{qU}{kT}} - 1)$$

w którym $0 < \beta \leq 1$.

3.6. Przebiecie złącza P-N

Przy dużych napięciach wstecznych prąd bardzo szybko osiąga duże wartości i następuje przebiecie. Wartość napięcia przebiecia, w zależności od oporności właściwej półprzewodnika, może się zawierać w bardzo sze-



Rys. 3-3. Lawinowe narastanie liczby nośników

rocznych granicach. Istota tego zjawiska jeszcze do niedawna nie była dostatecznie jasna. W chwili obecnej wiadomo, że w przypadku dostatecznie szerokich złącz, dla których koncentracje N_a, N_d są niezbyt duże, można przyjąć mechanizm lawinowego powielania nośników.

W silnym polu elektrycznym w złączu nośniki mogą nabrać tak znacznej energii, że przez zderzenie jonizują neutralne atomy. Aby jednak nośniki mogły uzyskać wystarczającą energię, złącze musi być dos-

tatecznie szerokie. Załóżmy, że liczba elektronów wprowadzonych w złącze P-N w punkcie $x = 0$ jest równa n_0 — rys. 3-3. Na drodze od 0 do x wytworzonych zostało n_1 elektronów. Podobnie na drodze od x do d powstało n_2 elektronów. Na odcinku dx mamy:

$$dn_1 = (n_0 + n_1) a_i dx + n_2 a_i dx = n a_i dx$$

przy czym $n = n_0 + n_1 + n_2$ — całkowita liczba elektronów,

a_i — szybkość jonizacji równa liczbie par dziura-elektron, wytworzonych przez nośnik na jednostkowym odcinku drogi.

Całkując powyższy wzór przy warunkach brzegowych:

$$n_1 = 0 \quad \text{dla } x = 0$$

$$n = n_1 + n_0 \quad \text{dla } x = d$$

otrzymujemy

$$1 - \frac{1}{M} = \int_0^d a_i dx \quad (3.18)$$

gdzie M — współczynnik powielania nośników:

$$M = \frac{n}{n_0} = \frac{1}{1 - \int_0^d a_i dx} \quad (3.19)$$

Dla $\int_0^d a_i dx \rightarrow 1$ współczynnik powielania dąży do nieskończoności i wówczas następuje przebiecie złącza P-N. W złączach wąskich przebiecie lawinowe nie nastąpi nawet przy przekroczeniu krytycznego natężenia pola elektrycznego. Wzór (3.18) został wyprowadzony przy założeniu, że nie zachodzi rekombinacja w złączu. Jest to uzasadnione tym, że czas przelotu nośnika przez złącze w silnym polu jest bardzo mały, rzędu 0,1 ns. Założyliśmy również, że szybkość jonizacji a_i zależy tylko od natężenia pola elektrycznego.

W złączach bardzo wąskich występuje znany już z rozdziału 2 mechanizm emisji zenerowskiej. Z nowszych badań wynika, że w złączach germanowych przebiecie zenerowskie występuje przy oporności właściwej:

$$\rho < 0,5 \Omega \text{ cm}$$

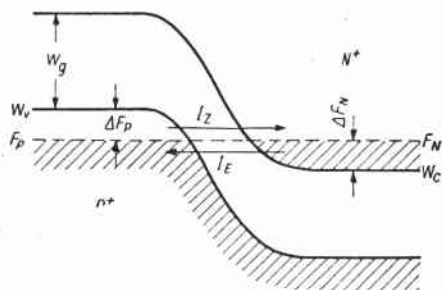
przy czym między ρ i napięciem przebiecia istnieje empiryczna zależność:

$$U_{prz.} = 88 \rho^{0,64}$$

4. WŁASNOŚCI DIOD TUNELOWYCH

4.1. Model pasmowy zdegenerowanego złącza P-N

Jak podano w rozdziale 2 w półprzewodnikach bardzo silnie domieszkowanych poziom Fermiego zajmuje położenie w pasmie przewodnictwa (typ N) lub w pasmie podstawowym (typ P). Dzięki temu, w stanie równowagi następuje w złączu P-N zachodzenie na siebie pasma podstawowego i pasma przewodnictwa. Z drugiej strony, na podstawie wzoru (3.10)



Rys. 4-1. Pasmowy model energetyczny złącza zdegenerowanego

można stwierdzić, że przy bardzo dużych koncentracjach domieszek szerokość warstwy ładunku przestrzennego staje się bardzo mała. Na przykład, dla $N_a = N_d = 5 \cdot 10^{19} \text{ cm}^{-3}$ otrzymujemy $d = 70 \text{ \AA}$. Przez tak wąskie złącze P-N mogą elektrony z łatwością przenikać na zasadzie zjawiska tunelowego. Mówimy, że przezroczystość bariery jest duża, lub że duże jest prawdopodobieństwo przenikania tunelowego. Te dwa czynniki warunkują działanie diody tunelowej.

Przebieg charakterystyki prądowo-napięciowej staje się zrozumiały w oparciu o pasmowy model energetyczny, przedstawiony na rys. 4-1. W stanie równowagi poziomy Fermiego znajdują się na tej samej wysokości. Dzięki zachodzeniu na siebie pasm możliwe jest „tunelowanie” z pasma podstawowego do pasma przewodnictwa. Wynikający stąd prąd nazywa się prądem Zenera I_Z . Prąd elektronów z pasma przewodnictwa do pasma podstawowego nazywa się prądem Esakiego I_E . W stanie równowagi ($U = 0$) oba prądy muszą być równe, tak że prąd wypadkowy:

$$I = I_E - I_Z \quad (4.1)$$

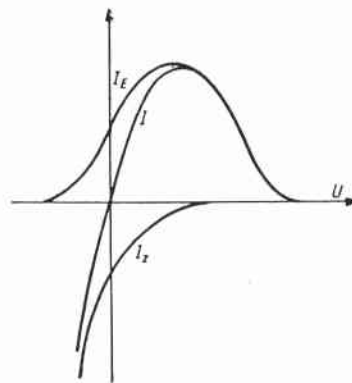
jest równy zeru.

Należy zwrócić uwagę, że w złączu zdegenerowanym bariera energii potencjalnej jest stosunkowo bardzo duża, znacznie większa niż w zwykłym złączu, jakie rozpatrywano w poprzednim rozdziale. Wysokość bariery w stanie równowagi jest qU_D , gdzie U_D wynika ze wzoru (3.11), tracącego słusność w przypadku półprzewodników zdegenerowanych. Na podstawie rys. 4-1 można stwierdzić, że wysokość bariery jest w przybliżeniu równa szerokości przerwy energetycznej półprzewodnika.

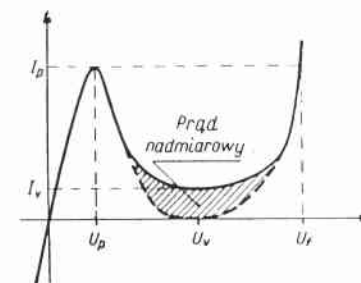
Zalóżmy obecnie, że do obszaru P zostało przyłożone dodatnie napięcie U i tym samym nastąpiło obniżenie pasm. Prąd I_E wzrośnie i jednocześnie zmniejszy się prąd Zenera, tak że prąd wypadkowy zostanie zwiększony. Przy pewnym napięciu dodatnim prąd wypadkowy osiągnie wartość maksymalną, zaś prąd Zenera spadnie praktycznie do zera. Dalsze zmniejszenie się prądu I uwarunkowane jest wyłącznie prądem I_E . Przy napięciu U równym:

$$U = \frac{\Delta F_N + \Delta F_P}{q} \quad (4.2)$$

co odpowiada warunkowi $W_c = W_v$, prąd powinien spaść do zera. Dwie składowe natężenia prądu w diodzie tunelowej przedstawione są na rys. 4-2. Tego rodzaju jakościowe wyjaśnienie przebiegu charakterystyki diody tunelowej tylko częściowo jest zgodne z rzeczywistą krzywą $I-U$.



Rys. 4-2. Dwie składowe natężenia prądu w diodzie tunelowej



Rys. 4-3. Prąd upływu w diodzie tunelowej

W diodzie tunelowej nie występuje spadek natężenia prądu do wartości zerowej (dla $W_c = W_v$). Nawet najstaranniej wykonaną diodę tunelową cechuje pewien prąd upływu i następnie ponowny wykładniczy wzrost prądu (rys. 4-3). Wzmiankowany prąd upływu nazwany został prądem nadmiarowym. Pochodzenie jego nie jest w chwili obecnej dostatecznie wyjaśnione. Wykładniczy wzrost prądu dla wyższych napięć stanowi prąd w kierunku przednim zwykłego złącza P-N zgodny ze wzorem Shockleya. Jego istnienie stanie się zrozumiałe, jeśli weźmiemy pod uwagę, że dla wyższych napięć dodatnich bariera staje się dostatecznie niska i wtedy następuje zwykły przepływ dyfuzyjny nośników.

Przy napięciach ujemnych przyłożonych do obszaru P pasma zostają przesunięte w górę i wtedy dominuje prąd Zenera, tak że mamy do czynienia z emisją elektronów z pasma podstawowego do pasma prze-

wodnictwa. Z rys. 4-3 widać, że w diodzie tunelowej płynie większy prąd w kierunku wstecznym niż w kierunku przednim, gdzie prąd osiąga wartość I_p i następnie maleje. Można uważać, że kierunek łatwego przepływu w diodzie tunelowej jest „odwrócony”. W związku z tym warto wspomnieć, że pierwsze teorie prostowników półprzewodnikowych opracowane przed trzydziestu laty zakładały istnienie mechanizmu tunelowego. Teorie te zostały jednak z czasem zarzucone, gdyż wskazywały na przeciwny kierunek łatwego przepływu prądu.

4.2. Parametry diod tunelowych

Krzywą $I-U$ diody tunelowej charakteryzują dwa punkty:

- szczytowy (I_p, U_p)¹⁾
- dolinowy (I_v, U_v)

Zakres charakterystyki dla napięć od 0 do U_p nazywamy zakresem niższych napięć, zaś dla napięć $U > U_p$ mamy zakres wyższych napięć. Zakres środkowy $U_p < U < U_v$ odpowiada oczywiście ujemnej oporności.

Napięcie U_p w przypadku diod germanowych wynosi najczęściej 55 mV, może jednak mieścić się w granicach 30 ÷ 80 mV. W diodach arseno-galowych $U_p \approx 100$ mV. Prąd szczytowy I_p zależy od gęstości prądu tunelowego oraz od powierzchni złącza. W diodach tunelowych znajdujących typowe zastosowania, I_p mieści się w granicach 1 ÷ 10 mA.

Napięcie doliny U_v jest różne dla diod wykonanych z różnych półprzewodników i zależy od wartości przerwy energetycznej. Typowe wartości U_v są następujące:

Ge	— 350 mV
Si	— 450 mV
GaAs	— 700 mV

Wartość prądu I_v zależy od powierzchni diody oraz od materiału półprzewodnikowego. Stosunek $\frac{I_p}{I_v}$ charakteryzuje w pewnym stopniu jakość diody tunelowej. Typowe wartości $\frac{I_p}{I_v}$ są następujące:

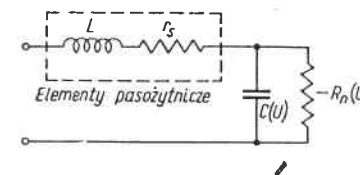
Ge	— 5 (max. 20)
Si	— 4
GaAs	— 15 (max. 70)

¹⁾ Od angielskich wyrazów: peak (szczyt) oraz valley (dolina).

Niekiedy pożyteczna jest znajomość wartości napięcia U_f ¹⁾ odpowiadająca natężeniu prądu $I_f = I_p$ w zakresie wyższych napięć. Typowe wartości U_f są następujące:

Ge	— 450 mV
Si	— 800 mV
GaAs	— 1000 mV
InSb	— 200 mV
GaSb	— 450 mV

Rys. 4-4. Układ zastępczy diody tunelowej



Napięcie U_f zależy, podobnie jak U_v , od wartości przerwy energetycznej półprzewodnika.

Oprócz parametrów I_p, I_v, U_v, U_p oraz ewentualnie U_f szczególnie ważne są parametry układu zastępczego diody dla małych sygnałów (rys. 4-4). Złącze $P^+ - N^+$ charakteryzuje ujemna oporność różniczkowa:

$$-R_n = \frac{1}{\frac{dI}{dU}} \quad (4.3)$$

albo ujemna przewodność różniczkowa:

$$-G_n = \frac{dI}{dU} \quad (4.4)$$

oraz pojemność C . Ujemna oporność $-R_n$ zależy w pierwszym rzędzie od wartości prądu szczytowego I_p oraz w mniejszym stopniu od U_p, U_v, I_v . Przybliżoną wartość minimalną oporności określa wzór:

$$R_{n \min} \approx 0,4 \frac{U_v - U_p}{I_p - I_v} \quad (4.5)$$

Wartość R_n mieści się najczęściej w granicach od kilku do kilkuset omów.

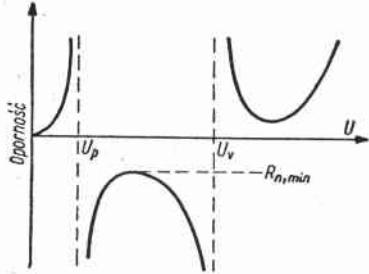
Interesujące zależności dla diody tunelowej wynikają z wykresów $G_n = f(U)$ lub $R_n = f(U)$, przedstawionych na rys. 4-5 i rys. 4-6. Widzimy, że w punktach U_p oraz U_v przewodność różniczkowa jest równa zero (oporność nieskończona). Oporność R_n jest nieliniowa; można jednak odróżnić na charakterystyce $I-U$ odcinek prawie liniowy, dla którego (umownie) średnia wartość $\frac{\Delta U}{\Delta I}$ (rys. 4-7) nie przekracza wartości oporności R_n o więcej niż np. 15%.

¹⁾ Od angielskiego wyrazu forward (przedni).

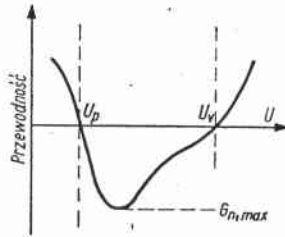
Pojemność C diody tunelowej jest proporcjonalna do powierzchni złącza i zależy od koncentracji domieszek po obu stronach złącza:

$$C \sim \sqrt{\frac{N_a N_d}{N_a + N_d}} \quad (4.6)$$

Najmniejsze uzyskane wartości C są rzędu 1 pF. Zarówno R_n jak i C zależą od napięcia (punktu pracy).

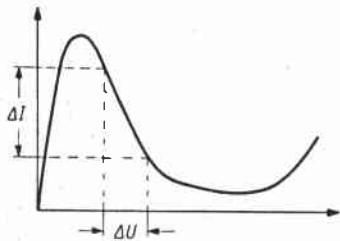


Rys. 4-5. Zależność oporności od napięcia w diodzie tunelowej



Rys. 4-6. Zależność przewodności od napięcia w diodzie tunelowej

Pozostałe elementy układu zastępczego mają charakter pasożytniczy. Są to oporność szeregową r_s oraz indukcyjność L . Szeregową oporność dodatnią r_s zależy od powierzchni złącza, oporności właściwej materiału



Rys. 4-7. Zakres liniowy oporności ujemnej

półprzewodnikowego w diodzie, oporności na styku bazy i doprowadzeń oraz rozptywu linii prądu. Typowa wartość r_s wynosi 0,5 Ω ; w diodach mikrofalowych o bardzo małej powierzchni złącza wartość r_s może wynosić nawet 8 Ω . Od oporności r_s zależy w pewnym stopniu częstotliwość graniczna diody:

$$f_{gr} \sim \frac{1}{R_n C} \sqrt{\frac{R_n}{r_s} - 1} \quad (4.7)$$

Wskutek zjawiska naskórkowego oporność r_s jest funkcją częstotliwości. Indukcyjność L diody odgrywa bardzo ważną rolę w praktycznym zastosowaniu diod. Dla zmniejszenia indukcyjności złącze P^+N^+ powinno być zmontowane bezpośrednio między dwiema płytkami, a to

w celu uniknięcia stosowania cienkich doprowadzeń wewnątrz oprawki diody. W produkowanych diodach indukcyjność L waha się w granicach od ułamka nanohenra do kilkunastu nH (1 nH = 10^{-9} H).

Innymi interesującymi parametrami diody tunelowej są: J_p , $R_n C$, $I_0 R_n$, $\frac{I_p}{C}$.

Gęstość prądu szczytowego $J_p = \frac{I_p}{S}$ zależy przede wszystkim od prawdopodobieństwa przenikania tunelowego, które jest duże dla złącz bardzo wąskich, a więc bardzo silnie domieszkowanych. J_p określa w dużej mierze ogólną jakość diody tunelowej. Duża wartość J_p pozwala uzyskać diody o małej powierzchni i dostatecznie małej wartości R_n . Uzyskiwane wartości J_p sięgają nawet dziesiątków tysięcy amperów na centymetr kwadratowy powierzchni złącza. Mała wartość iloczynu $R_n C$ (stała czasu) warunkuje możliwość pracy przy bardzo wielkich częstotliwościach, w szczególności zaś krótki czas przełączania oraz dużą wartość iloczynu wzmocnienia i szerokości pasma we wzmacniaczach tunelowych. Stała czasu $R_n C$ nie zależy zasadniczo od powierzchni S złącza (dla tej samej próbki półprzewodnika), ponieważ:

$$C \sim S \quad R_n \sim \frac{1}{S}$$

Małe wartości $R_n C$ łatwiej można uzyskać w półprzewodnikach o małej wartości przerwy energetycznej W_g .

Iloczyn prądu stałego diody I_0 oraz oporności R_n określa własności szumowe diody i dla pewnego napięcia optymalnego w zakresie oporności ujemnej przybiera wartość minimalną.

Iloraz $k = \frac{I_p}{C}$ ma wpływ na maksymalną moc wielkiej częstotliwości generowanej w diodzie tunelowej.

Maksymalne częstotliwości f w diodach tunelowych uzyskuje się dla możliwie małych wartości C , R_n , r_s . Granicznym czynnikiem warunkującym uzyskanie maksymalnego f jest czas relaksacji dielektrycznej w półprzewodniku:

$$\tau_{rd} = \epsilon_0 \epsilon \rho$$

Założmy dla przykładu, że $\rho = 7 \cdot 10^{-4} \Omega \text{cm} = 7 \cdot 10^{-6} \Omega \text{m}$. Ponieważ dla germanu $\epsilon = 16$, $\epsilon_0 = 8,854 \cdot 10^{-12} \text{ F/m}$ to:

$$\tau_{rd} = 7 \cdot 10^{-6} \cdot 1,4 \cdot 10^{-11}$$

czyli

$$\tau_{rd} = 10^{-16} \text{ s.}$$

Można również oszacować czas przelotu τ_p elektronu o średniej prędkości:

$$\bar{v} = \sqrt{\frac{3kT}{m^*}}$$

na drodze równej szerokości bariery d .

Jeśli na przykład $T = 300 \text{ K}$, $m^* = 0,2 m_0$, $d = 100 \text{ \AA} = 10^{-8} \text{ m}$, to:

$$\tau_p = \frac{d}{v} \approx 5 \cdot 10^{-14} \text{ s}$$

Tak krótkie odcinki czasu wynikają zarówno z małej oporności właściwej ρ półprzewodnika (co warunkuje małe d), jak i przede wszystkim stąd, że w procesie tunelowania biorą udział jedynie nośniki większościowe półprzewodnika.

4.3. Elementy teorii złącza zdegenerowanego P^+N^+

W zakresie napięć od 0 do U_v charakterystyka statyczna diody tunelowej różni się znacznie od charakterystyki zwykłej diody $P-N$. Jest to zrozumiałe, gdyż występuje zupełnie odmienne zjawisko fizyczne. Z tego też względu nie byłoby możliwe przeniesienie teorii Shockleya dotyczącej złącza $P-N$ na przypadek złącza zdegenerowanego P^+N^+ , przy czym mówiąc o degeneracji mamy na myśli degenerację obu obszarów złącza.

Wprowadźmy następujące oznaczenia:

- $D_c(W)$ gęstość stanów w pasmie przewodnictwa,
- $D_v(W)$ gęstość stanów w pasmie podstawowym,
- $f_c(W)$ funkcja rozkładu Fermiego-Diraca w pasmie przewodnictwa,
- $f_v(W)$ funkcja rozkładu Fermiego-Diraca w pasmie podstawowym,
- $P_{c \rightarrow v}$ prawdopodobieństwo przenikania tunelowego z pasma przewodnictwa do pasma podstawowego,
- $P_{v \rightarrow c}$ prawdopodobieństwo przenikania tunelowego z pasma podstawowego do pasma przewodnictwa.

Liczba elektronów w przedziale energii W , $W + dW$ w pasmie przewodnictwa wyraża się wzorem:

$$D_c(W) f_c(W) dW$$

Temu samemu przedziałowi energii dW odpowiada liczba nieobsadzonych stanów w pasmie podstawowym:

$$D_v(W) [1 - f_v(W)]$$

Zakresem energii, dla którego zachodzi tunelowanie, jest $W_v - W_c$, tak że prąd Esakiego I_E (rys. 4-8) wyrazi się wzorem:

$$I_E = A \int_{W_c}^{W_v} P_{c \rightarrow v} D_c(W) f_c(W) D_v(W) [1 - f_v(W)] dW$$

Podobnie liczba elektronów w przedziale energii dW w pasmie podstawowym wyraża się wzorem:

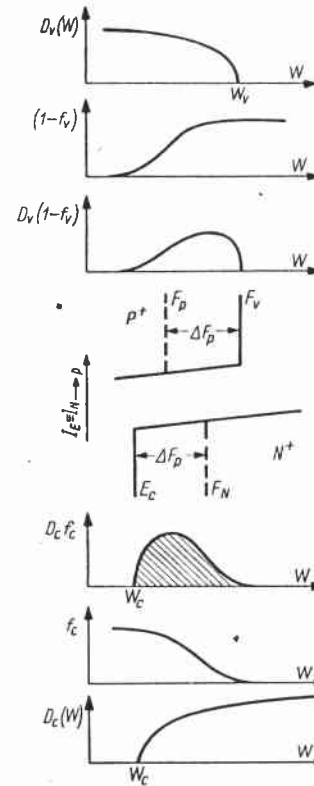
$$D_v(W) f_v(W) dW$$

zaś liczba nieobsadzonych stanów w pasmie przewodnictwa

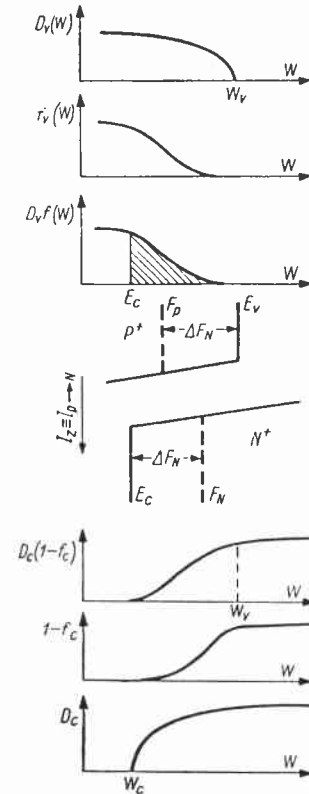
$$D_c(W) [1 - f_c(W)] dW$$

Zatem prąd Zenera I_Z (rys. 4-9) z pasma podstawowego do pasma przewodnictwa będzie określony wyrażeniem:

$$I_Z = A \int_{W_c}^{W_v} P_{v \rightarrow c} D_v(W) f_v(W) D_c(W) [1 - f_c(W)] dW$$



Rys. 4-8. Prąd Esakiego



Rys. 4-9. Prąd Zenera

Występująca tu stała A określa powierzchnię złącza, prędkość elektronu oraz jego ładunek.

Ponieważ można przyjąć, że:

$P_{c \rightarrow v} = P_{v \rightarrow c} = P = \text{const}$, to wzór na prąd wypadkowy przyjmie postać:

$$I = I_E - I_Z = AP \int_{W_c}^{W_v} D_c(W) D_v(W) [f_c(W) - f_v(W)] dW \quad (4.8)$$

W powyższych wzorach:

$$D_c(W) \sim \sqrt{W - W_c} \quad (4.9)$$

$$D_v(W) \sim \sqrt{W_v - W} \quad (4.10)$$

$$f_c(W) = \frac{1}{e^{\frac{W - F_N}{kT}} + 1} \quad (4.11)$$

$$f_v(W) = \frac{1}{e^{\frac{W - E_F}{kT}} + 1} \quad (4.12)$$

analogicznie $1 - f_c(W) = \frac{1}{e^{\frac{F_N - W}{kT}} + 1} \quad (4.13)$

$$1 - f_v(W) = \frac{1}{e^{\frac{F_P - W}{kT}} + 1} \quad (4.14)$$

W stanie równowagi $F_N = F_P$, tak że:

$$f_c(W) = f_v(W)$$

i tym samym prąd wypadkowy jest równy zeru.

Podobnie dla $W_v = W_c$ całka we wzorze (4.3) przyjmuje wartość zerową i prąd $I = 0$. Z rysunku (4.1) widać, że:

$$W_v = \Delta F_N + \Delta F_P - qU \quad (4.15)$$

Warunek $W_v = W_c$ wymaga, by:

$$U = \frac{\Delta F_N + \Delta F_P}{q} \quad (4.16)$$

Podstawiając wyrażenia (4.9) ÷ (4.14) do wzoru (4.8) otrzymujemy:

$$I = B(W_v - W_c)^2 (e^{\frac{qU}{kT}} - 1) \int_0^1 \frac{\sqrt{y(1-y)} dy}{me^{\alpha y} + ne^{\alpha(1-y)} + \beta} \quad (4.17)$$

gdzie:

$$m = e^{-\frac{\Delta F_N}{kT}}$$

$$n = e^{-\frac{\Delta F_P}{kT}}$$

$$\alpha = \frac{W_v - W_c}{kT}$$

$$\beta = e^{\frac{qU}{kT}} + 1$$

Korzystając z podstawienia trygonometrycznego $y = \cos^2 \varphi$ całkę (4.17) redukuje się do postaci:

$$I = B(W_v - W_c)^2 (e^{\frac{qU}{kT}} - 1) \int_0^{\frac{\pi}{2}} \frac{\sin^2 \varphi d\varphi}{me^{\frac{\alpha}{2} \cos \varphi} + ne^{-\frac{\alpha}{2} \cos \varphi} + e^{-\frac{\alpha}{2} \beta}}$$

Przy założeniu $\Delta F_N, \Delta F_P \approx 2kT$ otrzymujemy następujące przybliżenie:

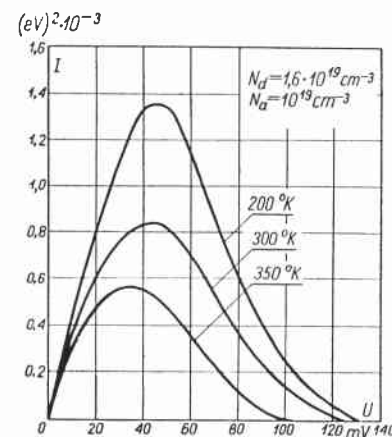
$$I = I_E - I_Z = C \frac{(W_v - W_c)^2 (e^{\frac{qU}{kT}} - 1)}{\left(e^{\frac{qU}{kT}} + 1\right) + \left(e^{-\frac{\Delta F_N}{kT}} + e^{-\frac{\Delta F_P}{kT}}\right) e^{\frac{W_v - W_c}{2kT}}} \quad (4.18)$$

w którym C — stała.

Przyjmując $W_c = 0$ i podstawiając wzór (4.10) otrzymujemy ostatecznie:

$$I = C \frac{(\Delta F_N + \Delta F_P - qU)^2 (e^{\frac{qU}{kT}} - 1)}{\left(e^{\frac{qU}{kT}} + 1\right) + e^{-\frac{\Delta F_P - \Delta F_N}{kT}} + e^{-\frac{\Delta F_N - \Delta F_P}{kT}}} \quad (4.19)$$

Znając ΔF_N i ΔF_P po obu stronach złącza zdegenerowanego, można wykreślić zależność $I = f(U)$. Rys. 4-10 obrazuje zależność (4.19) dla trzech temperatur: 200 °K, 300 °K, 350 °K. Krzywa ciągła poprowadzona jest



Rys. 4-10. Wykres zależności (4.19) dla różnych temperatur

dla $N_d = 10^{19} \text{ cm}^{-3}$ oraz $N_a = 1,6 \cdot 10^{19} \text{ cm}^{-3}$.

Napięcie szczytowe U_p można wyznaczyć z następującego wzoru:

$$U_p = \frac{1}{6q} [2\Delta F_P - \Delta F_N + \sqrt{(2\Delta F_P + \Delta F_N)^2 + 4\Delta F_N \Delta F_P}] \quad (4.20)$$

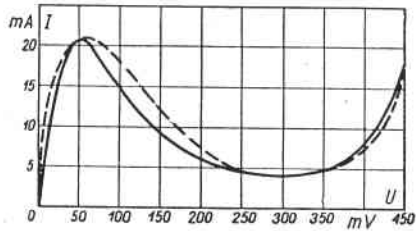
wyprowadzonego przy założeniu, że obszar P^+ jest silniej domieszkowany niż obszar N^+ . Na przykład dla $\Delta F_P = 3kT$ i $\Delta F_N = 2kT$ otrzymujemy $U_p \cong 55 \text{ mV}$.

Pełną charakterystykę prądowo-napięciową diody tunelowej można przedstawić za pomocą następującego wzoru, zawierającego sześć odpowiednio dobranych stałych A_j :

$$I = A_1 U + A_2 U^2 + A_3 U e^{-\frac{2}{3} \left(\frac{U}{A_4} \right)^2} + A_5 (e^{A_6 U} - 1) \quad (4.21)$$

Innym przybliżeniem analitycznym jest:

$$I = (A_1 U + A_2 U^3 + A_3 U^5 + A_4 U^7) e^{-\frac{U}{U_p}} + A_5 (e^{A_6 U} - 1) \quad (4.22)$$



Rys. 4-11. Charakterystyka diody tunelowej w przybliżeniu analitycznym (4.21)

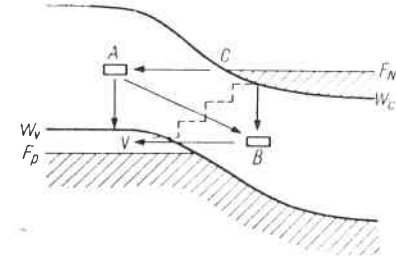
Rys. 4-11 przedstawia charakterystykę diody oraz zależność $I - U$ (krzywa przerywana) wg wzoru (4.22) dla następujących wartości stałych A_j :

$$\begin{aligned} A_1 &= 950,2 & A_4 &= 2,907 \cdot 10^6 \\ A_2 &= 702,6 & A_5 &= 3,442 \cdot 10^{-7} \\ A_3 &= 1133 \cdot 10^5 & A_6 &= 39 \end{aligned}$$

4.4. Prąd nadmiarowy

Przy $W_v < W_c$ pasmo podstawowe i pasmo przewodnictwa w złączu zdegenerowanym nie pokrywają się i wobec tego prąd tunelowy powinien zmaleć do zera. W rzeczywistości jednak płynie pewien prąd warunkujący skończoną wartość stosunku $\frac{I_p}{I_v}$ (rys. 4-3). Prąd ten może być uzależniony od strat energii elektronów tunelujących z pasma przewodnictwa wskutek wzajemnego oddziaływania tych elektronów i fononów lub nośników swobodnych. Najbardziej jednak prawdopodobne wydaje się tu tunelowanie do poziomów energetycznych w pasmie zabronionym. Tego rodzaju poziomy pochodzą od głębokich domieszek, względnie dyslokacji. Wskutek bardzo silnego domieszkowania w półprzewodnikach zdegenerowanych mogą istnieć pasma domieszkowe, których szerokość powinna być uzależniona od koncentracji domieszek.

Rys. 4-12 przedstawia możliwy mechanizm powstawania prądu nadmiarowego. Elektron w pasmie przewodnictwa C może tunelować do poziomu lokalnego A w pasmie zabronionym, z którego następnie spada do poziomu podstawowego V . Można również sądzić, że elektron spada do poziomu B i następnie tuneluje do pasma V . Odmienne tunelowanie odby-



Rys. 4-12. Model prądu nadmiarowego

wałoby się na drodze $CABV$, na której elektron traci energię przy przejściu z A do B . Można sobie wyobrazić również schodkowe przejście z C do V (zaznaczone na rysunku linią przerywaną), składające się z szeregu poziomych przejść tunelowych połączonych z przejściami pionowymi, przy których elektron traci energię. Najbardziej prawdopodobne wydaje się przejście na drodze CBV .

Przyjmując istnienie pasm domieszkowych o szerokości ΔW_a , ΔW_d odpowiednio dla domieszek akceptorowych i donorowych, można wyznaczyć napięcie doliny U_v :

$$U_v = \frac{1}{q} (\Delta W_a + \Delta W_d + \Delta F_N + \Delta F_P) \quad (4.23)$$

Doświadczenie potwierdza, że położenie minimum prądu doliny zależy od poziomu domieszkowania, innymi słowy od szerokości pasm domieszkowych (ΔW_a , ΔW_d).

4.5. Źródła szumu w diodach tunelowych

W diodzie tunelowej prąd wypadkowy jest różnicą dwóch oddzielnych i niezależnych prądów:

$$I = I_E - I_Z$$

Ponieważ prądy I_E , I_Z nie są skorelowane, to każdy z nich przyczynia się do fluktuacji emisji (efekt śrutowy). Zgodnie ze wzorem Schottky'ego mamy:

$$\overline{i^2} = 2q (|I_E| + |I_Z|) B \quad (4.24)$$

gdzie B — szerokość pasma, w którym rozważamy problem szumu.

W stanie równowagi $I = 0$; dioda tunelowa jako element oporowy wykazuje szum Johnsona, dla którego:

$$\overline{u^2} = 4kTRB \quad (4.25)$$

Korzystając z zależności:

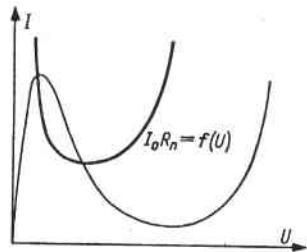
$$\frac{\overline{u^2}}{R} = \frac{\overline{i^2}}{G}$$

w której

$$G = \frac{1}{R} = \left(\frac{dI}{dU} \right)_{U=0}$$

otrzymujemy:

$$\overline{i^2} = 4kTGB \quad (4.26)$$



Rys. 4-13. Zależność $I_0 R_n$ od napięcia

Można wykazać, że dla $U = 0$ wzór (4.24) przechodzi w (4.26). W rozdziale 10 wykazemy, że współczynnik szumu diody tunelowej jest proporcjonalny do iloczynu $I_0 R_n$ (I_0 — prąd polaryzujący diodę). Zależność $I_0 R_n$ od napięcia pokazana jest na rys. 4-13.

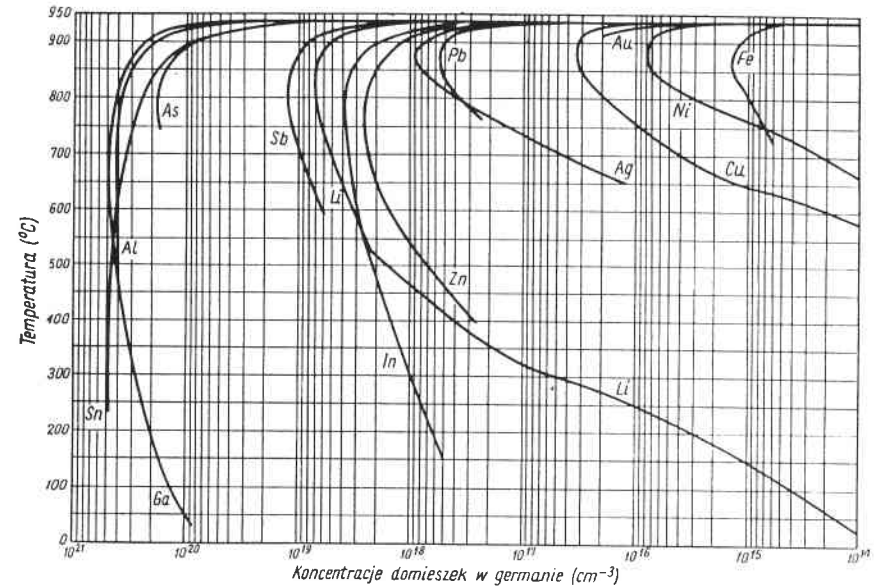
5. ELEMENTY TECHNOLOGII DIOD TUNELOWYCH

Zagadnienia technologiczne wytwarzania diod tunelowych sprowadzają się do uzyskania:

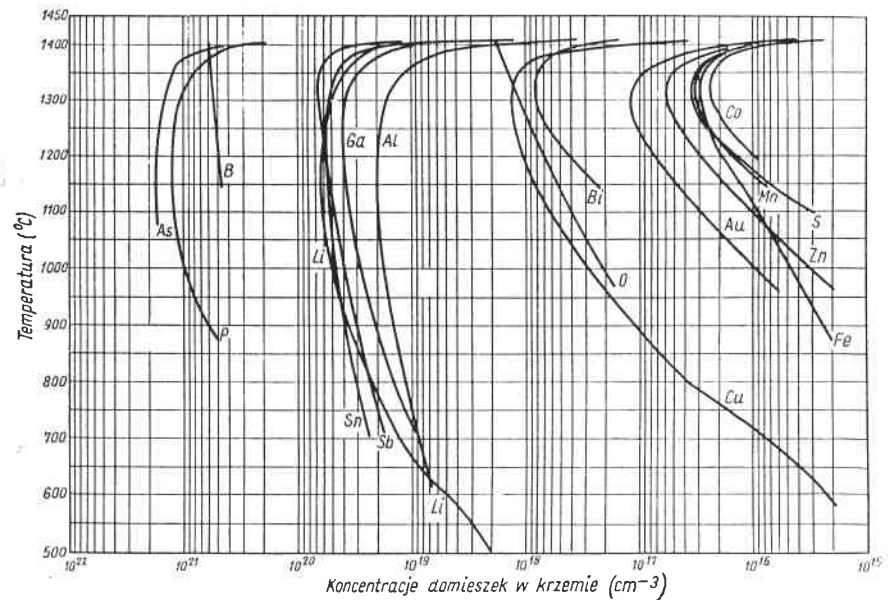
- monokryształów półprzewodników zdegenerowanych,
- złącz wykazujących ujemną oporność.

5.1. Wytwarzanie półprzewodników zdegenerowanych

Jak już wiadomo, dla uzyskania półprzewodników zdegenerowanych koncentracja domieszek musi być rzędu $10^{19} \div 10^{20} \text{ cm}^{-3}$. Biorąc pod uwagę, że kryształ germanu zawiera $4,5 \cdot 10^{22}$ atomów w 1 cm^3 trzeba stwierdzić, że są to koncentracje bardzo duże, sięgające $0,2 \div 2\%$. Tak dużych koncentracji nie można osiągnąć dla wszystkich domieszek.

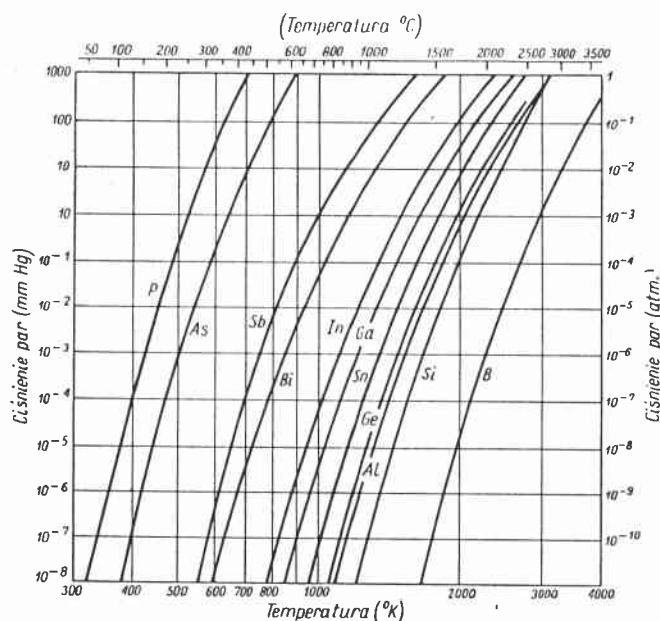


Rys. 5-1. Rozpuszczalność domieszek w germanie w funkcji temperatury



Rys. 5-2. Rozpuszczalność domieszek w krzemie w funkcji temperatury

W szczególności, w przypadku najczęściej stosowanych w germanie domieszek antymonu (Sb) i indu (In), maksymalne ich rozpuszczalności są zbyt małe dla uzyskania odpowiednio silnie domieszkowanego materiału. Na rys. 5-1 i rys. 5-2 uwidocznione są rozpuszczalności domieszek w germanie i krzemie w funkcji temperatury. Z wykresów tych widzimy, że w przypadku krzemu i germanu mogą być stosowane jedynie fosfor (P), arsen (As) oraz gal (Ga) i bor (B). Rozpuszczalność aluminium (Al) jest bardzo duża w germanie, lecz nie zbyt duża w krzemie.



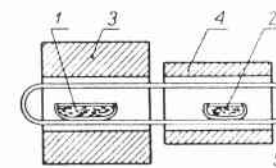
Rys. 5-3. Wykres ciśnienia par niektórych pierwiastków chemicznych w funkcji temperatury

Pozornie może się wydawać, że dla uzyskania zdegenerowanego germanu typu N wystarczy dodać do ciekłego germanu ($\sim 950^\circ\text{C}$) odpowiednią ilość fosforu lub arsenu. W rzeczywistości jednak zarówno fosfor jak i arsen w temperaturze tej wykazują olbrzymie ciśnienia par, co uniemożliwia dowolne domieszkowanie w procesie wyciągania monokryształu. Uzyskanie bardzo silnie domieszkowanego materiału typu P nie przedstawia trudności, gdyż — jak widać z wykresu na rys. 5-3 — ciśnienie par galu w temperaturze topnienia germanu jest bardzo małe.

Trudność uzyskania germanu zdegenerowanego typu N pokonano w procesie dyfuzji par fosforu (lub arsenu) do ciekłego materiału wyjściowego. W metodzie tej uzyskuje się tzw. stop pośredni, w którym temperatura fosforu nie przekracza 350°C i tym samym ciśnienie nie prze-

kracza 1 atm. Schemat służącej do tego celu aparatury przedstawiony jest na rys. 5-4. W odpompowanej i zatopionej rurze kwarcowej znajdują się dwie łądki, zawierające odpowiednio german i fosfor. Rura umieszczona jest wewnątrz dwóch pieców oporowych. Większy piec służy do stopienia germanu, mniejszy do podgrzewania fosforu. Temperatury tych pieców utrzymywane są na poziomie 960° i 350°C . Dyfuzja par fosforu odbywa się w ciągu kilku godzin, przy czym ciśnienie par nie przekracza 1 atm. Otrzymany stop pośredni jest materiałem polikrystalicznym,

Rys. 5-4. Aparatura do wytwarzania germanu zdegenerowanego
1 — german, 2 — łądka z fosforem,
3 — piec do topienia germanu,
4 — piec mniejszy, 5 — rura kwarcowa



z którego następnie zostaje wyciągnięty monokryształ w aparaturze typu Czochralskiego. Pewną odmianę klasycznej metody wytwarzania monokryształów germanu przedstawili F. A. Trumbore i E. M. Porbansky [B3]. W metodzie tej obracana (z szybkością ok. 150 obr/min) zarodek monokryształu germanu zostaje doprowadzona do zetknięcia z ciekłym germanem, lecz nie jest wyciągana. Podczas wyparowywania arsenu ze stopu ciekłego o zawartości $5 \div 10\%$ As na kryształ-rodzi wytrąca się lub rośnie german silnie domieszkowany arsenem. W wyniku tego procesu przebiegającego w czasie ok. 20 godzin, powstaje jednorodny monokryształ o ciężarze ok. $30 \div 35$ g.

Pierwsze monokryształy zdegenerowanego germanu, uzyskane w laboratoriach Sony Corporation, domieszkowano fosforem indu (InP). Ze względu na znacznie większą rozpuszczalność fosforu kryształ był typu N.

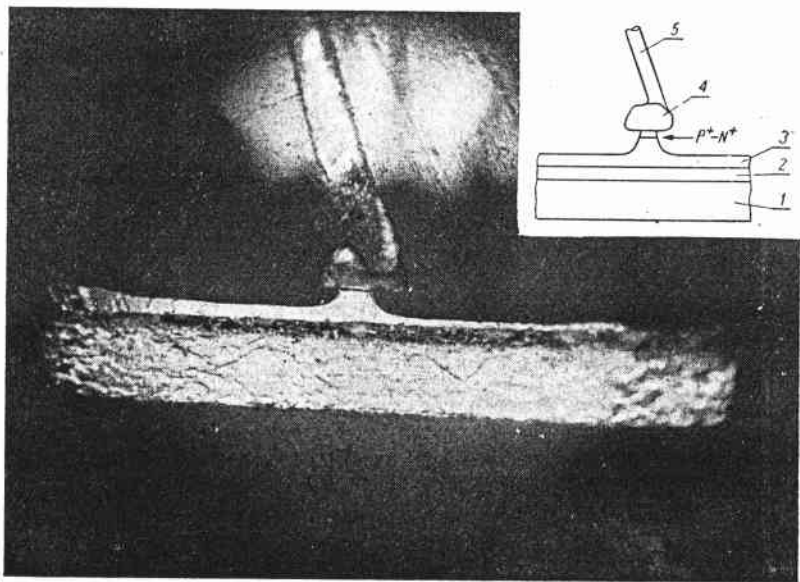
Wytwarzanie monokryształów arsenku galu przedstawia również pewną trudność ze względu na duże ciśnienie par arsenu w temperaturze topnienia GaAs (1238°C). Domieszkami donorowymi w GaAs są selen (Se), tellur (Te), zaś akceptorowymi cynk (Zn) lub kadm (Cd).

Dla uzyskania materiału półprzewodnikowego do eksperymentalnej produkcji diod tunelowych może być wykorzystany materiał o niewystarczająco małej oporności właściwej, do którego wdyfundowano odpowiednie domieszki. W ten sposób można np. z GaAs o oporności właściwej $0,06 \Omega \text{ cm}$ ($N = 5 \cdot 10^{16} \text{ cm}^{-3}$) uzyskać materiał zdegenerowany.

Jako kryterium koncentracji domieszki służy pomiar oporności właściwej łącznie z pomiarem stałej Halla. Ze względu na duże zawartości domieszek, ich liczbę i przede wszystkim identyfikację chemiczną można wyznaczyć za pomocą znanych metod chemicznych, jak np. spektroskopia optyczna, kolorymetria, neutronowa analiza aktywacyjna itp.

5.2. Wytwarzanie diod tunelowych

Diody tunelowe otrzymuje się prawie wyłącznie techniką stopową, która — jak to zaznaczono w rozdziale 3 — pozwala uzyskać złącza o bardzo stromym rozkładzie domieszek. Złącze w półprzewodniku typu *N* uzyskuje się przez wtopienie kulki indowej zawierającej domieszkę galu (0,5 ÷ 0,8 ‰), lub domieszkę galu i cynku. Własności elektryczne złą-



Rys. 5-5. Przekrój germanowego złącza tunelowego
1 — baza niklowa; 2 — warstwa lutu; 3 — german; 4 — stop ind-gal; 5 — drucik srebrny

cza określa praktycznie wyłącznie gal. W przypadku germanu typu *P* wtopia się kulkę indową zawierającą fosfor (0,5 ÷ 1 ‰). Często diody tunelowe wytwarza się nie przez wtopienie kulek, lecz przez zgrzanie lub uformowanie impulsami elektrycznymi styku drucika aluminiowego lub srebrnego (zawierającego odpowiednią domieszkę) z półprzewodnikiem.

W Zakładzie Elektroniki I.P.P.T. PAN pierwsze partie złącz tunelowych uzyskano w następujący sposób [B5]. Płytki germanowe monokrystaliczne¹⁾ o grubości ok. 100 μ po oszlifowaniu i trawieniu w roztworze CP-4 umieszczano łącznie z kulkami indo-galowymi (ϕ 130 μ) w grafitowym szablonie i następnie zgrzewano w piecu wodorowym w temperaturze ok. 500 °C. W procesie tym istotne było szybkie narastanie temperatury oraz szybkie studzenie. Bazę niklową przylutowywa-

¹⁾ Zadawalające wyniki uzyskiwano również przy użyciu germanu polikrystalicznego.

no do płytki półprzewodnikowej za pomocą niskotopliwego lutu w piecu wodorowym w temperaturze ok. 350 °C. Drugą elektrodę (anodę) stanowił cienki drucik srebrny przylutowany do kulki. Po zmontowaniu diody następowało trawienie elektrochemiczne impulsami prądowymi w rozcieńczonym (5 ÷ 10 ‰) roztworze wodorotlenku potasu. Podczas trawienia malała powierzchnia złącza, a tym samym i prąd szczytowy I_p (aż do zadanej wartości, kontrolowanej za pomocą charakterografu). Trawienie usuwało powierzchniowe zanieczyszczenia i przyczyniało się do zwiększenia stosunku $\frac{I_p}{I_v}$, zmniejszenia napięcia U_p i zwiększenia napięcia U_v . Po płukaniu w wodzie zdejonizowanej, osuszeniu i zamaskowaniu lakierem, diodę zamykano hermetycznie. Rys. 5-5 przedstawia zdjęcie przekroju germanowej diody tunelowej z widoczną wyraźnie granicą przejścia *P-N*. Średnica złącza wynosi 35 μ.

Złącza krzemowe można wytwarzać przez wtopienie drutu aluminiowego z domieszką boru do krzemu typu *N*.

Złącza arseno-galowe wykonuje się przez wtopienie kulek cynkowych lub kadmowych do GaAs typu *N*.

Eksperymentalne złącza w antymonku galu wykonywano przez elektryczne formowanie metalicznego styku ostrzowego z domieszką selenu lub telluru do GaSb typu *P*.

Złącza tunelowe w antymonku indu wykonywano przez wtopienie kulek o zawartości 99,9 ‰ indu oraz 0,1 ‰ kadmu do kryształu o stosunkowo małej zawartości domieszek (rzędu 1,5 ÷ 6 · 10¹⁷ cm⁻³).

Poszczególni producenci (jak i odbiorcy) ustalają określone wymagania próbne dla diod tunelowych. Diody te muszą wytrzymywać odpowiednie cykle temperaturowe, przejść przez określone próby wytrzymałościowe na zginanie, skręcanie, wstrząsy itp., oraz muszą wykazać odporność na przeciążenia elektryczne.

6. UKŁAD ZASTĘPCZY DIODY TUNELOWEJ I ZAGADNIENIE STABILNOŚCI

6.1. Układ zastępczy

Układ zastępczy diody tunelowej dla małych sygnałów przedstawia rys. 4.4. Składa się on z czterech elementów: oporności ujemnej diody $-R_n$, pojemności złącza *C*, oporu szeregowego r_s oraz indukcyjności *L* doprowadzeń do złącza. W przypadku diod o bardzo małej pojemności złącza należałoby dodać jeszcze pojemność oprawki. Zazwyczaj jednak pojemność oprawki jest znikomo mała (rzędu ułamka pikofarada) i dlatego się ją pomija. Nie bierze się również pod uwagę pojemności dyfuzyjnej diody, gdyż przy stosunkowo niewielkich napięciach polaryzujących diodę pojemność dyfuzyjna nie odgrywa większej roli. Jest rzeczą

oczywista, że zamiast operować opornością ujemną, możemy mówić o przewodności ujemnej — G_n zgodnie z zależnością:

$$-G_n = -\frac{1}{R_n} \quad (6.1)$$

6.2. Impedancja diody

Impedancję $Z(\omega) = R + jX$ diody tunelowej wyraża zgodnie z układem zastępczym z rys. 4-4 następujący wzór:

$$Z(\omega) = r_s - \frac{R_n}{\omega^2 C^2 R_n^2 + 1} + j \left(\omega L - \frac{\omega C R_n^2}{\omega^2 C^2 R_n^2 + 1} \right) \quad (6.2)$$

Część rzeczywista $\text{Re}Z(\omega)$ jest równa:

$$R = r_s - \frac{R_n}{\omega^2 C^2 R_n^2 + 1} \quad (6.3)$$

Część urojona $\text{Im} Z(\omega)$ jest równa:

$$X = \omega L - \frac{\omega C R_n^2}{\omega^2 C^2 R_n^2 + 1} \quad (6.4)$$

Dla prądu stałego ($\omega \rightarrow 0$) otrzymujemy:

$$Z = r_s - R_n \quad (6.5)$$

Częstotliwość, dla której znika część rzeczywista, nazywa się częstotliwością odcięcia opornościowego. Z warunku $\text{Re} Z(\omega) \equiv R = 0$ znajdujemy:

$$\omega_R = \frac{1}{R_n C} \sqrt{\frac{R_n}{r_s} - 1} \quad (6.6)$$

lub inaczej

$$f_R = \frac{1}{2\pi R_n C} \sqrt{\frac{R_n}{r_s} - 1}$$

Najczęściej oporność R_n jest znacznie większa od oporności szeregowej r_s diody. Jako typowe wartości można podać $R_n = 100 \Omega$ i $r_s = 1 \Omega$. Wzór (6.6) można więc przepisać w postaci:

$$f_R \approx \frac{1}{2\pi C \sqrt{R_n r_s}} \quad (6.7)$$

Podobnie definiujemy częstotliwość samorezonansową ω_X jako częstotliwość, dla której znika część urojona impedancji, czyli reaktancja. Z warunku $\text{Im} Z = X = 0$ znajdujemy:

$$\omega_X = \frac{1}{\sqrt{LC}} \sqrt{1 - \frac{L}{C R_n^2}} \quad (6.8)$$

lub inaczej

$$f_X = \frac{1}{2\pi \sqrt{LC}} \sqrt{1 - \frac{L}{C R_n^2}}$$

Dla $L \ll C R_n^2$ otrzymujemy wzór przybliżony:

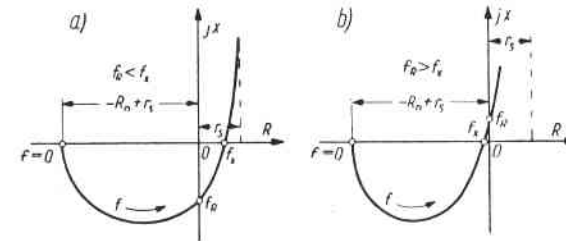
$$f_X \approx \frac{1}{2\pi \sqrt{LC}} \quad (6.9)$$

Poniżej częstotliwości samorezonansowej reaktancja diody jest pojemnościowa, zaś powyżej tej częstotliwości — indukcyjna. Należy również pamiętać, że obie częstotliwości f_R i f_X są funkcjami punktu pracy diody tunelowej.

Z łatwością można sprawdzić, że dla $r_s = \frac{L}{C R_n}$, $f_X = f_R$. Impedancję diody w funkcji częstotliwości można przedstawić schematycznie jak na rys. 6-1. Rozróżniamy tu dwa przypadki szczególne:

$$\omega_R < \omega_X \quad \text{oraz} \quad \omega_R > \omega_X$$

Na wykresie tym na osi odciętych odkładamy oporność rzeczywistą, zaś na osi rzędnych reaktancję. Częstotliwość jest parametrem rosnącym w kierunku strzałki.



Rys. 6-1. Zależność impedancji diody tunelowej od częstotliwości

W analizie obwodów z diodami tunelowymi wygodnie jest niekiedy posługiwać się częstotliwością zespoloną $s = \sigma + j\omega$. Wzór (6.2) przyjmie w tym przypadku postać:

$$Z(s) = \frac{s^2 L C R_n + s(r_s R_n C - L) + R_n - r_s}{s C R_n - 1} \quad (6.10)$$

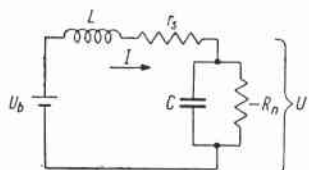
6.3. Zagadnienie stabilności

Zagadnienie stabilności diod tunelowych ma pierwszorzędne znaczenie zarówno przy projektowaniu układów jak i w miernictwie diod tunelowych. W celu zanalizowania warunków stabilności i ustalenia rela-

cji wiążących poszczególne parametry układu z diodą tunelową rozpatrzmy możliwie najprostszy obwód, w którym dioda tunelowa podłączona jest do bezstratnego i bezindukcyjnego źródła prądu stałego (rys. 6-2). Równania różniczkowe dla tego obwodu będą następujące:

$$U_b = L \frac{dI}{dt} + r_s I + U \quad (\text{napięcia}) \quad (6.11)$$

$$C \frac{dU}{dt} = I - f(U) \quad (\text{prądy}) \quad (6.12)$$



Rys. 6-2. Obwód zawierający diodę tunelową i źródło prądu stałego (I' — prąd płynący przez $-R_n$)

Punkty osobliwe ¹⁾ określamy z zależności:

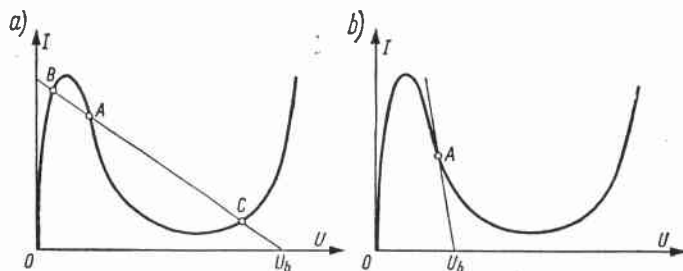
$$U_b = r_s I_0 + U_0 = r_s I'_0 + U_0 \quad (6.13)$$

$$I_0 = f(U_0) = I'_0 \quad (6.14)$$

w których: I_0, U_0 — wartości osobliwe.

W punkcie osobliwym $I_0 = I'_0$; prąd płynący do pojemności C jest równy zeru.

Linia obciążenia (rys. 6-3) przechodzi przez oś U w punkcie $U = U_b$. Nachylenie określone jest opornością r_s . Przecięcie linii obciążenia z krzywą



Rys. 6-3. Przecięcie linii obciążenia i charakterystyki statycznej diody tunelowej: a) w trzech punktach; b) w jednym punkcie

wą $I' = f(U)$ wyznacza właśnie punkty osobliwe. Istnieją tu dwie możliwości. Przy dużych wartościach U_b i r_s otrzymuje się trzy punkty przecięcia A, B, C — rys. 6-3a. W przypadku małych wartości U_b i r_s otrzymuje się jeden punkt przecięcia linii obciążenia z krzywą $I' = f(U)$ — rys. 6-3b.

¹⁾ Patrz Uzupełnienie II (str. 120).

W pobliżu osobliwości można przedstawić prądy i napięcia następująco:

$$\begin{aligned} I &= I_0 + i \\ U &= U_0 + u \end{aligned} \quad (6.15)$$

gdzie i oraz u oznaczają małe zmiany.

Równania (6.11) i (6.12) — uwzględniając (6.15) — przyjmą postać:

$$\left. \begin{aligned} \frac{di}{dt} &= -\frac{r_s}{L} i - \frac{1}{L} u \\ \frac{du}{dt} &= \frac{1}{C} i + \frac{1}{CR_n} u \end{aligned} \right\} \quad (6.16)$$

Porównując (6.16) ze wzorem (3) ¹⁾ z Uzupełnienia II widzimy, że:

$$a = \frac{1}{C} \quad b = \frac{1}{CR_n} \quad c = -\frac{r_s}{L} \quad d = -\frac{1}{L} \quad (6.17)$$

Pierwiastki charakterystyczne λ_1 oraz λ_2 będą wyrażone wzorem:

$$\lambda_{1,2} = \frac{1}{2} \left(\frac{1}{CR_n} - \frac{r_s}{L} \right) \pm j \sqrt{\frac{1}{LC} \left(1 - \frac{r_s}{R_n} \right) - \frac{1}{4} \left(\frac{1}{CR_n} - \frac{r_s}{L} \right)^2} \quad (6.18)$$

Pierwiastki charakterystyczne reprezentują zespolone częstotliwości $s = \sigma + j\omega$. Mogą one być rzeczywiste, zespolone lub urojone, zależnie od wartości parametrów r_s, R_n, C, L . Jeżeli jedna z dwóch częstotliwości λ_1, λ_2 ma dodatnią część rzeczywistą, obwód będzie niestabilny. Jeżeli pierwiastki są rzeczywiste, to początkowe zaburzenie będzie rosło wykładniczo lub malało do stałej wartości. Pierwiastki zespolone warunkują rosnące lub malejące przebiegi sinusoidalne.

Ze wzoru (6.18) widać, że warunkiem stabilności będzie: $\frac{1}{CR_n} - \frac{r_s}{L} < 0$ stąd

$$r_s > \frac{L}{CR_n} \quad (6.19)$$

Równocześnie jednak $1 - \frac{r_s}{R_n} > 0$, lub inaczej:

$$r_s < R_n \quad (6.20)$$

Jest to warunek stabilności dla prądu stałego. Zwróćmy uwagę, że warunek (6.19) oznacza, iż $f_R < f_X$. Warunek stabilności można więc wyrazić w postaci:

$$\frac{L}{CR_n} < r_s < R_n \quad (6.21)$$

¹⁾ W drugim równaniu (6.16) korzystamy z rozwinięcia

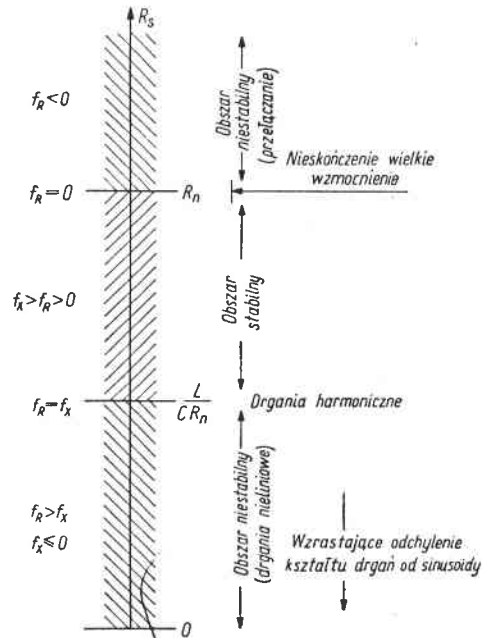
$$f(U) = f(U_0) + \frac{df(U)}{dU} u \quad \text{oraz definicji} \quad \frac{df(U)}{dU} = -\frac{1}{R_n}$$

W konkretnym układzie musimy jednak brać pod uwagę dodatkowe oporności i indukcyjności. Jeżeli przez R_s oznaczymy całkowitą oporność szeregową, a przez L_s całkowitą indukcyjność, to warunek (6.21) przyjmie postać:

$$\frac{L_s}{CR_n} < R_s < R_n \quad (6.22)$$

Ponieważ $R_n > \frac{L_s}{CR_n}$, to $R_n^2 > \frac{L_s}{C}$ i tym samym

$$R_n > \sqrt{\frac{L_s}{C}} \quad (6.23)$$



Rys. 6-4. Zakresy pracy diody tunelowej

Powyższe warunki można przedstawić poglądowo w sposób podany na rys. 6-4. Wyrażenie (6.18) zawiera cztery parametry układu zastępczego diody tunelowej. Przez podstawienie:

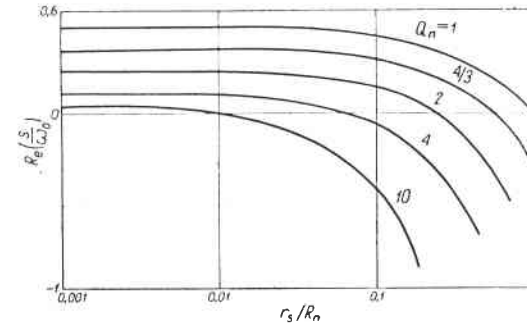
$$\omega_0 = \frac{1}{\sqrt{LC}} \quad (6.24)$$

oraz

$$Q_n = \omega_0 R_n C = \frac{R_n}{\sqrt{\frac{L}{C}}} \quad (6.25)$$

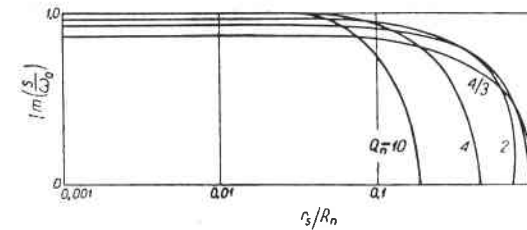
redukuje się liczbę parametrów do dwóch:

$$\frac{s_{1,2}}{\omega_0} = \frac{1}{2Q_n} \left(1 - \frac{r_s}{R_n} Q_n^2 \right) \pm j \sqrt{\left(1 - \frac{r_s}{R_n} \right) - \frac{1}{4Q_n^2} \left(1 - \frac{r_s}{R_n} Q_n^2 \right)^2} \quad (6.26)$$



Rys. 6-5. Zależności części rzeczywistej $\frac{s}{\omega_0}$ od stosunku $\frac{r_s}{R_n}$

Hines [G10] podał wykresy powyższej zależności, osobno dla części rzeczywistej i osobno dla części zespolonej w funkcji stosunku $\frac{r_s}{R_n}$ (rys. 6-5 i rys. 6-6). Parametrem jest Q_n . Z wykresu widać, że układ będzie nie-

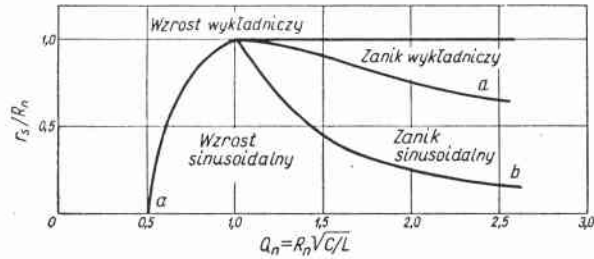


Rys. 6-6. Zależność części urojonej $\frac{s}{\omega_0}$ od stosunku $\frac{r_s}{R_n}$

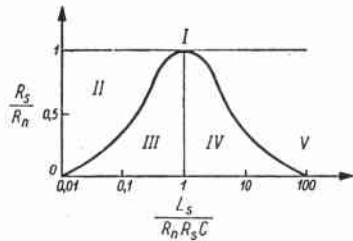
stabilny w przypadku, gdy stosunek $\frac{r_s}{R_n}$ będzie zbyt mały lub zbyt duży, co odpowiada dodatnim wartościom $\frac{s}{\omega_0}$, lub w przypadku zbyt dużych wartości $\frac{L}{C}$ (rys. 6-5). W. W. Anderson [G10] przedstawił interesujący wykres dotyczący zależności stosunku $\frac{r_s}{R_n}$ od parametru Q_n . Mamy tu (rys. 6-7) cztery obszary. Dla krzywej a) $-\frac{r_s}{R_n} = \frac{2}{Q_n} - \frac{1}{Q_n^2}$. Dla krzywej b) $-\frac{r_s}{R_n} = \frac{1}{Q_n^2}$. Nieco odmienny wykres przedstawił Cox [14]. Została tu

przedstawiona zależność stosunku $\frac{R_s}{R_n}$ od $\frac{L_s}{R_n R_s C}$ (rys. 6-8). W obszarze I całkowita oporność szeregową R_s przewyższa wartość oporności R_n i może mieć miejsce jedynie przełączanie. W obszarach II, III i IV mamy $R_s < R_n$. Jeżeli całkowita indukcyjność szeregową jest mała, tak że

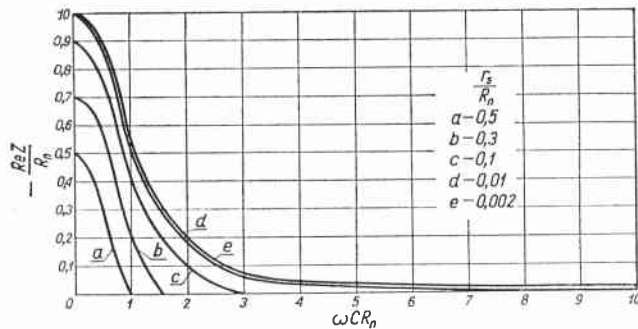
$$L_s < R_n R_s C$$



Rys. 6-7. Wykres Andersona



Rys. 6-8. Wykres Coxa



Rys. 6-9. Zależność $-\frac{\text{Re}Z}{R_n}$ od ωCR_n

to obwód będzie stabilny i oscylacje nie wystąpią. Można odróżnić tu dwa przypadki:

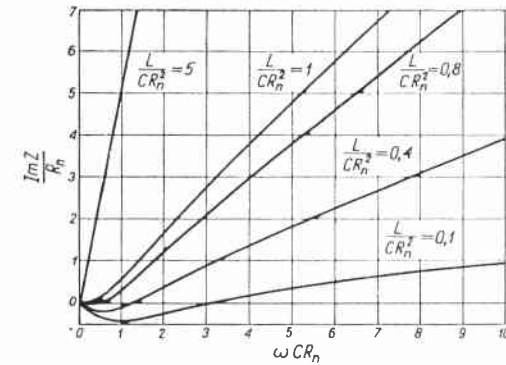
$$R_n < \frac{L_s}{2\sqrt{L_s C} - R_s C} \quad (\text{pojedynczy impuls wykładniczy})$$

$$R_n > R_n R_s C \quad (\text{możliwość wzmacniania})$$

Jeżeli $L_s > R_n R_s C$, to podobnie można odróżnić dwa przypadki:

$$R_n > \frac{L_s}{2\sqrt{L_s C} - R_s C} \quad (\text{drgania sinusoidalne})$$

$$R_n < \frac{L_s}{2\sqrt{L_s C} - R_s C} \quad (\text{drgania relaksacyjne})$$



Rys. 6-10. Zależność $\frac{\text{Im}Z}{R_n}$ od ωCR_n

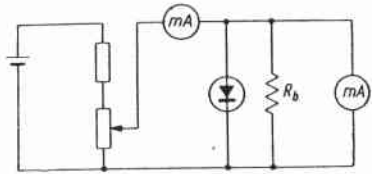
Niekiedy, zwłaszcza przy projektowaniu obwodów generacyjnych wielkiej częstotliwości, pomocne mogą być inne wykresy. Jako przykład może służyć wykres zależności $-\frac{\text{Re}Z}{R_n}$ od ωCR_n (rys. 6-9) oraz $\frac{\text{Im}Z}{R_n}$ od ωCR_n (rys. 6-10). Parametrem pierwszego wykresu jest stosunek $\frac{r_s}{R_n}$, drugiego zaś — stosunek $\frac{L}{CR_n^2}$.

7. POMIARY DIOD TUNELOWYCH

7.1. Pomiary statyczne

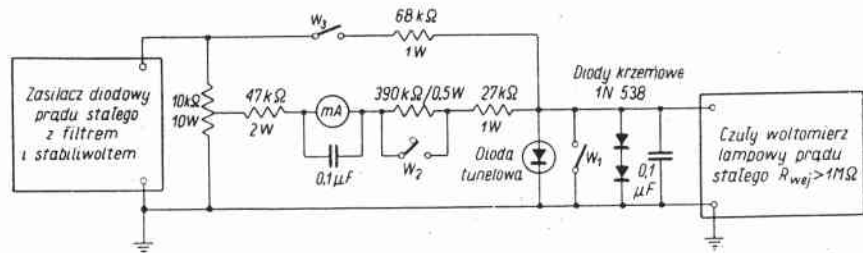
Najbardziej istotne informacje o własnościach diod tunelowych można uzyskać z charakterystyki statycznej przedstawiającej zależność natężenia prądu diody od przyłożonego napięcia. Z charakterystyki statycznej można odczytać wartości punktu szczytowego (I_p, U_p) oraz punktu doliny (I_v, U_v), wyznaczyć graficznie ujemną oporność $-R_n$, oraz zorientować się o własnościach szumowych, które zależą od iloczynu $I_0 R_n$. Również w badanych zjawiskach fizycznych zachodzących w diodach tunelowych najwięcej informacji daje charakterystyka statyczna $I - U$.

Charakterystykę statyczną diody tunelowej można w zasadzie zdjąć stosując możliwie najprostszy układ pomiarowy (rys. 7-1). W zakresie napięć od U_p do U_v dioda musi być zbocznikowana dodatnim bezindukcyjnym opornikiem R_b podłączonym możliwie jak najbliżej końcówek diody. Dla wykreślenia charakterystyki należy odjąć od mierzonej wartość



Rys. 7-1. Układ do pomiaru charakterystyki statycznej diody tunelowej

ci prądu prąd płynący przez bocznikujący opornik dodatni R_b . Ten sposób postępowania szczególnie ułatwia wykorzystanie przyrządu piszącego we współrzędnych $X - Y$, przy czym prąd płynący przez bocznikujący opornik może być automatycznie odejmowany.

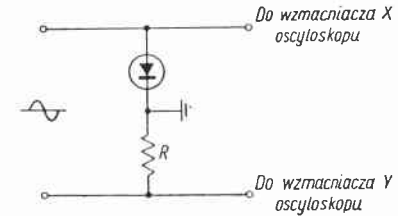


Rys. 7-2. Schemat miernika parametrów I_p , I_v , U_p , U_v

Szczególnie interesujące parametry diody tunelowej I_p , I_v , U_p , U_v można wyznaczyć możliwie dokładnie za pomocą specjalnego miernika (rys. 7-2). Zwiększając prąd płynący przez diodę aż do momentu przełączania można odczytać prąd szczytowy na miliamperomierzu. Wartość napięcia odczytuje się na miliwoltmierz lampowym o dużej oporności wewnętrznej. Wyłącznik W_2 umożliwia wprowadzenie do układu opornika szeregowego dla zredukowania prądu przy pomiarze I_v . Wyłącznik W_3 powoduje doprowadzenie do diody napięcia ustalającego punkt charakterystyki w zakresie wyższych napięć. Następnie prąd diody zostaje stopniowo zredukowany, aż napięcie „przeskoczy” do rosnącej charakterystyki, odpowiadającej niższym napięciom. Równocześnie odczytuje się wartość napięcia punktu doliny.

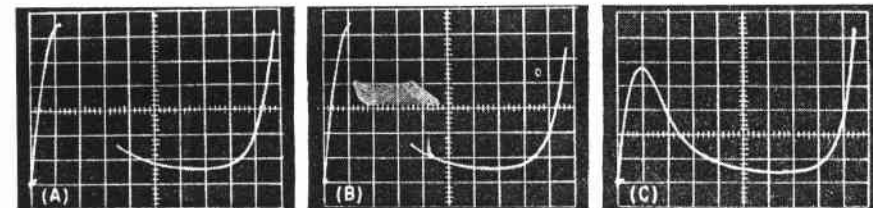
7.2. Oscyloskopowe metody pomiaru

Najwygodniejszy i najczęściej spotykany sposób zdjęcia charakterystyki $I - U$ polega na wykorzystaniu kalibrowanego oscyloskopu z odpowiednio zaprojektowaną przystawką. Schemat najprostszej przystawki



Rys. 7-3. Schemat ideowy przystawki oscyloskopowej

oscyloskopowej przedstawiony jest na rys. 7-3. Napięcie na diodzie zostaje podane do wzmacniacza X, zaś napięcie proporcjonalne do prądu płynącego przez diodę — do wzmacniacza Y. Pożądane jest, aby dla uniknięcia różnic fazowych wzmacniacze X i Y były identyczne, mimo że wymagana jest znacznie mniejsza czułość wzmacniacza X.



Rys. 7-4. Charakterystyki napięciowo-prądowe diod tunelowych: a) przełączanie; b) oscylacje; c) pełna charakterystyka

Interesuje nas zazwyczaj tylko część charakterystyki odpowiadającej doprowadzonym napięciom dodatnim. Dzięki jednokierunkowemu wyprostowaniu zmiennych napięć zasilających przystawkę, punkt początkowy zaznaczony jest na ekranie lampy oscyloskopowej silnie naświetloną plamką. Otrzymany obraz najczęściej jednak bądź nie przedstawia części charakterystyki odpowiadającej ujemnej oporności, bądź też odtwarza charakterystykę zniekształconą przez oscylacje lub działanie przełączające diody (np. rys. 7-4 a, b, c). Uzyskanie prawidłowego wykresu $I - U$ nie jest zatem łatwe.

Ciekawy układ mostkowy przystawki do oscyloskopu opracował A. M. Goodman [F1]. Schemat ideowy tego mostka uwidoczniony jest na rys. 7-5. Dioda tunelowa jest tu zbocznikowana opornikiem o bardzo małej

oporności $R_b < R_{n \min}$. Z łatwością można obliczyć zależność pomiędzy prądem diody I_d a napięciem U_y doprowadzanym na płytki pionowe. Prąd stały płynący przez diodę jest równy:

$$I_d = I_1 - I_2$$

lub inaczej

$$I_d = \frac{\left(U_1 - \frac{R_1}{R_b}\right) U_x}{R_1}$$

Ponieważ $I_2 R_b = U_x$ oraz $I_1 R_1 = U_1$

to

$$U_y = U_x - \left(\frac{R_3}{R_2 + R_3}\right) (U_1 + U_x)$$

$$U_y = - \frac{\left[U_1 - \left(\frac{R_2}{R_3}\right) U_x\right]}{\left(\frac{R_2}{R_3} + 1\right)}$$

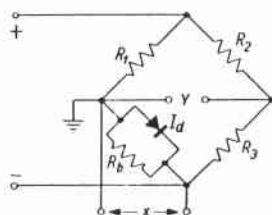
Z warunku równowagi mostka przy odłączonej diodzie tunelowej:

$$\frac{R_1}{R_b} = \frac{R_2}{R_3}$$

otrzymujemy:

$$I_d = - \left(\frac{R_2 + R_3}{R_1 R_3}\right) U_y$$

Pełny schemat przystawki mostkowej przedstawiony jest na rys. 7-6. Przy odłączonej diodzie można uzyskać równowagę mostka za pomocą po-

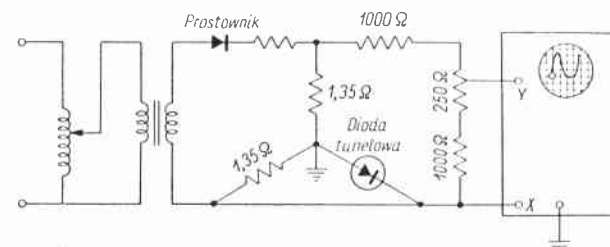


Rys. 7-5. Schemat ideowy mostka Goodmana (I_1 , I_2 prądy płynące odpowiednio przez R_1 , R_b)

tencjometru 250Ω , dzięki czemu na oscyloskopie otrzymuje się linię poziomą. Wspomniany układ mostkowy umożliwi obserwowanie całej charakterystyki diod tunelowych o dostatecznie małych wartościach R_n , a tym samym o dużym prądzie szczytowym.

Podobna przystawka została opisana również przez J. A. Naruda i T. A. Fyfe'a [F6]. Dioda tunelowa (rys. 7-7) jest tu zbocznikowana

opornikiem $R_b = 3,3 \Omega$ podłączonym możliwie jak najbliżej jej końcówek. Napięcie na oporniku $R_s = 5 \Omega$ jest proporcjonalne do całkowitego prądu, a prąd płynący przez opornik R_b , — kompensowany przez doprowa-

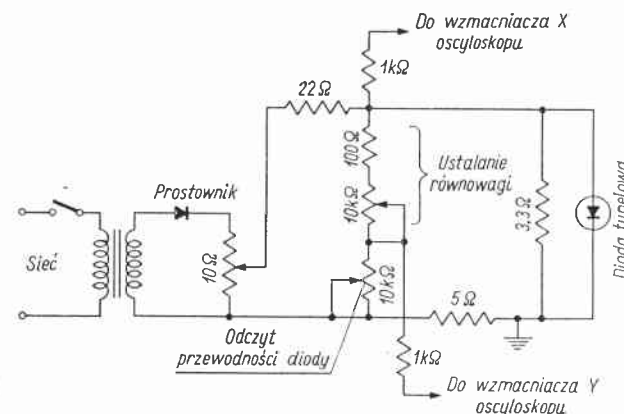


Rys. 7-6. Pełny schemat przystawki mostkowej

żeniem części napięcia na diodzie do wzmacniacza pionowego Y. Dokładną kompensację uzyskuje się przez odłączenie diody z układu i wyregulowanie wartości opornika R_1 ($10 \text{ k}\Omega$). W tych warunkach spełniona jest zależność:

$$\frac{R_2}{R_1} = \frac{R_s}{R_b}$$

i na ekranie występuje linia pozioma. Z pomocą tego układu można określić również wartość ujemnej przewodności diody tunelowej.

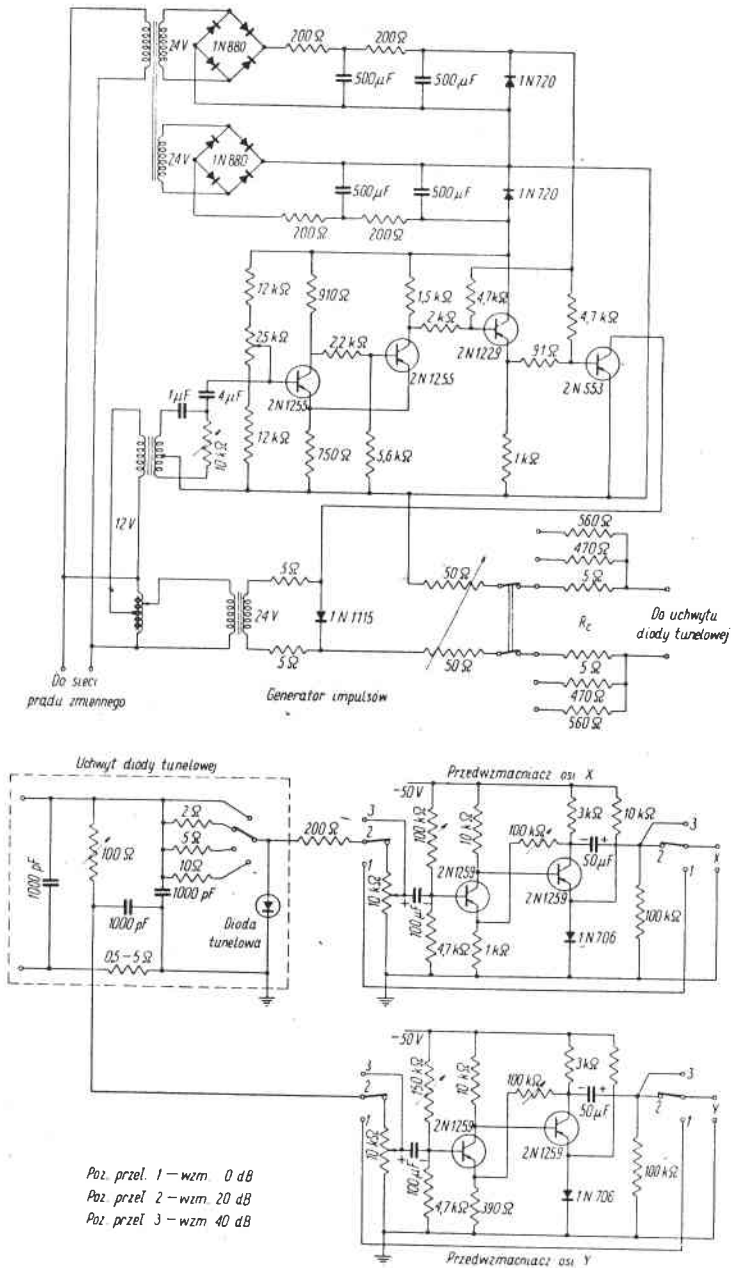


Rys. 7-7. Schemat przystawki oscyloskopowej Naruda i Fyfe'a

Dobierając wartość R_2 ($10 \text{ k}\Omega$) tak, aby nachylenie charakterystyki było równe zero (linia pozioma w ujemnej części), oblicza się przewodność ujemną G_n ze wzoru:

$$G_n = \frac{R_2}{R_1 R_s} - \frac{1}{R_b}$$

Tak więc pomiędzy G_n i R_2 istnieje zależność liniowa.



Rys. 7-8. Pełny schemat przystawki Dilla i Macphersona

Opornik stabilizujący R_b powinien spełniać następujący warunek (rozdz. 6);

$$\frac{L}{CR_n} < r_s + R_b < R_n$$

Opornik o wartości 22 Ω „izoluje” diodę tunelową od układu zasilającego. Maksymalna oporność opornika stabilizującego R_b jest proporcjonalna do oporności R_n diody, a oporność minimalna — do stosunku $\frac{L}{CR_n}$. Dla zmniejszenia wartości L dioda jest podłączona do specjalnego uchwytu z mosiężnych bloków.

W niektórych przypadkach, na przykład przy pomiarach charakterystyki diod w zależności od temperatury, opornik stabilizujący R_b nie powinien być umieszczony zbyt blisko diody tunelowej, gdyż zmiany je-



Rys. 7-9. Charakterograf do diod tunelowych

go wartości mogą spowodować niestabilność lub błędy pomiarowe. Wówczas też można podłączyć dodatkowy kondensator C_d i tym samym pominać dodatkową indukcyjność L_d wynikającą z użycia dłuższych przewodów. Jeżeli warunek $r_s > \frac{L}{CR_n}$ nie jest spełniony, można zwiększyć wartość r_s przez dodanie bezindukcyjnego opornika r_d .

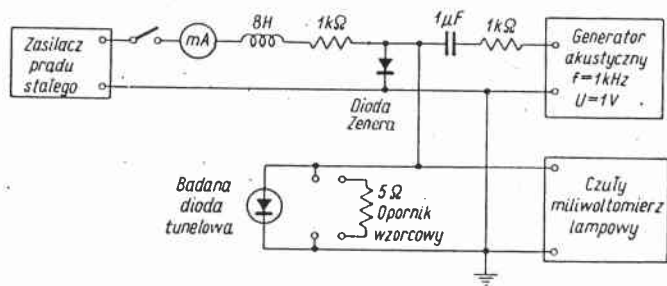
Tęgo rodzaju koncepcję przedstawili H. G. Dill i M. R. Macpherson [F4]. Rys. 7-8 przedstawia schemat przystawki oraz specjalnie za-

projektowanych przedwzmacniaczy X i Y. Opornik $R_s = 0,5 \div 5 \Omega$ powoduje spadek napięcia proporcjonalny do prądu diody. Kondensatory 1000 pF nie są niezbędne, lecz zwiększają stabilność, zapobiegając niepożądanym oscylacjom. Pojemności 1000 pF są wykonane z odizolowanych od siebie miką i olejem silikonowym bloków miedzianych uchwyty diody. Dla uzyskania mniejszego rozpraszania mocy w badanych diodach zastosowano dla zasilania nie połówki sinusoidalnego napięcia, lecz jedynie ich części. W tym celu, w szereg ze źródłem zasilającym podłączono tranzystor mocy, który jest przełączany tylko w chwili narastania sinusoidy. Układ spustowy Schmitta wytwarza napięcie kwadratowe dla sterowania układem przełączającym. Czas przełączania reguluje przesuwnik fazowy oraz impuls o zmiennej szerokości z układu Schmitta. Oporniki R_c mają wartość odpowiednio dużą, tak że generator nie wpływa na stabilność uchwytu diodowego.

Rys. 7-9 pokazuje wygląd zewnętrzny charakterografu do diod tunelowych zbudowanego w Zakładzie Doświadczalnym P. I. E.

7.3. Pomiar oporności szeregowej diody

Oporność szeregową r_s diody tunelowej można wyznaczyć ze zdjętej statycznie charakterystyki wstecznej, lub za pomocą przystawki oscyloskopowej przy zmienionym kierunku podłączenia diody. Jest rzeczą jasną,



Rys. 7-10. Układ do pomiaru oporności szeregowej

że ze względu na występowanie oporności dodatniej zagadnienie niestabilności staje się nieistotne.

Sam pomiar opiera się na tej zasadzie, że dla dużych prądów oporność bariery dąży do zera i tym samym impedancja diody uwarunkowana jest jedynie opornością materiału półprzewodnikowego, rozplywem linii prądu i ewentualnie opornością na styku bazy. Ze względu na to, że oporność półprzewodnika (czynnik dominujący) jest bardzo mała, występują również małe wartości r_s . Dla dostatecznie dużych prądów charakterystyka wsteczna przebiega liniowo i tym samym odwrotność tan-

gensa kąta nachylenia wyznacza wartość r_s . Obszar omowy osiąga się jednak dopiero przy natężeniach prądu rzędu $(50 \div 100) I_p$, co może niekiedy powodować rozproszenie zbyt dużej mocy, mimo że napięcie na diodzie będzie stosunkowo niewielkie (np. rzędu 100 mV). Z tego względu pożądane byłoby opracowanie impulsowej metody pomiaru oporności szeregowej diod tunelowych.

Prosty układ pomiarowy [C13] składa się z zasilacza prądu stałego, akustycznego generatora sygnałowego oraz dokładnego miliwoltomierza lampowego (rys. 7-10). Prąd stały z zasilacza polaryzuje diodę tunelową do napięcia w zakresie omowym. Równocześnie do diody zostaje doprowadzony sygnał o napięciu 1 mV i o częstotliwości 1 kHz. Z miliwoltomierza (ewentualnie wycechowanego w omach) odczytuje się napięcie na diodzie. Tego rodzaju miernik można łatwo wycechować przez podłączenie zamiast diody wzorcowego kilkuomowego opornika.

7.4. Pomiar oporności ujemnej

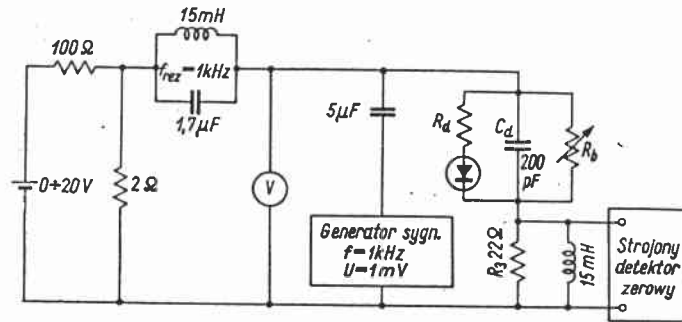
Przy omawianiu przystawki charakteroskopu Dilla i Macphersona zwrócono uwagę na możliwość wyznaczenia oporności ujemnej przez dołączenie dodatkowego opornika dla uzyskania na ekranie oscyloskopu linii poziomej w zakresie napięć od U_p do U_v . W tym przypadku zmierzona wartość R_n jest jednak mało dokładna; reprezentuje ona raczej $R_n \min$. Użytkownika diod tunelowych często jednak będzie interesowała możliwość wyznaczenia wartości R_n w punkcie odpowiadającym określonemu napięciu stałemu polaryzującemu diodę tunelową.

Zwróćmy uwagę, że R_n jest parametrem wewnętrznym, co zresztą wynika z układu zastępczego diody (rys. 4-4). Wyznaczenie wartości oporności ujemnej z charakterystyki statycznej $I-U$ daje raczej sumę $-R_n + r_s$. Boczniując diodę opornikiem R_b o wartości równej $R_n - r_s$ można metodą zerową dokonać pomiaru R_n , przy czym podobnie jak przy oscyloskopowym wykreślaniu charakterystyki $I-U$ muszą być spełnione warunki stabilności.

Jeżeli częstotliwość odcięcia opornościowego jest mniejsza od częstotliwości samorezonansowej: $f_R < f_N$, to wtedy na zaciski diody należy włączyć dostatecznie dużą pojemność w celu zapewnienia warunków stabilności dla prądu zmiennego, a jednocześnie dostatecznie małą, aby można ją było pominąć przy częstotliwości pomiarowej ($f_{pom} = 1 \text{ kHz}$). Jeżeli natomiast częstotliwość samorezonansowa jest mniejsza od częstotliwości odcięcia opornościowego: $f_N < f_R$, to wówczas wystarczy włączyć w szereg z diodą dodatkowy opornik R_d . Zwiększając w ten sposób wartość r_s , obniża się częstotliwość f_R zgodnie ze wzorem:

$$f_R = \frac{1}{2\pi R_n C} \sqrt{\frac{R_n}{r_s + R_d} - 1}$$

Na rys. 7-11 przedstawiony jest schemat miernika oporności R_n . Napięcie polaryzujące diodę otrzymuje się z dzielnika $100\Omega/2\Omega$ i mierzy się je woltmierzem. Obwód rezonansowy 1 kHz eliminuje wpływ obciążenia generatora opornikiem $R = 2\Omega$. Sygnał pomiarowy o częstotliwości $f = 1\text{ kHz}$ i możliwie małej amplitudzie w celu zwiększenia dokładności



Rys. 7-11. Schemat miernika do pomiaru oporności ujemnej

zostaje doprowadzony do diody. Spadek napięcia na cewce 15 mH jest znikomo mały. Napięcie na oporniku R_3 mierzy się za pomocą bardzo czułego, selektywnego detektora zerowego. Opornik R_d włączony jest w przypadku $f_R > f_X$. Pojemność C_d stanowią dwa wejściowe odizolowane od siebie miki blokowe mosiężne, do których podłączona jest dioda tunelowa (przy zachowaniu możliwie najkrótszych końcówek).

Miernik ten może służyć do:

- pomiaru napięcia stałego, przy którym R_n przyjmuje pewną z góry określoną wartość,
- wyznaczenia minimalnej wartości R_n oraz odpowiadającego jej napięcia doprowadzonego,
- wykreślenia zależności $R_n = f(U)$,
- wyznaczenia napięć: szczytowego U_p i doliny U_v .

W pierwszym przypadku oporność R_b musi być równa:

$$R_b = R_n - r_s$$

Napięcie polaryzujące diodę zmienia się aż do uzyskania zerowego wskazania na detektorze. Jeżeli oporność R_n jest większa od minimalnej wartości $R_{n, min}$, wystąpią dwa położenia zerowe.

W drugim przypadku, chcąc znaleźć wartość $R_{n, min}$, należy zmniejszać R_b aż do otrzymania jednego położenia zerowego.

W przypadku trzecim należy dobrać odpowiednią wartość napięcia polaryzującego i zmierzyć R_b , przy której uzyskuje się zerowe wskazania na detektorze. Wartość R_n wyznacza się z sumy $R_b + r_s$ (+ ewentualnie dodana oporność R_d).

Wreszcie w ostatnim przypadku korzystamy z tego, że oporność różniczkowa w punktach U_p i U_v ma wartość nieskończoną.

UZUPELNIENIE I
TABLICA CAŁKI FERMIEGO-DIRACA

φ	$F_{1/2}(\varphi)$	φ	$F_{1/2}(\varphi)$	φ	$F_{1/2}(\varphi)$
-4,0	0,016 128	1,0	1,396 375	6,0	10,144 28
-3,9	0,017 812	1,1	1,489 372	6,1	10,387 03
-3,8	0,019 670	1,2	1,586 323	6,2	10,631 90
-3,7	0,021 721	1,3	1,687 226	6,3	10,878 86
-3,6	0,023 984	1,4	1,792 068	6,4	11,127 89
-3,5	0,026 480	1,5	1,900 833	6,5	11,378 98
-3,4	0,029 233	1,6	2,013 496	6,6	11,632 11
-3,3	0,032 269	1,7	2,130 027	6,7	11,887 26
-3,2	0,035 615	1,8	2,250 391	6,8	12,144 40
-3,1	0,039 303	1,9	2,374 548	6,9	12,403 54
-3,0	0,043 366	2,0	2,502 458	7,0	12,664 64
-2,9	0,047 842	2,1	2,634 072	7,1	12,927 69
-2,8	0,052 770	2,2	2,769 344	7,2	13,192 67
-2,7	0,058 194	2,3	2,908 224	7,3	13,459 58
-2,6	0,064 161	2,4	3,050 659	7,4	13,728 39
-2,5	0,070 724	2,5	3,196 598	7,5	13,999 10
-2,4	0,077 938	2,6	3,345 988	7,6	14,271 68
-2,3	0,085 864	2,7	3,498 775	7,7	14,546 12
-2,2	0,094 566	2,8	3,654 905	7,8	14,822 41
-2,1	0,104 116	2,9	3,814 326	7,9	15,100 53
-2,0	0,114 588	3,0	3,976 985	8,0	15,380 48
-1,9	0,126 063	3,1	4,142 831	8,1	15,662 24
-1,8	0,138 627	3,2	4,311 811	8,2	15,945 80
-1,7	0,152 373	3,3	4,483 876	8,3	16,231 14
-1,6	0,167 397	3,4	4,658 977	8,4	16,518 26
-1,5	0,183 802	3,5	4,837 066	8,5	16,807 14
-1,4	0,201 696	3,6	5,018 095	8,6	17,097 76
-1,3	0,221 193	3,7	5,202 020	8,7	17,390 13
-1,2	0,242 410	3,8	5,388 795	8,8	17,684 23
-1,1	0,265 471	3,9	5,578 378	8,9	17,980 04
-1,0	0,290 501	4,0	5,770 726	9,0	18,277 56
-0,9	0,317 630	4,1	5,965 80	9,1	18,576 77
-0,8	0,346 989	4,2	6,163 56	9,2	18,877 68
-0,7	0,378 714	4,3	6,363 96	9,3	19,180 26
-0,6	0,412 937	4,4	6,566 98	9,4	19,484 51
-0,5	0,449 793	4,5	6,772 57	9,5	19,790 41
-0,4	0,489 414	4,6	6,980 70	9,6	20,097 96
-0,3	0,531 931	4,7	7,191 34	9,7	20,407 15
-0,2	0,577 470	4,8	7,404 45	9,8	20,717 97
-0,1	0,626 152	4,9	7,620 01	9,9	21,030 42
0,0	0,678 094	5,0	7,837 97	10,0	21,344 47
0,1	0,733 403	5,1	8,058 32		
0,2	0,792 181	5,2	8,281 03		
0,3	0,854 521	5,3	8,506 06		
0,4	0,920 505	5,4	8,733 39		
0,5	0,990 209	5,5	8,962 99		
0,6	1,063 694	5,6	9,194 85		
0,7	1,141 015	5,7	9,428 93		
0,8	1,122 215	5,8	9,665 21		
0,9	1,307 327	5,9	9,903 67		