621.315.592:546.28

TADEUSZ JANICKI

Dyfuzja donorów i akceptorów w krzemie

Rękopis dostarczono 14.12.1962

Duża precyzja przy wykonywaniu płaskich i równoległych przejść $p-n$, przy jednoczesnym braku wad charakterystycznych dla techniki stopowej, uczyniły z dyfuzji przodującą metodę wykonywania przyrządów półprzewodnikowych.

Zjawisko maskującego działania warstwy tlenków krzemu na dyfuzję niektórych domieszek pozwala ograniczyć ją do wybranych obszarów powierzchni krzemu, pozbawionych pokrycia warstwą SiO_2, co pozwala w efekcie uzyskać przejście p—n o żądanej konfiguracji.

Możliwość przeprowadzania procesu dyfuzji domieszek do krzemu w gazach nośnych znacznie upraszcza aparaturę, a zależność parowalności domieszek od rodzaju gazu nośnego ułatwia regulację koncentracji powierzchniowej.

W artykule podano — obok praw rządzących procesem dyfuzji — sposoby pomiaru i wykonywania obliczeń podstawowych parametrów dyfuzji. W części technologicznej zamieszczono opis aparatury, podano technologię dyfuzji szeregu pierwiastków donorowych i akceptorowych oraz zastosowanie dyfuzji do wytwarzania pojedynczych i wielokrotnych przejść $p-n$.

1. WSTĘP

Technika dyfuzji domieszek donorowych i akceptorowych w germanie i krzemie zajęła w ostatnich latach pod względem stosowalności w technologii elementów półprzewodnikowych nie tylko równorzędne miejsce z techniką stopową, lecz także wyprzedziła tę ostatnią w dziedzinie technologii germanowych tranzystorów wielkiej częstotliwości, a już zdecydowaną przewagę nad metodami stopowymi uzyskała w technologii tranzystorów krzemowych. Przyczyną tej przewagi w przypadku krzemu jest konieczność wykonywania procesów stopowych w temperaturach na ogół znacznie wyższych, niż ma to miejsce w przypadku germanu. Okoliczność ta w połączeniu ze specyfiką wtapianych materiałów stwarza bardzo duże trudności w uzyskaniu w sposób powtarzalny jednorodnych i płaskich przejść p-n o zadanym kształcie. Drugą, nie mniej ważną przyczyną są niezadowalające właściwości mechaniczne wtopionych obszarów, wywołane kruchością eutektyków i wprowadzanych przez nie naprężeń.

Powyższe wady czynią technikę stopową w przypadku technologii tranzystorów krzemowych bardzo mało atrakcyjną, zwłaszcza że nie mo-

że ona konkurować z tak bezsprzecznymi zaletami techniki dyfuzyjnej, jak możliwość uzyskania żądanej koncentracji domieszek i ich gradientu w danym punkcie wykonywanego przyrządu półprzewodnikowego, dużej powtarzalności procesu, a w szczególności bardzo wysokiej precyzji. Precyzja ta pozwala na uzyskanie w sposób całkowicie powtarzalny przejść typu p-n-p (lub n-p-n) o grubości obszaru typu n (lub p) nie przekraczającej 2 μ, co byłoby nie do pomyślenia w przypadku zastosowania techniki stopowej.

Połączenie techniki dyfuzyjnej z techniką fotolitograficzną pozwala na uzyskanie przejść p-n o określonej konfiguracji i bardzo małych wymiarach, dzięki czemu możliwe jest wykonywanie w sposób powtarzalny tranzystorów o częstotliwościach granicznych f_α sięgających kilkuset megaherców.

Ze względu na swoje zalety technika dyfuzyjna odgrywa dominującą rolę nie tylko w technologii krzemowych tranzystorów wielkiej częstotliwości, lecz także w technologii takich elementów krzemowych, jak tranzystory mocy, diody prostownicze, diody regulacyjne (Zennera), diody sterowane (tyratrony) czy wreszcie ogniwa fotowoltaiczne. Ostatnio technika dyfuzyjna wkroczyła w jeszcze jedną dziedzinę elektroniki półprzewodnikowej, a mianowicie w dziedzinę obwodów monolitycznych, gdzie odgrywa dominującą rolę jako technika służąca nie tylko do wykonywania tranzystorów i diod, będących częściami składowymi danego obwodu monolitycznego, lecz także oporników i kondensatorów, wchodzących w skład tego obwodu.

1. OGÓLNE WIADOMOŚCI O PROCESIE DYFUZJI

1.1. Podstawowe prawa dyfuzji

Najbardziej istotnym zagadnieniem dyfuzji domieszek z punktu widzenia jej zastosowania w technologii przyrządów półprzewodnikowych jest znajomość rozkładu koncentracji wdyfundowanych domieszek. Zagadnienie to w większości przypadków daje się rozwiązać analitycznie wystarczająco dokładnie w odniesieniu do zastosowań praktycznych. W niektórych przypadkach, zawierających szczególnie skomplikowane warunki na wartości brzegowe, dokładne rozwiązanie zagadnienia jest bardzo trudne lub wręcz niemożliwe. Tym niemniej, po wprowadzeniu pewnych uproszczeń, możliwe są rozwiązania dające w wyniku wartości przybliżone. Ponieważ w większości przypadków będziemy się interesowali położeniem przejść p-n równoległych do jednej z powierzchni półprzewodnika, z której odbywa się dyfuzja, więc zawężimy zagadnienie do przypadku dyfuzji w jednym tylko wymiarze.

Rozważmy element półprzewodnika, w którym dyfuzja domieszek ze źródła o koncentracji N_s, znajdującego się na powierzchni S, powoduje zmianę koncentracji domieszek wewnątrz rozpatrywanego elementu z wartości N_1 (przed dyfuzją) na wartość N_x, zależną od odległości badanej powierzchni X od powierzchni S.

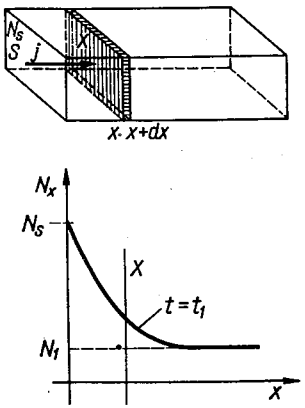
Podstawowym prawem rządzącym procesem dyfuzji jest prawo Ficka, które stwierdza, że gęstość strumienia j dyfundującej domieszki jest proporcjonalna do gradientu koncentracji tej domieszki:

$$j = -D \left.\frac{\partial N_x}{\partial x}\right|_x , \tag{1}$$

gdzie N_x jest koncentracją dyfundującej domieszki [at/cm^3] przy badanej powierzchni X.

Występujący w zależności (1) współczynnik proporcjonalności D nosi nazwę stałej dyfuzji i ma wymiar [cm^2/sec].

W rozważaniach naszych ograniczymy się do dyfuzji w monokrysztale, dzięki czemu odpadną dodatkowe warunki wynikające z zakłóceń na granicach ziaren kryształów. Ponadto będziemy rozpatrywali krańcowo małe ilości domieszek przy stosunkowo niskich koncentracjach, dzięki



Rys. 1. Objaśnienie do rozwiązania równania dyfuzji

czemu będziemy mogli zaniedbać zmianę wartości współczynnika dyfuzji, który, jak wiadomo, nie jest wielkością stałą przy wysokich koncentracjach dyfundującego czynnika. Założenie dotyczące małej ilości domieszek pozwala dodatkowo na zaniedbanie zmian w wymiarach półprzewodnika spowodowanych dyfuzją i przyjęcia jego powierzchni jako płaszczyzny odniesienia, względem której będziemy mierzyli interesujące nas wielkości dyfuzji. Po wprowadzeniu powyższych uproszczeń obliczymy przyrost koncentracji domieszek w elemencie o długości dx i przekroju poprzecznym S, prostopadłym do kierunku dyfuzji (rys. 1).

Ilość domieszek wchodzących do rozpatrywanego elementu w czasie dt wyrazi się jako:

$$Q_1 = -D\left(\frac{\partial N_x}{\partial x}\right)_x \cdot S \cdot dt,$$

zaś ilość domieszek wychodzących z tego elementu w tym samym czasie jako:

$$Q_2 = -D\left(\frac{\partial N_x}{\partial x}\right)_{x+dx} \cdot S \cdot dt.$$

Różnica obu wyrażeń określająca przyrost ilości domieszek w rozpatrywanym elemencie w czasie dt, dla warunku $dx \to 0$ da się sprowadzić po prostych przekształceniach do następującej zależności:

$$\frac{\partial N_x}{\partial t} = D\frac{\partial^2 N_x}{\partial x^2}. \tag{2}$$

Zależność (2) nosi nazwę drugiego prawa Ficka.

Rozwiążemy teraz równanie (2) dla przypadku dyfuzji ze źródła dającego stałą koncentrację powierzchniową (rys. 1).

Warunki brzegowe dla takiego przypadku będą następujące:

1) w czasie $t = 0$ $N_x = N_1$ dla $0 \leqslant x < \infty$,
2) dla $x = 0$ $N_x = N_s$ dla $0 < t < \infty$.

Wprowadzając nową zmienną:

$$\lambda = \frac{x}{\sqrt{t}}$$

możemy przekształcić równanie (2) na równanie różniczkowe zwyczajne o postaci:

$$2D\frac{d^2N}{d\lambda^2} + \lambda\frac{dN}{d\lambda} = 0, \tag{3}$$

skąd po rozwiązaniu otrzymamy wyrażenie na N_x:

$$N_x = A \cdot \int_0^{\lambda} e^{-\lambda^2/4D} \cdot d\lambda + B,$$

gdzie A i B są stałymi całkowania.
Wprowadzając z kolei dla wygody nową zmienną:

$$\xi = \lambda/2D$$

i korzystając z uprzednio wprowadzonej zależności:

$$\lambda = \frac{x}{\sqrt{t}}$$

otrzymamy następujące wyrażenie na wartość koncentracji w punkcie x:

$$N_x = A' \cdot \int_0^{x/2\sqrt{Dt}} e^{-\xi^2} \cdot d\xi + B, \tag{4}$$

gdzie $A' = 2\sqrt{D} \cdot A$.

Po wyznaczeniu stałych całkowania A' i B, równanie (4) przyjmie postać:

$$\frac{N_x - N_1}{N_s - N_1} = 1 - \frac{2}{\pi} \int_0^{x/2\sqrt{Dt}} e^{-\xi^2} d\xi. \tag{5}$$

Wprowadzając pojęcie funkcji błędu:

$$erf\, t \overset{\text{df}}{=} \int_0^t e^{-\tau^2} d\tau$$

oraz

$$erfc\, t \overset{\text{df}}{=} 1 - erf\, t$$

możemy napisać równanie (5) w postaci najczęściej spotykanej w literaturze, a mianowicie:

$$\frac{N_x - N_1}{N_s - N_1} = erfc\,(x/2\sqrt{Dt}). \tag{6}$$

Jeżeli w rozpatrywanym półprzewodniku koncentracja domieszek przed dyfuzją, N_1, jest równa zeru, wówczas równanie (6) przyjmie postać:

$$N_x = N_s\, erfc\,(x/2\sqrt{D \cdot t}), \tag{7}$$

Rozważony przypadek dyfuzji ze źródła o stałej koncentracji domieszek jest jednym z najczęściej spotykanych w praktyce.

Inny, spotykany również w praktycznych zastosowaniach, jest przypadek dyfuzji ze źródła posiadającego ograniczoną ilość atomów domieszki na powierzchni półprzewodnika.

Rozwiązanie równania dyfuzji (2) dla takiego przypadku ma postać:

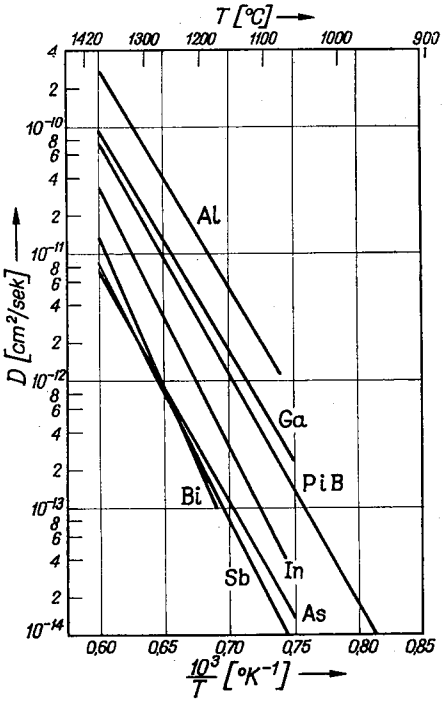
$$N_x = Q\,(\pi \cdot D \cdot t)^{-1/2} \cdot \exp\,(-x^2/4D \cdot t), \tag{8}$$

gdzie Q jest liczbą atomów domieszki przypadających na 1 cm^2 powierzchni półprzewodnika przed rozpoczęciem procesu dyfuzji. Równania (6), (7) i (8) nie podają bezpośrednio zależności koncentracji N_x od temperatury, w której przeprowadzany jest proces dyfuzji. Zależność ta kryje się przede wszystkim w stałej dyfuzji D, związanej z temperaturą przez następujące równanie:

$$D = D_0 \exp\,(-\Delta W/RT), \tag{9}$$

gdzie:

D_0 — stała wartość dla danego rodzaju domieszki,
ΔW — energia aktywacji dyfuzji [kal],
R — stała gazowa 1,98 [kal/°K mol],
T — temperatura absolutna [°K].



Rys. 2. Zależność stałych dyfuzji D domieszek donorowych i akceptorowych w krzemie od temperatury dyfuzji

Tablica 1

Wartości stałych D_0 i ΔH dla dyfuzji domieszek w krzemie

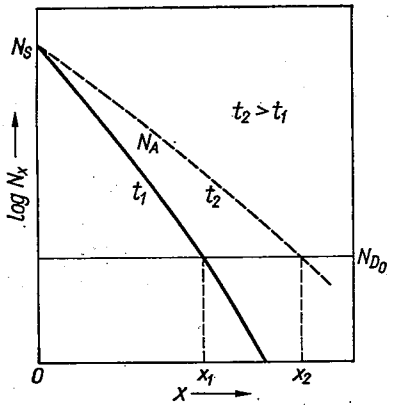
Donory			Akceptory		
Nazwa	D_0 cm²/s	ΔH [kcal]	Nazwa	D_0 [cm²/sek]	ΔH [kcal]
Antymon	5,6	91	Aluminium	8,0	80
Arsen	0,32	82	Bor	10,5	85
Bizmut	1033	107	Gal	3,6	81
Fosfor	10,5	85	Ind	16,5	90

Przebieg zależności stałych dyfuzji w funkcji temperatury dla najczęściej stosowanych w przypadku krzemu domieszek donorowych i akceptorowych podano na rys. 2, zaś wartości energii aktywacji ΔW i stałych D_0 zamieszczono w tablicy 1.

1.2. Zastosowanie dyfuzji do wytwarzania przejść p-n

Jednym z zasadniczych zastosowań dyfuzji w technologii elementów półprzewodnikowych jest wytwarzanie przy jej pomocy przejść p-n. Uzyskanie przejścia p-n w materiale o początkowej koncentracji donorów N_{D_0} (lub koncentracji akceptorów N_{A_0}) jest możliwe dzięki dyfuzji domieszek akceptorowych (lub donorowych), o ile koncentracja wdyfundowanych domieszek będzie większa od koncentracji początkowej N_{D_0} (lub N_{A_0}).

Rysunek 3 przedstawia rozkład koncentracji domieszek w próbce krzemowej o koncentracji początkowej donorów N_{D_0} po dyfuzji domieszek



Rys. 3. Rozkład koncentracji domieszek akceptorowych wdyfundowanych w krzem typu n o koncentracji donorów N_{D_0}, dla dwu różnych czasów dyfuzji

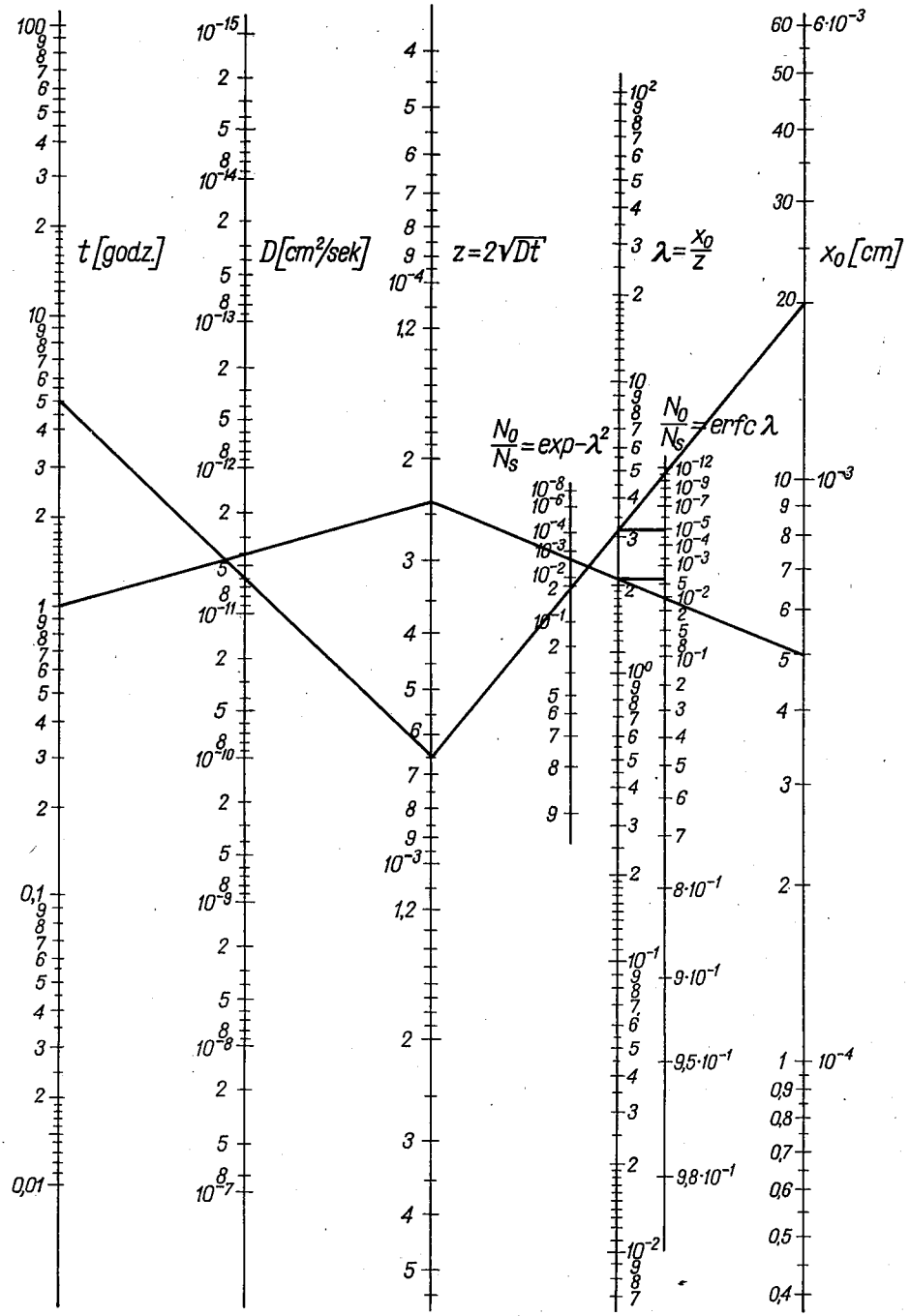
akceptorowych w temperaturze T_1 i czasie t_1. Linią przerywaną oznaczono rozkład koncentracji akceptorów, jaki powstałby gdyby dyfuzję przeprowadzano w czasie $t_2 > t_1$ przy zachowaniu takich samych pozostałych parametrów procesu.

Ponieważ przejście p-n powstaje w miejscu, gdzie koncentracja wdyfundowanych akceptorów jest równa koncentracji donorów, zmiana czasu dyfuzji umożliwia otrzymanie wymaganej odległości przejścia p-n od powierzchni płytki. Odległość ta jest, ogólnie biorąc, proporcjonalna do $\sqrt{t}$. Znając odległość przejścia p-n od powierzchni płytki X_1, uzyskaną po czasie dyfuzji t_1, możemy określić czas potrzebny do osiągnięcia wymaganej odległości X_2, o ile pozostałe parametry procesu pozostaną stałe.

Czas ten wyrazi się następującą prostą zależnością:

$$t_2 = t_1 \frac{X_2^2}{X_1^2}. \tag{10}$$

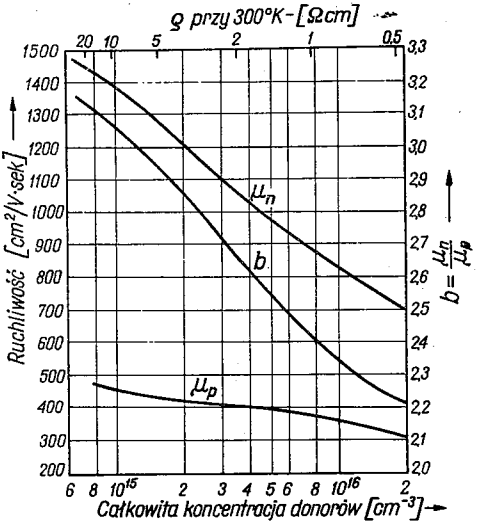
Wymaganą odległość przejścia p-n można uzyskać również przy stałej wartości t i D wykorzystując zależność rozkładu koncentracji dyfundowanych domieszek od koncentracji powierzchniowej N_s. Zakres tych zmian



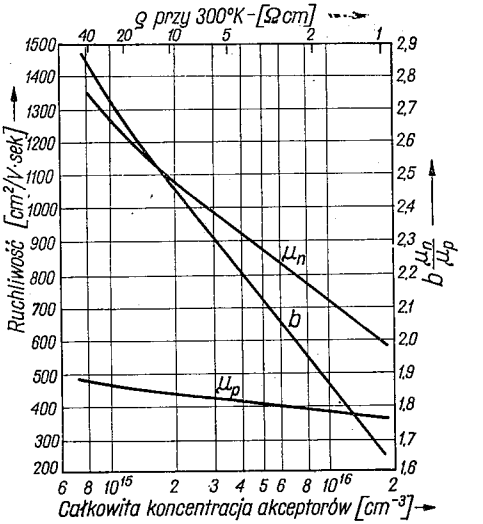
Rys. 4. Nomogram do rozwiązywania równania dyfuzji

jest oczywiście ograniczony przez możliwości zmian koncentracji powierzchniowej, której wielkość przy stałej temperaturze dyfuzji zależy od rodzaju użytej domieszki, jej temperatury oraz od wybranego sposobu doprowadzenia par domieszki do powierzchni krzemu.

Dokładna znajomość odległości przejścia p-n od powierzchni płytki może posłużyć do obliczenia koncentracji powierzchniowej wdyfundowanych domieszek, o ile znane są pozostałe parametry dyfuzji, tzn. czas trwania procesu, stała dyfuzji oraz oporność właściwa krzemu. Wykonanie tego



Rys. 5. Zależność ruchliwości nośników w krzemie typu n oraz jego oporu właściwego od koncentracji donorów



Rys. 6. Zależność ruchliwości nośników w krzemie typu p oraz jego oporu właściwego od koncentracji akceptorów

obliczenia ułatwia nam fakt istnienia w miejscu przejścia p-n równych koncentracji donorów i akceptorów. Jeżeli więc dyfuzja akceptorów w krzem, zawierający stałą koncentrację donorów N_{D_0} spowodowała powstanie przejścia p-n w odległości X_1 od powierzchni płytki, wówczas równanie (7) po wstawieniu w miejsce N_x wartości N_{D_0}, a w miejsce x wartości X_1 pozwoli na obliczenie szukanej wartości koncentracji powierzchniowej akceptorów. Wykonanie tego rodzaju obliczenia znakomicie się uprości, jeżeli skorzystamy z nomogramu przedstawionego na rys. 4. Sposób posługiwania się nomogramem zilustrujemy za pomocą dwóch przykładów.

Przykład 1. Obliczyć koncentrację powierzchniową galu wdyfundowanego w krzem typu n o oporności 2 Ω cm, o ile wiadomo, że po czasie dyfuzji galu $t = 1$ h, wykonanej w temperaturze 1200 °C głębokość przejścia p-n wynosiła 5 μ.

A oto sposób postępowania:

Korzystając z wykresu zamieszczonego na rys. 2 znajdujemy wartość

stałej dyfuzji galu w temperaturze 1200 °C, a z wykresu zamieszczonego na rys. 5 odczytujemy koncentrację donorów dla krzemu o oporności $\varrho = 2\ \Omega$ cm.

Mamy więc

$$D = 4 \cdot 10^{-12}\ \text{cm}^2/\text{sec},$$

$$N_{D_0} = 3{,}5 \cdot 10^{15}\ \text{at/cm}^3.$$

Odkładając na odpowiednich osiach wartości czasu dyfuzji t oraz stałej dyfuzji D znajdujemy parametr $z = 2\sqrt{Dt}$, a następnie łącząc linią uzyskaną wartość „z" z wartością $X_0 = 5\,\mu$ otrzymujemy wartość parametru $\lambda = \frac{x_0}{z} \approx 2{,}1$.

Jeżeli proces dyfuzji został wykonany przy stałej koncentracji powierzchniowej galu, wówczas jest on opisany przez funkcję $erfc\,(\lambda)$. Dla takiego przypadku szukaną przez nas wartość $erfc\,(\lambda) = \frac{N_0}{N_s}$ znajdziemy wystawiając z punktu $\lambda = 2{,}1$ prostopadłą w kierunku osi $erfc\,(\lambda)$. Daje to w wyniku wartość $\frac{N_0}{N_s} \simeq 4 \cdot 10^{-3}$.

Ponieważ w naszym przypadku koncentracja donorów wynosiła $3{,}5 \cdot 10^{15}$, więc szukana wartość koncentracji powierzchniowej równa $\frac{N_0}{erfc\,(\lambda)}$ wyniesie około $8{,}75 \cdot 10^{17}$ at/cm^3.

Przykład 2. Jak długo należy przeprowadzać dyfuzję fosforu przy temperaturze 1300 °C, w krzem typu p o oporności 10 Ω cm, ażeby otrzymać przejście p-n na głębokości 20 μ. Wiadomo ponadto, że warunki procesu zostały tak dobrane, aby otrzymać stałą koncentrację powierzchniową fosforu, wynoszącą $2 \cdot 10^{20}$ at/cm^3.

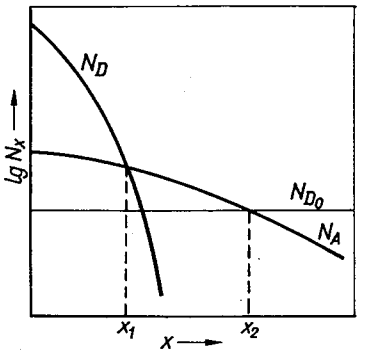
Korzystając z wykresu zamieszczonego na rys. 6 znajdujemy wartość koncentracji akceptorów dla krzemu typu p o oporności 10 Ω cm. Wynosi ona około $1{,}9 \cdot 10^{15}$ at/cm. Obliczając wartość stosunku $\frac{N_0}{N_s} = \frac{1{,}9 \cdot 10^{15}}{2 \cdot 10^{20}} \approx 10^{-5}$ i odkładając ją na osi, znajdujemy wartość parametru $\lambda = 3{,}1$, a następnie po połączeniu tego punktu z punktem $X_0 = 20\,\mu$ znajdujemy potrzebną wartość parametru z. Biorąc następnie z wykresu zamieszczonego na rys. 2 wartość stałej dyfuzji fosforu w temperaturze 1300 °C (wynosi ona około $6 \cdot 10^{-12}$ cm^2/sec) znajdujemy na osi t wymagany czas dyfuzji, wynoszący w naszym przypadku około 5 godzin.

1.3. Wytwarzanie wielokrotnych przejść typu n-p-n lub p-n-p przy pomocy dyfuzji dwóch domieszek

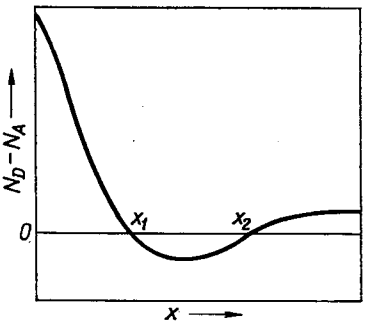
Dyfuzja dwóch domieszek ma istotne znaczenie dla wytwarzania wielokrotnych przejść typu n-p-n lub p-n-p, przydatnych do wykonywania

tranzystorów. Właściwie dobrana para domieszek akceptor-donor pozwala na otrzymanie tego rodzaju przejść nawet przy jednoczesnej dyfuzji obu tych domieszek. Jest to możliwe dzięki temu, że pierwiastki akceptorowe posiadają w krzemie znacznie większe stałe dyfuzji niż pierwiastki donorowe, a więc odwrotnie niż ma to miejsce w przypadku germanu. Odstępstwo od tej reguły wykazuje jedynie fosfor, którego stała dyfuzji w krzemie jest, praktycznie biorąc, taka sama, jak stała dyfuzji boru. Cechą donorów w krzemie jest natomiast to, że dają one duże wartości koncentracji powierzchniowych.

Na rys. 7 przedstawiono schemat ilustrujący zasadę uzyskiwania struktury typu n-p-n przy pomocy dyfuzji pierwiastka akceptorowego i donorowego w krzem o stałej koncentracji donorów N_{D_0}. Pierwsze przejście (n-p) powstaje w miejscu, gdzie koncentracja akceptorów, re-



Rys. 7. Dyfuzja domieszki akceptorowej i donorowej w krzem typu n dla uzyskania struktury typu n-p-n



Rys. 8. Przebieg wypadkowej koncentracji domieszek dla przypadku przedstawionego na rys. 7

prezentowana przez krzywą N_A, jest równa wypadkowej koncentracji donorów ($N_D + N_{D_0}$). Drugie przejście (p-n) powstaje w punkcie X_2, w którym koncentracja akceptorów jest równa koncentracji donorów N_{D_0}.

Przebieg wypadkowej koncentracji domieszek w funkcji odległości od powierzchni płytki, dla rozkładu domieszek z rys. 7, przedstawiono na rysunku 8.

Struktury typu n-p-n, przedstawione na rys. 7 i 8 można otrzymać przy pomocy dwóch metod.

Pierwsza z metod polega na jednoczesnej dyfuzji domieszki akceptorowej i donorowej. Pożądaną strukturę osiąga się tu dzięki różnicy stałych dyfuzji i koncentracji powierzchniowych wybranych domieszek. Najczęściej stosowanymi domieszkami są tu gal — jako akceptor — oraz arsen względnie antymon — jako donor.

Druga metoda polega na sukcesywnej dyfuzji obu domieszek. W pierwszej kolejności dyfunduje się domieszkę o małej szybkości dyfuzji, a następnie dopiero domieszkę szybko dyfundującą. Postępowanie to ma na

celu uniknięcie dużych zmian w przewidzianym rozkładzie koncentracji, które powstałyby, gdyby domieszkę szybko dyfundującą dyfundowano w pierwszej kolejności.

Metoda sukcesywnej dyfuzji pozwala na większą swobodę w wyborze zarówno temperatury dyfuzji, jak i temperatury domieszek, dzięki czemu możliwe jest otrzymanie wymaganych koncentracji powierzchniowych oraz stałych dyfuzji. Przy pomocy tej metody możliwe jest otrzymanie przejść typu *n-p-n* lub *p-n-p* nawet w przypadku, gdy stałe dyfuzji obu dyfundowanych domieszek, akceptorowej i donorowej, są równe. Czynnikiem regulującym położenie każdego ze złącz jest czas dyfuzji jednej i drugiej domieszki oraz wielkość koncentracji powierzchniowej. Przy obliczaniu krzywych rozkładu wdyfundowanych domieszek przyjmuje się, że każda domieszka dyfunduje niezależnie od drugiej i że oddziaływanie między nimi jest do pominięcia. Założenie to, aczkolwiek nie zawsze dostatecznie ścisłe, zwłaszcza w bezpośredniej bliskości powierzchni krzemu, gdzie koncentracje dyfundowanych domieszek są duże, okazuje się najczęściej wystarczająco dokładne dla zastosowań dyfuzji w technice tranzystorowej. Przyjęcie tego założenia umożliwia nam obliczenie krzywych rozkładu donorów i akceptorów, o ile znane są temperatura i czas dyfuzji, oporność właściwa krzemu oraz odległość obu przejść *p-n* od powierzchni płytki.

Przebieg postępowania jest następujący:

1) Jeżeli mamy do czynienia z przypadkiem dyfuzji podobnym do tego, jaki przedstawiono na rys. 7, wówczas znajomość N_{D_0}, odległości X_2 oraz czasu i temperatury dyfuzji pozwoli na obliczenie koncentracji powierzchniowej akceptorów N_{A_s}.

2) Znając N_{A_s} możemy obliczyć koncentrację akceptorów w punkcie X_1, a stąd i koncentrację wdyfundowanych donorów, ponieważ w punkcie X_1 $N_A = N_D + N_{D_0}$.

3) Znajomość N_D w punkcie X_1 pozwala z kolei na obliczenie koncentracji powierzchniowej dyfundowanych donorów, a w konsekwencji na obliczenie całej krzywej rozkładu domieszek.

Jest rzeczą oczywistą, że znajomość krzywych rozkładu umożliwia także obliczenie gradientów koncentracji domieszek w poszczególnych punktach, co jest bardzo ważne przy konstruowaniu urządzeń półprzewodnikowych.

1.4. Pomiar głębokości przejścia *p-n*

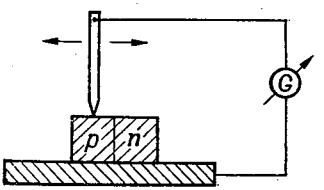
Znajomość odległości przejścia *p-n* od powierzchni płytki krzemowej bądź od innej płaszczyzny odniesienia, np. płaszczyzny drugiego przejścia *p-n*, jest bardzo ważnym czynnikiem nie tylko w technice wykorzystującej proces dyfuzji do wykonywania elementów półprzewodnikowych, lecz

także w badaniach podstawowych dotyczących parametrów samego procesu dyfuzji w półprzewodnikach, jak np. przy wyznaczaniu stałej dyfuzji, czy też, jak to pokazano uprzednio — przy określaniu koncentracji powierzchniowej dyfundowanych domieszek.

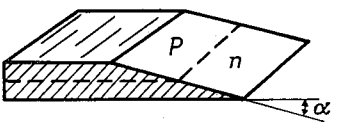
Dla przybliżonego zlokalizowania położenia przejścia p-n można posłużyć się metodą sondy termoelektrycznej, używanej powszechnie do określania typu przewodnictwa materiałów półprzewodnikowych. Zasada tego pomiaru jest pokazana na rys. 9.

Zmiana kierunku wychylenia galwanometru na przeciwny oznacza, że między tym i poprzednim położeniem termosondy znajduje się przejście p-n.

Przygotowanie płytki do pomiaru polega na zeszlifowaniu jej pod kątem α, wynoszącym zazwyczaj $3 \div 5^{\circ}$, dzięki czemu wymiary obszarów typu p i n na zeszlifowanej płaszczyźnie zostają znacznie powiększone (rys. 10). Szlifowanie płytki pod wybranym kątem może być wykonane



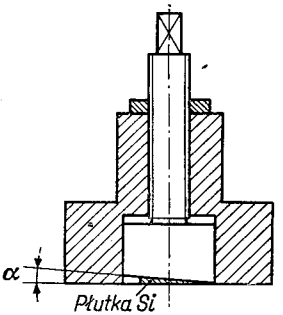
Rys. 9. Zasada przybliżonego określania położenia przejścia p-n



Rys. 10. Próbka z wykonanym skośnym szlifem

bardzo łatwo w adapterku pokazanym na rys. 11. Najczęściej stosowanymi materiałami ściernymi do szlifowania próbek są: proszek diamentowy, bardzo drobny karborund i tlenek glinu (alundum).

Wyznaczanie głębokości przejścia p-n przy pomocy termosondy daje nam tylko jej przybliżoną wartość w jednym punkcie próbki.



Rys. 11. Szablon do szlifowania próbek pod kątem α

Wyznaczenie przebiegu linii przejścia p-n na całej powierzchni szlifu wymagałoby bardzo dużej ilości pomiarów nie dając w zamian dokładnego obrazu przejścia p-n, a w szczególności ewentualnych lokalnych zmian głębokości dyfuzji wywołanych defektami kryształu. Jedną z zasadni-

czych wad tej metody jest ponadto jej mała dokładność wynikająca ze stosunkowo dużych wymiarów średnicy samego ostrza sondy.

Z powyższych względów opisana metoda ustępuje znacznie innej metodzie lokalizacji przejścia p-n, polegającej na ujawnieniu linii stanowiącej granicę dwóch obszarów o odmiennym typie przewodnictwa. Linię te otrzymuje się dzięki różnemu zabarwieniu obszaru typu p w porównaniu z obszarem typu n badanej próbki krzemowej po zanurzeniu jej w pewnych kąpielach barwiących.

W celu zwiększenia dokładności późniejszych pomiarów, a często w ogóle w celu ich umożliwienia (np. w przypadku wyznaczania głębokości złącza rzędu pojedynczych mikronów), badaną próbkę szlifuje się pod kątem $3 \div 5°$, podobnie jak w metodzie opisanej uprzednio. Przy szlifowaniu próbki należy zwrócić szczególną uwagę na to, by powierzchnia szlifu była możliwie gładka, gdyż wszelkie rysy i zadrapania powierzchni szlifu utrudniają w znacznej mierze uzyskanie dużej ostrości linii przejścia p-n. Ostrość linii granicznej przejścia p-n uzyskanej za pomocą barwienia krzemu zależy w dużym stopniu od czasu barwienia, który dla danych warunków procesu (skład roztworu, oświetlenie, temperatura) należy dobrać eksperymentalnie.

Po uzyskaniu wyraźnej linii przejścia p-n próbkę należy wyjąć z roztworu i szybko przemyć wodą destylowaną, gdyż dalsze pozostawienie roztworu barwiącego na powierzchni płytki powoduje rozmazywanie otrzymanego obrazu.

Dokładność, jaką uzyskuje się przy wyznaczaniu głębokości przejścia p-n w przypadku szlifowania próbki pod kątem, a następnie barwienia jej, zależy w bardzo znacznym stopniu od dokładności, z jaką potrafimy zmierzyć ten kąt. Dla uzyskania dużej dokładności w określeniu głębokości przejścia p-n nie można polegać na wartości kąta, którą reprezentuje szablon do szlifowania, ponieważ sama warstewka kleju używanego do mocowania próbki w szablonie może zmienić jej kąt ustawienia względem poziomej płaszczyzny szablonu.

Pomimo tego, że istnieje szereg metod pozwalających na stosunkowo dokładny pomiar wartości kąta, pod jakim została zeszlifowana próbka (np. metoda odbicia promienia świetlnego), w przypadkach, w których szczególnie zależy nam na dokładnym ustaleniu głębokości przejścia p-n, poleca się pomiar przy pomocy mikroskopu interferencyjnego. Tego rodzaju pomiar jest bardzo prosty i wymaga jedynie zabarwienia zeszlifowanej pod kątem próbki, a następnie zmierzenia głębokości badanego obszaru z dokładnością niezależną od kąta, pod jakim próbka została zeszlifowana.

1.5. Metody barwienia

Uprzednio wspomniano, że dzięki odmiennemu zabarwieniu obszarów typu p i typu n badanej próbki krzemowej, przejście p-n może być przed-

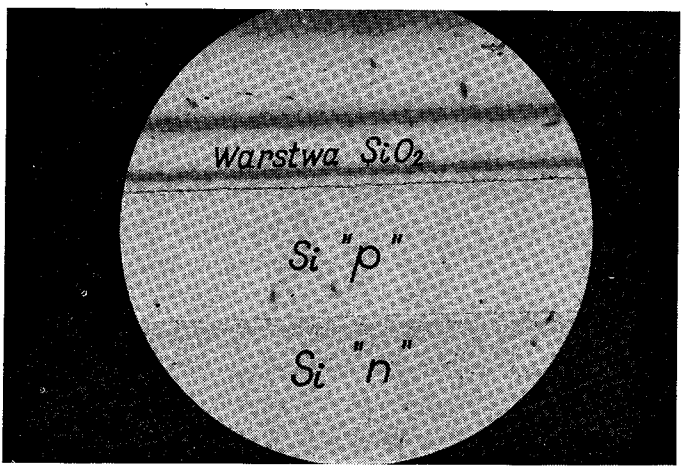
stawione na płaszczyźnie przekroju w postaci linii, stanowiącej granicę tych dwóch obszarów.

Omówimy teraz kilka z najczęściej stosowanych metod barwienia. Jedną z nich jest metoda podana przez Fullera i Ditzenbergera. Polega ona na pokryciu badanej powierzchni płytki kroplą fluorowodoru zawierającego niewielką domieszkę kwasu azotowego (0,1—0,2 ml 69,5% HNO na 100 ml 48,5% HF). Obszary typu p barwią się w takim roztworze na ciemno w stosunku do obszarów typu n.

Należy zaznaczyć, że w metodzie tej czas barwienia, jak również ilość roztworu i koncentracja użytych składników wpływają nie tylko na dokładność procesu, lecz także mogą spowodować inwersję, dzięki której na ciemno zostaną zabarwione obszary typu n, a nie obszary typu p.

Whoriskey opracował metodę, dzięki której otrzymuje się większą powtarzalność procesu barwienia. Metoda ta polega na silnym oświetleniu badanej próbki, zanurzonej w skoncentrowanym fluorowodorze. Pod skośnym oświetleniem obszary typu p mają wygląd srebrzysty, zaś obszary typu n mają zabarwienie czarne.

Bardzo dobre wyniki w odwzorowywaniu linii przejścia p-n uzyskuje się przez barwienie krzemu w roztworach zawierających sole metali oraz fluorowodór. Inne zabarwienie obszaru typu p w stosunku do obszaru typu n uzyskuje się tu dzięki różnym szybkościom osadzania się metalu (którego sól jest w roztworze) na obu typach krzemu (rys. 12).



Rys. 12. Przykład ujawnienia linii przejścia p-n w roztworze $AgNO_3$ z domieszką HF

Szybkość osadzania się metalu na powierzchni krzemu jest tym większa, im bardziej elektrododatni jest ten metal oraz im większe jest stężenie fluorowodoru w roztworze. Wzrost koncentracji HF w roztworze powoduje co prawda szybsze osadzanie się metalu, jednakże bardzo ujemnie wpływa na definicję linii przejścia p-n.

W tablicy 2 podano skład roztworów barwiących z tak dobranymi ilościami fluorowodoru, ażeby skompensować szybsze osadzanie się metali bardziej elektrododatnich i otrzymać w efekcie jednakowe szybkości osadzania się metali z poszczególnych roztworów. Nie wszystkie metale wymienione w tablicy 2 dają jednakowo dobre i mocne pokrycia. Względna adhezja metalu jest tym mniejsza, im bardziej elektrododatni jest metal.

Tablica 2

Składy rozworów służących do ujawnienia przejść *p-n* w krzemie

Metal osadzony	sól metalu	g/l	Ilość cm^3 48% HF/l
Sb	SbF_3	400	500
Cu	$CuSO_4 \cdot 5H_2O$	200	10
Ag	$AgNO_3$	40	4
Pt	$H_2PtCl_6 6H_2O$	3	2
Au	$AuCl_3$	10	0,3

Złoto daje pokrycia tak słabe, że mogą być one zniszczone nie tylko podczas przemywania próbki, lecz także nawet wskutek działania napięcia powierzchniowego roztworu. Najbardziej mocny depozyt dają: miedź i antymon, zaś najlepszą definicję złącza otrzymuje się w przypadku osadzania miedzi i srebra.

Roztwory podane w tablicy 2 dają w przybliżeniu taką samą grubość depozytu na krzemie typu *p*, jak na krzemie typu *n*, czyniąc definicję przejścia bardzo słabą lub wręcz niemożliwą. Jeżeli jednak osadzanie metalu będzie zachodziło w roztworze przy jednoczesnym silnym oświetleniu przejścia *p-n*, wówczas większa część metalu osadzi się na krzemie typu *p* dając wystarczająco ostrą linię rozgraniczającą oba obszary.

Warunkiem tego, aby w opisanej metodzie metal osadzał się szybciej na krzemie typu *p* jest zanurzenie silnie oświetlonego przejścia *p-n* w dużej objętości roztworu, gdyż pokrycie płytki tylko cienką warstwą roztworu powoduje szybsze osadzanie się metalu na krzemie typu *n*. Przyczyny tej odwrotności nie są dotychczas znane.

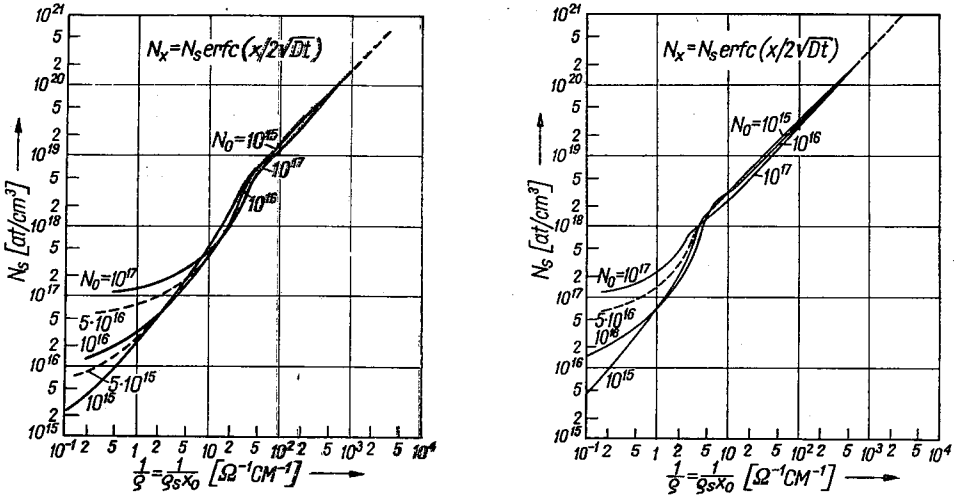
Obecność silnego światła nie jest niezbędna dla wytworzenia różnicy w prędkości osadzania się metalu na krzemie typu *p* i *n*, o ile do obszaru typu *p* dołączy się elektrodę tworzącą z nim kontakt omowy. Elektroda ta, wykonana z takiego samego metalu jak metal osadzany, powinna być zanurzona w roztworze wraz z krzemem.

Metoda ta daje bardzo ostrą granicę przejścia *p-n*. Pozwala ona także na wykrycie gwałtownych zmian w koncentracji domieszek w kryształach krzemu.

Opisane wyżej metody barwienia w połączeniu z pomiarami mikroskopowymi pozwalają na uzyskanie stosunkowo dużych dokładności w określeniu głębokości przejścia p-n. Głębokości przejść rzędu kilkunastu-kilkudziesięciu mikronów mogą być wyznaczone z błędem nie przekraczającym wartości $\pm 5\%$, zaś głębokości przejść rzędu pojedynczych mikronów — z błędem mniejszym niż $\pm 10\%$.

1.6. Pomiar koncentracji powierzchniowej

Opiszemy po krótce jedną spośród znanych metod pomiaru koncentracji powierzchniowej. Metoda ta polega na pomiarze przewodności warstwy uzyskanej wskutek dyfuzji domieszek w półprzewodnik o przeciwnym w stosunku do charakteru dyfundowanych domieszek typie przewodnictwa.

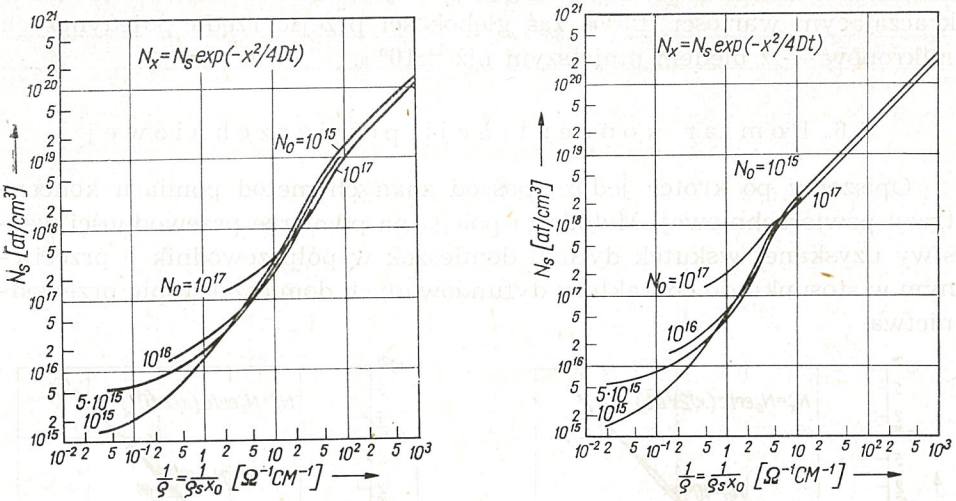


Rys. 13 i 14. Wykresy do obliczania koncentracji powierzchniowych wdyfundowanych warstw dla rozkładu domieszek wg funkcji erfc (λ)
Rys. 13 dla warstw typu n, rys. 14 — dla warstw typu p

Uzyskane w wyniku dyfuzji przejście p-n działa jako izolująca płaszczyzna równoległa do powierzchni próbki. O ile znany jest rozkład domieszek, przewodność wdyfundowanej warstwy oraz głębokość przejścia p-n X_0, wówczas, korzystając z krzywych opracowanych przez Backenstossa (rysunki 13—16), można znaleźć wartość koncentracji powierzchniowej N_s. Występującą na rysunkach 13—16 oporność ϱ_s mierzy się najczęściej przy użyciu sondy 4-punktowej. Jeżeli różnicę potencjału między ostrzami pomiarowymi oznaczamy literą U, zaś prąd płynący przez próbkę literą I, wówczas oporność warstwy wyrazi się jako

$$\varrho_s = \frac{\pi}{\ln 2} \frac{U}{I} \approx 4{,}53 \frac{U}{I}.$$

W przypadku gdy rozkład domieszek nie jest znany, należy zastosować metodę podaną przez Fullera i Ditzenbergera, polegającą na stopniowym zeszlifowywaniu wdyfundowanej warstwy i pomiarach jej przewodności.



Rys. 15 i 16. Wykresy do obliczania koncentracji powierzchniowych wdyfundowanych warstw dla gausowskiego rozkładu domieszek
Rys. 15 — dla warstw typu *n*, rys. 16 — dla warstw typu *p*

2. TECHNOLOGIA DYFUZJI

2.1. Obróbka powierzchni

Sprawa dokładnej obróbki powierzchni krzemu stanowi bardzo ważne zagadnienie w technologii dyfuzji. Wszelkiego rodzaju defekty powierzchni w postaci lokalnych pęknięć, rys i zadrapań powodują w efekcie nieregularności w głębokości dyfuzji, a w konsekwencji uniemożliwiają uzyskanie płaskich przejść *p-n*.

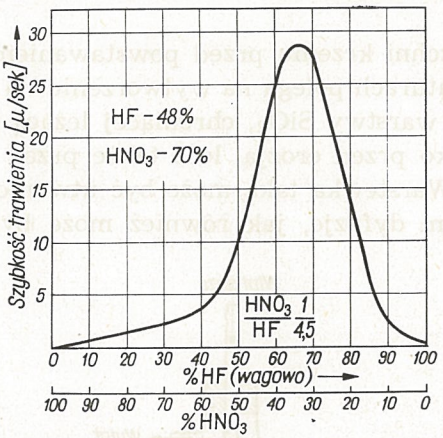
Przygotowanie płytek do dyfuzji polega na bardzo dokładnym oszlifowaniu ich przy pomocy takich materiałów szlifierskich, jak proszek diamentowy, karborund lub tlenek glinu, a następnie odtłuszczeniu powierzchni i wytrawieniu.

Spośród różnych mieszanek trawiących używanych do trawienia krzemu, o składzie dobieranym w zależności od żądanej szybkości trawienia i wymagań dotyczących stanu powierzchni, najczęściej stosowane są mieszaniny kwasu azotowego i fluorowodoru.

Rysunek 17 przedstawia wykres szybkości trawienia krzemu w mieszaninie HNO_3 + HF w zależności od jej składu. Jak widać, największą szybkość trawienia posiada mieszanka, w której stosunek wagowy HNO_3 do HF wynosi w przybliżeniu 1 : 4,5. Duża szybkość trawienia jest jednak bardzo niewygodna, ze względu na trudności dokładnej kontroli

w przypadku trawienia płytki do pożądanej grubości, a ponadto nie daje zadowalającej gładkości powierzchni. Z tego względu najczęściej używaną mieszanką do trawienia krzemu jest mieszanka składająca się z 5 części (objętościowo) skoncentrowanego HNO_3 i 1 części HF.

Gładkość powierzchni, jaką otrzymuje się po trawieniu chemicznym, nie zawsze jest wystarczająca do wykonywania niektórych elementów półprzewodnikowych. Dotyczy to szczególnie tranzystorów krzemowych wielkiej częstotliwości, w których grubość bazy wynosząca zazwyczaj $2 \div 3\,\mu$ narzuca szczególnie ostre warunki na gładkość powierzchni, ponieważ nierówność frontu dyfuzji może doprowadzić do miejscowego po-



Rys. 17. Zależność szybkości trawienia krzemu od składu mieszanki trawiącej HF-HNO_3

łączenia się obszarów emitera i kolektora. Oprócz tego trawienie chemiczne nie zapewnia wystarczająco płaskiej powierzchni na całej płytce z tego względu, że szybkość trawienia krzemu na brzegach jest zazwyczaj większa niż w środku płytki. Otrzymane po takim trawieniu płytki o poduszkowatym kształcie uniemożliwiają późniejsze zastosowanie do wykonywania tranzystorów dyfuzyjnych techniki fotolitograficznej. Z powyższych względów obróbka powierzchni krzemu w przypadkach, w których zależy nam szczególnie na gładkiej i płaskiej powierzchni, polega na dokładnym oszlifowaniu płytek krzemowych, a następnie na wypolerowaniu ich przy pomocy metod i materiałów stosowanych w optyce.

Przed przystąpieniem do procesu dyfuzji wypolerowane płytki należy bardzo dokładnie odtłuścić.

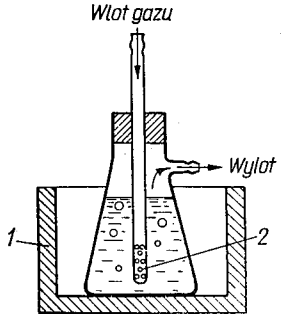
2.2. Ochrona powierzchni płytki

Podczas dyfuzji niektórych domieszek, takich jak np. gal, na powierzchni krzemu pojawiają się liczne drobne zagłębienia naruszające strukturę powierzchni i psujące jej właściwości. Poważne uszkodzenia

powierzchni powstają także w przypadku grzania krzemu w temperaturach wyższych od 1000 °C w atmosferze takich gazów, jak np. suchy wodór lub azot, powszechnie stosowanych w bardzo popularnej metodzie dyfuzji, wykorzystującej strumień gazu jako czynnik niosący domieszkę.

Defekty powierzchni powstające podczas grzania płytek krzemowych w wymienionych gazach są spowodowane silną erozją powierzchni krzemu. Erozja ta nie występuje natomiast nawet w temperaturze 1350 °C, o ile gazem nośnym jest O_2, CO lub inny gaz utleniający. Erozja powierzchni nie występuje również podczas grzania płytek w H_2 lub N_2, jeśli gazy te będą zawierały odpowiednią ilość pary wodnej jako czynnika utleniającego.

Ochrona powierzchni krzemu przed powstawaniem defektów i erozją w wysokich temperaturach polega na wytworzeniu na powierzchni płytek jednolitej i spoistej warstwy SiO_2, chroniącej leżącą pod nią powierzchnię krzemu nie tylko przed erozją, lecz także przed parowaniem i dalszym utlenianiem. Warstewka taka może być utworzona zarówno w procesie poprzedzającym dyfuzję, jak również może być hodowana jednocześnie wraz z procesem dyfuzji, jeśli jest to dyfuzja w gazach utleniających takich, jak np. O_2 lub CO bądź dyfuzja w H_2 lub N_2 zawierających parę wodną.



Rys. 18. Urządzenie do nawilżania gazu nośnego
1 — termostat, *2* — otwory Ø 0,5 mm

Zawartość pary wodnej musi być kontrolowana, ponieważ zarówno jej niedobór, jak i nadmiar nie sprzyjają tworzeniu się jednolitej i spoistej warstwy SiO_2. Najprostszy sposób kontrolowania zawartości pary wodnej w gazie polega na przepuszczeniu go przez wodę podgrzaną do określonej temperatury (zob. rys. 18).

Wzrost temperatury wody powoduje wzrost zawartości pary wodnej w gazie, a w konsekwencji wzrost grubości warstwy SiO_2. C. J. Frosch i L. Derick podają, że jednolitą warstwę SiO_2 można uzyskać w zakresie temperatur wody od 30 do 80 °C. Ograniczenie to wynika z dwóch zjawisk.

1) Zawartość pary wodnej w gazie przepuszczonym przez wodę o temperaturze niższej od 30 °C jest niewystarczająca do wytworzenia jednolitej warstwy tlenków na powierzchni krzemu.

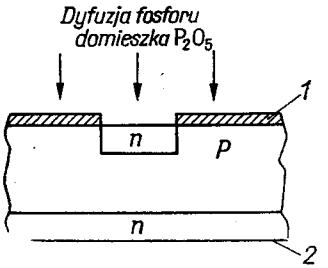
2) Zbyt duża zawartość pary wodnej, jaką otrzymuje się przy podgrzewaniu wody powyżej 80 °C, daje niepożądane efekty objawiając się w postaci mikroskopijnych wgłębień na utlenionej powierzchni krzemu. Powstawanie tych wgłębień może być wytłumaczone niemożliwością utworzenia jednolitej i ciągłej warstwy SiO_2 w obecności tak dużej ilości pary wodnej, ze względu na zjawisko częściowej rozpuszczalności SiO_2 w wodzie przy bardzo wysokich temperaturach.

Grubość warstwy tlenków utworzonej na powierzchni krzemu dzięki obecności pary wodnej w gazie zależy od temperatury krzemu, czasu grzania oraz składu gazu. Dla przykładu można podać, że grubość warstwy SiO_2 utworzonej w czasie $t = 1$ h w wilgotnym azocie (temperatura wody 30 °C) przy podgrzaniu krzemu do temperatury 1200 °C wynosi około 2500 Å.

2.3. Maskujące działanie warstwy tlenków

Obecność odpowiednio grubej warstwy tlenków na powierzchni krzemu powoduje w przypadku dyfuzji niektórych donorów i akceptorów spadek ich koncentracji powierzchniowej nawet o kilka rzędów wielkości w stosunku do koncentracji tych samych domieszek na powierzchni nie utlenionej.

Zjawisko to ma ogromne znaczenie ze względu na możliwość dyfundowania domieszek w pewnych, z góry określonych miejscach uzyskanych przez usunięcie z nich warstewki SiO_2 (rys. 19).



Rys. 19. Zlokalizowanie dyfuzji fosforu w wyniku maskującego działania warstwy SiO_2

1 — warstewka SiO_2, *2* — powierzchnia nie utleniona

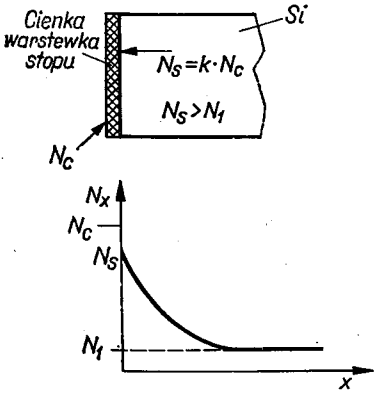
Drugą korzyścią wynikającą z omawianego zjawiska jest możliwość uzyskania dowolnej koncentracji powierzchniowej w przypadku dyfuzji w gazie, ponieważ efekt maskowania przez warstwę tlenków zależy, między innymi, od temperatury krzemu oraz od rodzaju i składu gazu.

Wpływ rodzaju gazu na maskujące działanie tlenków nie jest dotychczas znany. Sugeruje się jednak pewne zmiany właściwości cząstek do-

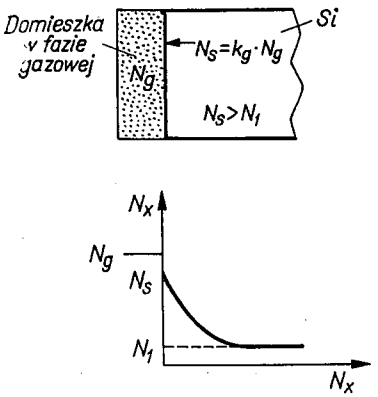
mieszki w różnych gazach zmieniających ich aktywność w stosunku do SiO_2 lub też roztwarzalność w warstwie SiO_2. Należy zaznaczyć, że maskujące działanie warstwy tlenków na dyfuzję domieszek ma charakter selektywny, tzn. odnosi się tylko do dyfuzji niektórych donorów i akceptorów. Selektywność tego zjawiska może być wykorzystana do otrzymania korzystnych, ze względu na geometrię i właściwości, struktur typu n-p-n i p-n-p.

2.4. Klasyfikacja źródeł domieszek

Niezbędnym warunkiem przeprowadzenia dyfuzji domieszek do wnętrza półprzewodnika jest wytworzenie na jego powierzchni źródła domieszek o koncentracji powierzchniowej N_s, większej od koncentracji tego rodzaju domieszek wewnątrz półprzewodnika. Istnieje kilka zasadniczych metod wytworzenia takich źródeł, różniących się między sobą zarówno pod względem wielkości uzyskiwanych koncentracji powierzchniowych N_s, jak również pod względem zachowania się tych koncentracji w czasie trwania procesu dyfuzji.



Rys. 20. Dyfuzja domieszek z warstwy stopu



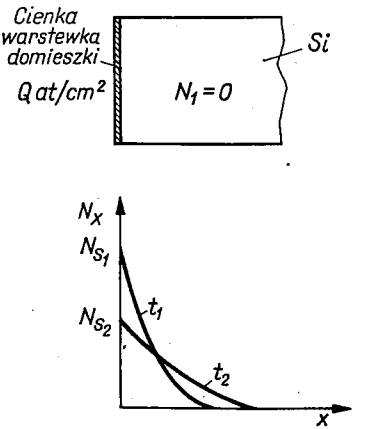
Rys. 21. Dyfuzja domieszek z fazy gazowej

W praktycznych zastosowaniach dyfuzji do celów elektroniki półprzewodnikowej spotykamy się najczęściej z dyfuzją ze źródła o stałej koncentracji powierzchniowej. Przykładem takiego źródła może być cienka warstewka stopu utworzona na powierzchni krzemu dzięki stopieniu z nim materiału stanowiącego żądaną domieszkę (rys. 20). Ponieważ podczas dyfuzji stop ten będzie znajdował się w fazie ciekłej, koncentracja powierzchniowa N_s dla takiego przypadku będzie równa koncentracji domieszek, jaka istnieje na powierzchni krzemu graniczącej z cieczą, i wyrazi się jako:

$$N_s = Nk_c,$$

gdzie k jest stałą segregacji danej domieszki w krzemie, zaś N_c jest koncentracją atomów tej domieszki w utworzonym na powierzchni krzemu stopie.

Ze względu na to, że ilość atomów domieszki, które dyfundują do wnętrza, jest na ogół znikoma w porównaniu z ilością atomów pozostających w cieczy, można założyć z dostateczną dokładnością, że źródło takie daje stałą koncentrację powierzchniową. Podobne warunki można uzyskać również w przypadku źródła domieszek powstającego na po-



Rys. 22. Przykład źródła o ograniczonej ilości domieszek

wierzchni krzemu w wyniku reakcji chemicznej między krzemem i związkami chemicznymi, w postaci których domieszki znajdują się w styczności z krzemem. Jeśli np. na powierzchnię krzemu naniesiemy w odpowiednio wysokiej temperaturze pięciotlenek fosforu, wówczas w wyniku reakcji chemicznej powstanie na powierzchni krzemu cienka warstewka cieczy zawierająca P_2O_5, SiO_2 oraz wolne atomy fosforu.

Reakcja ma przebieg następujący:

$$2P_2O_5 + 5Si \rightleftarrows 4P + 5SiO_2.$$

Reakcja ta jest odwracalna i w stanie równowagi określonym dla danej temperatury dyfuzji T w utworzonej warstewce cieczy znajduje się zawsze taka sama koncentracja wolnych atomów fosforu. Koncentracja powierzchniowa N_s będzie w tym przypadku określona podobnie jak w przypadku dyfuzji z warstwy stopu, a mianowicie:

$$N_s = kN_c,$$

gdzie k jest współczynnikiem segregacji fosforu w krzemie w temperaturze T, zaś N_c jest koncentracją wolnych atomów fosforu w cieczy.

Źródła domieszek powstające w wyniku reakcji chemicznych między krzemem i materiałem domieszki mogą być również uważane za źródła dające stałą koncentrację powierzchniową N_s, pod warunkiem, że szyb-

kość reakcji jest dostatecznie duża, aby skompensować ubytek koncentracji domieszek w cieczy spowodowany ich dyfuzją do wnętrza krzemu. Dodatkowym warunkiem jest oczywiście obecność dostatecznej ilości czynnika reagującego.

Należy zaznaczyć, że metoda otrzymywania na powierzchni krzemu źródła domieszek przy pomocy reakcji chemicznych jest w chwili obecnej najbardziej popularną i najchętniej stosowaną ze względu na stosunkową łatwość otrzymywania żądanych koncentracji powierzchniowych, dużej prostocie procesu i wielu innym zaletom, które omówimy obszerniej w dalszej części artykułu.

Inny sposób, pozwalający na uzyskanie na powierzchni krzemu źródła domieszek dla celów dyfuzji, polega na umieszczeniu krzemu w parach materiału domieszki, którą chcemy wdyfundować (rys. 21). W wyniku adsorpcji domieszek na powierzchni krzemu i ich dyfuzji do jego wnętrza wytwarza się między parą i powierzchnią krzemu pewien stan równowagi reprezentowany przez koncentrację równowagi na powierzchni krzemu — N_s. Jest rzeczą oczywistą, że ten stan równowagi nie zostaje osiągnięty natychmiast. Jednakże jeżeli zostanie on osiągnięty w czasie znacznie krótszym od czasu trwania dyfuzji, wówczas możemy uważać, że dyfuzja odbywa się ze źródła o stałej koncentracji domieszek N_s.

W praktycznych zastosowaniach dyfuzji do celów techniki półprzewodnikowej można się spotkać również z przypadkiem dyfuzji ze źródła o ograniczonej ilości domieszek (rys. 22). Ten rodzaj źródła uzyskuje się przez nałożenie na powierzchnię krzemu cienkiej warstwy domieszek w ilości Q atomów na 1 cm^2 powierzchni krzemu. Jeżeli warstwa ta zostanie całkowicie wdyfundowana do wnętrza krzemu, wówczas koncentracja domieszek N_x w punktach leżących w odległości x od powierzchni wyrazi się jako:

$$N_x = Q(\pi Dt)^{-1/2} \exp(-x^2/4Dt).$$

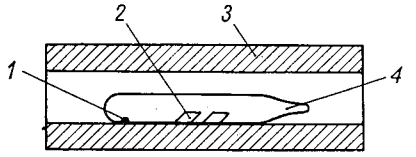
gdzie D jest stałą dyfuzji danej domieszki w krzemie, zaś t czasem trwania procesu dyfuzji.

2.5. Aparatura do dyfuzji

Istnieją dwie zasadnicze metody przeprowadzania procesu dyfuzji. Pierwsza z nich polega na dyfuzji domieszek w próżni, druga w strumieniu gazu przy ciśnieniu atmosferycznym.

Metoda pierwsza może mieć dwa warianty. W pierwszym z nich (rys. 23) dyfuzja odbywa się w odpompowanej do ciśnienia mniejszego niż 10^{-3} Tr zamkniętej ampułce kwarcowej, zawierającej płytki krzemowe oraz niewielką ilość materiału domieszki. Ampułkę tę, dla przeprowadzenia procesu dyfuzji umieszcza się w piecu umożliwiającym podgrzanie jej do wymaganej temperatury.

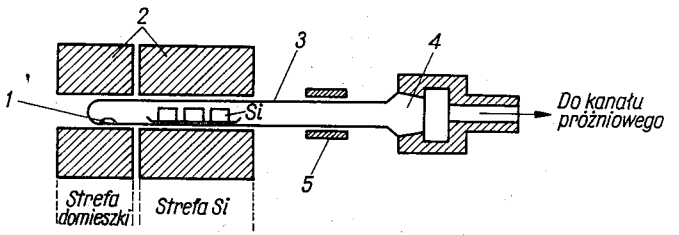
Drugi wariant metody dyfuzji w próżni jest przedstawiony na rys. 24. Rura kwarcowa zawierająca materiał domieszki oraz płytki krzemowe jest odpompowywana przez cały czas trwania procesu dyfuzji. Niezbędna



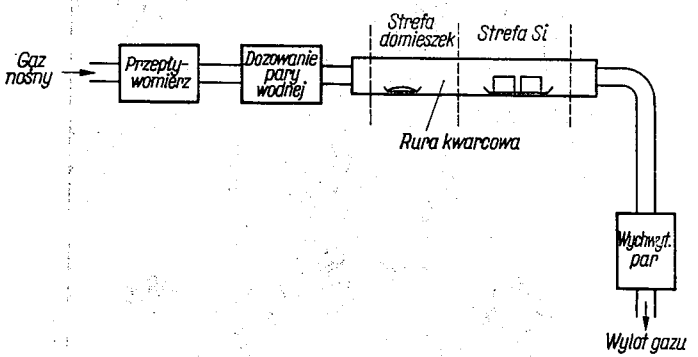
Rys. 23. Dyfuzja w odpompowanej ampułce kwarcowej
1 — domieszka, *2* — płytki Si, *3* — piec, *4* — zalutowana rurka kwarcowa

dla potrzeb dyfuzji ilość par domieszki jest uzyskiwana dzięki podgrzewaniu jej w oddzielnej strefie grzejnej.

Oba opisane wyżej warianty metody dyfuzji w próżni są bardziej kłopotliwe w porównaniu z dyfuzją domieszek w strumieniu gazu. Apa-



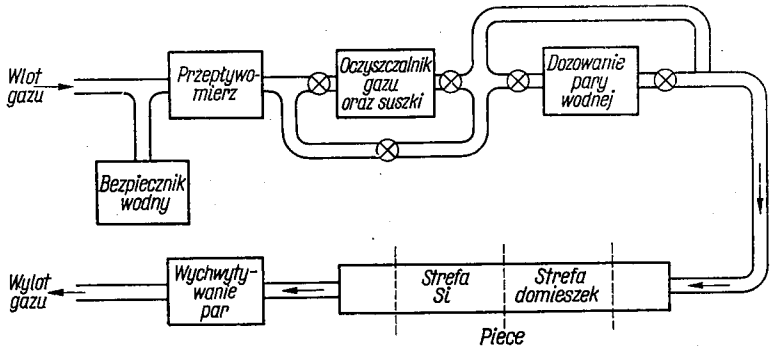
Rys. 24. Dyfuzja w próżniowym układzie dynamicznym
1 — domieszka, *2* — piece, *3* — rura kwarcowa, *4* — szlif, *5* — chłodnica



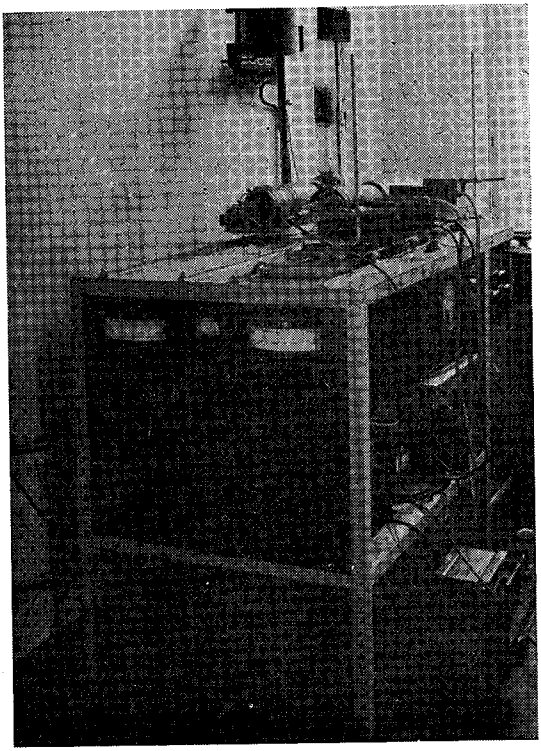
Rys. 25. Zasadniczy schemat aparatury do dyfuzji w gazie

ratura do dyfuzji w gazie (rys. 25 i 26) składa się z rury kwarcowej umieszczonej w piecu wyposażonym w dwie lub trzy strefy grzejne, urządzenia do regulacji zawartości pary wodnej w strumieniu gazu oraz miernika szybkości jego przepływu. Ze względu na to, że strumień gazu jest tu czynnikiem przenoszącym domieszkę z oddzielnej strefy grzejnej w bezpośrednie sąsiedztwo płytek krzemowych, gaz ten przyjęto nazywać gazem nośnym.

Na rysunku 27 przedstawiono fragment aparatury do dyfuzji zbudowanej w Zakładzie Elektroniki IPPT PAN. Aparatura ta składa się z dwóch części, mianowicie części próżniowej, umożliwiającej przeprowadzenie dyfuzji w odpompowywanej rurze kwarcowej, oraz części przed-



Rys. 26. Schemat zbudowanego w IPPT PAN urządzenia do dyfuzji w gazie



Rys. 27. Fragment zbudowanej w IPPT PAN aparatury do dyfuzji

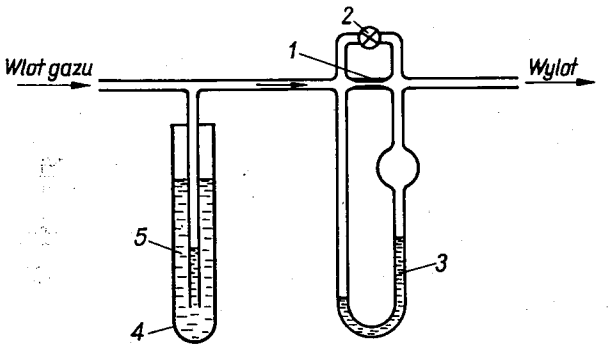
stawionej na pierwszym planie zamieszczonej fotografii — do dyfuzji w strumieniu gazu.

Część próżniowa tej aparatury jest przystosowana jednocześnie do naparowywania cienkich metalicznych warstw na gorące lub zimne płytki

krzemowe, w celu uzyskania niskoomowych kontaktów nieprostujących, potrzebnych do badania właściwości przejść *p-n*, otrzymywanych przy pomocy dyfuzji.

Druga część aparatury, przeznaczona do dyfuzji w gazie, została wyposażona dodatkowo (w porównaniu ze schematem przedstawionym na rys. 25) w bezpiecznik wodny, chroniący przed nadmiernym wzrostem ciśnienia gazu dostarczanego z butli oraz w osadnik dla kondensacji par domieszek, w celu uniemożliwienia im wydostania się na zewnątrz aparatury przez wylot przeznaczony dla gazu nośnego.

Opisywana aparatura ma ponadto wbudowane oczyszczalniki oraz suszki i wymrażarkę, umożliwiające usunięcie tlenu i pary wodnej z takich gazów, jak H_2, N_2 i A oraz ewentualne osuszenie O_2 i CO_2.



Rys. 28. Schemat bezpiecznika wodnego i przepływomierza
1 — kapilara, *2* — kran przelotowy, *3* — ciecz, *4* — rura szklana, *5* — woda

Wprowadzanie kontrolowanej zawartości pary wodnej do gazu nośnego odbywa się w podobny sposób, jak to przedstawiono na rys. 12. Przepływomierz gazu, którego szkic wraz z bezpiecznikiem wodnym pokazano na rys. 28, został wyskalowany dla przepływów od wartości 0,4 do 1,5 l/min.

W zależności od średnicy wewnętrznej użytej do dyfuzji rury kwarcowej tak dobierano wielkość przepływu, aby zachować w niej tę samą ustaloną prędkość gazu nośnego, wynoszącą około 5 cm/sek, co np. w przypadku rury o średnicy wewnętrznej $\phi = 17$ mm odpowiada przepływowi gazu o wielkości 0,68 l/min.

Piece pozwalające na uzyskanie temperatur aż do 1300 °C posiadają spiralne elementy grzejne wykonane z taśmy kanthalowej o szerokości 10 mm i grubości 1 mm. Elementy grzejne są umieszczone w rurach ceramicznych osłoniętych z kolei grubą warstwą materiału stanowiącego izolację cieplną pieca.

Regulację temperatury pieca uzyskiwano przez zmianę napięcia na wejściu transformatora zasilającego 220 V/24 V — 150 A. Pomiaru temperatury dokonywano przy pomocy termopar Pt—Pt Rh 10% i dostosowanych do nich wskaźników temperatury, produkcji zakładów A-3.

Do dyfuzji użyto rur kwarcowych, wykonanych z kwarcu typu transparent, o średnicach wewnętrznych od 15 do 20 mm. Płytki krzemowe umieszczono w łódeczkach kwarcowych o profilu pozwalającym na swobodny dostęp gazu nośnego do powierzchni krzemu. Materiał domieszki znajdował się w łódkach kwarcowych lub alundowych, w zależności od rodzaju użytej domieszki. Rodzaj domieszki decydował tu dlatego iż te domieszki, które są rodnikami lub modyfikatorami szkła (np. B_2O_3), nie mogą być umieszczane w wysokiej temperaturze w łódeczkach kwarcowych.

2.6. Uwagi dotyczące higieny i bezpieczeństwa pracy

Większość materiałów używanych jako domieszki, w przypadku dyfuzji w krzemie, posiada dla organizmu ludzkiego właściwości trujące. Biorąc to pod uwagę należy przy pracy z tymi materiałami zachować jak największą ostrożność, a w szczególności:

1) unikać bezpośredniej styczności materiału domieszki ze skórą,

2) unikać ewentualnego wdychania pyłów domieszki podczas przygotowywania jej do procesu dyfuzji,

3) nie dopuszczać, by w czasie trwania procesu dyfuzji pary domieszek lub ich pyły wydostawały się na zewnątrz aparatury do pomieszczeń zajmowanych przez personel,

4) zapewnić w pomieszczeniu dostatecznie dobrą wentylację, zwłaszcza gdy gaz nośny nie jest wydalany bezpośrednio na zewnątrz pomieszczenia,

5) stosować się do przepisów bezpieczeństwa i higieny pracy dotyczących używania gazów technicznych, a w szczególności wodoru i tlenu.

2.7. Dyfuzja jednej domieszki

W niniejszym rozdziale zajmiemy się technologią dyfuzji niektórych donorów (A_s, Sb, P) i akceptorów (Al, B, Ga) w krzemie z uwzględnieniem wpływu parametrów procesu na wielkość uzyskiwanych koncentracji powierzchniowych oraz efektywność maskowania przez warstwę tlenków.

2.7.1. Dyfuzja arsenu

Arsen jako domieszka donorowa w krzemie bywa najczęściej używany w postaci tlenku (As_2O_3), zwanego potocznie arszenikiem i odznaczającego się silnymi własnościami toksycznymi.

As_2O_3 jest stosowany zarówno do dyfuzji w próżni, jak i w gazie. W przypadku dyfuzji w zamkniętej ampułce kwarcowej, gdzie ilość domieszki jest znikoma (zazwyczaj około 0,1 mg), As_2O_3 znajduje się w tej samej temperaturze co i płytki krzemowe. W przypadku dyfuzji w próżni

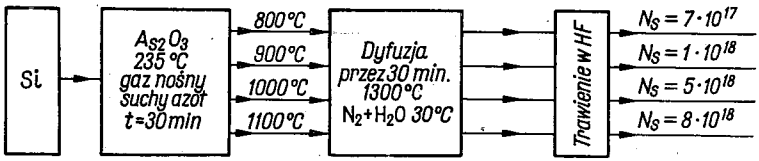
w układzie dynamicznym (rys. 24) lub w przypadku dyfuzji w strumieniu gazu nośnego As_2O_3 jest umieszczony w strefie grzejnej o temperaturze wynoszącej zazwyczaj 220 ÷ 235 °C.

Arsen, podobnie jak i pozostałe donory podlega maskującemu działaniu warstwy tlenków.

Warstwa SiO_2 o grubości 3000 Å powoduje przeszło dziesięciokrotne obniżenie koncentracji powierzchniowej arsenu w porównaniu z koncentracją na nie utlenionej powierzchni krzemu. Maskujące działanie warstwy SiO_2 jest wykorzystywane jako jeden ze sposobów kontrolowania koncentracji powierzchniowej nie tylko w przypadku arsenu, lecz także w przypadku innych domieszek, które podlegają maskującemu działaniu warstwy tlenków krzemu.

Drugi sposób umożliwiający zmianę koncentracji powierzchniowej polega na uprzednim utworzeniu na powierzchni krzemu cienkiej warstwy domieszki, w temperaturze zazwyczaj niższej niż temperatura dyfuzji, a następnie na wdyfundowaniu tej domieszki do wnętrza krzemu.

Na rysunku 29 pokazano zależność koncentracji powierzchniowej arsenu od temperatury nakładania As_2O_3. As_2O_3 umieszczony w temperatu-



Rys. 29. Wpływ temperatury nakładania As_2O_3 na wielkość koncentracji powierzchniowej

rze 235 °C był nakładany na powierzchnię płytek krzemowych w strumieniu suchego azotu przez 30 minut. Po utworzeniu cienkiej warstwy domieszki płytki poddano dyfuzji w wilgotnym azocie (temperatura wody 30 °C) w temperaturze 1300 °C. Czas dyfuzji wynosił we wszystkich przypadkach 30 minut.

Jak widać z rysunku 29, metoda uprzedniego nakładania domieszki pozwala na zmianę koncentracji powierzchniowej w granicach przekraczających jeden rząd wielkości. Metoda ta jest stosowana nie tylko w przypadku dyfuzji arsenu, lecz także w przypadku dyfuzji innych domieszek.

Dyfuzja w strumieniu gazu nośnego pozwala na uzyskanie dodatkowego sposobu regulacji koncentracji powierzchniowej, dzięki jej zależności od rodzaju gazu nośnego i temperatury dyfuzji. Dane obrazujące tę zależność w przypadku dyfuzji arsenu oraz fosforu i antymonu, a także dane dotyczące skuteczności maskowania przez warstwę SiO_2 są zamieszczone w tablicy 3. Tablica 3 podaje również wartości uzyskiwanych koncentracji powierzchniowych w zależności od temperatury dla dyfuzji arsenu, antymonu i fosforu w zamkniętych ampułkach kwarcowych.

Tablica 3

Dane dotyczące dyfuzji donorów w krzemie zaczerpnięte z pracy C. S. Fullera i J. A. Ditzenbergera (dyfuzja w próżni) oraz C. J. Froscha i L. Dericka (dyfuzja w gazie)

Dyfuzja w próżni. Domieszki w postaci tlenków umieszczono wraz z krzemem w zalutowanych ampułkach kwarcowych			Dyfuzja w gazach nośnych					
Donor	Temperatura dyfuzji [°C]	N_s [at/cm³]	Donor	Rodzaj gazu	Rodzaj i temperatura domieszki	Temperatura dyfuzji [°C]	N_s [at/cm³]	N_s pod warstwą SiO_2 o grubości 2500 Å [at/cm³]
Antymon Sb	1095	$1{,}4 \cdot 10^{18}$	Antymon Sb	$N_2 + H_2O$ 30°C	Sb_2O_4 900°C	1300	$6 \cdot 10^{17}$	$3 \cdot 10^{17}$
	1200	$5{,}7 \cdot 10^{19}$		$H_2 + H_2O$ 30°C		1300	$4 \cdot 10^{17}$	$< 10^{15}$
	1250	$9{,}2 \cdot 10^{21}$	Arsen As	$N_2 + H_2O$ 30°C	As_2O_3 235°C	1100	$5 \cdot 10^{18}$	$< 10^{15}$
Arsen As	1200	$5 \cdot 10^{17}$				1200	$7 \cdot 10^{18}$	$3 \cdot 10^{17}$
	1250	$5{,}9 \cdot 10^{17}$		$H_2 + H_2O$ 30°C		1300	$1 \cdot 10^{18}$	$3 \cdot 10^{17}$
	1380	$4{,}5 \cdot 10^{18}$				1200	$1 \cdot 10^{18}$	$< 10^{15}$
Fosfor P	950	$2{,}5 \cdot 10^{21}$		O_2 (suchy)		1300	$2 \cdot 10^{18}$	$3 \cdot 10^{17}$
	1030	$3 \cdot 10^{21}$				1300	$5 \cdot 10^{17}$	$3 \cdot 10^{17}$
	1148	$9 \cdot 10^{21}$	Fosfor P	$N_2 + H_2O$ 30°C	Fosfor czerwony 220°C	1100	$1 \cdot 10^{18}$	$4 \cdot 10^{17}$
						1300	$4 \cdot 10^{17}$	$8 \cdot 10^{16}$
				$H_2 + H_2O$ 30°C		1300	$4 \cdot 10^{17}$	$1 \cdot 10^{17}$
				O_2 (suchy)	P_2O_5 220°C	1100	$7 \cdot 10^{19}$	$3 \cdot 10^{17}$
						1300	$3 \cdot 10^{20}$	$3 \cdot 10^{20}$

2.7.2. *Dyfuzja antymonu*

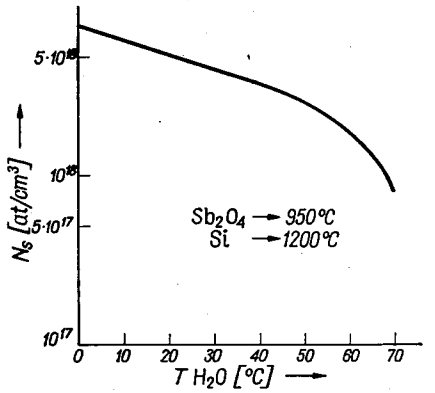
Najczęściej używaną postacią antymonu jako domieszki, szczególnie w przypadku dyfuzji w strumieniu gazu, jest tlenek antymonu Sb_2O_4. W przypadku dyfuzji w próżni używa się zarówno Sb, jak i Sb_2O_4. Antymon, podobnie jak arsen, podlega maskującemu działaniu warstwy SiO_2, jednakże efektywność maskowania jest tu nieco mniejsza niż w przypadku arsenu.

Ze względu na mniejszą lotność niż As_2O_3 tlenek antymonu wymaga stosowania wyższych temperatur w strefie domieszki. Zwykle zawierają się one w granicach 900—1000 °C, w zależności od wymagań dotyczących wielkości koncentracji powierzchniowej oraz sposobu przeprowadzania procesu dyfuzji.

Najczęściej stosowanymi gazami nośnymi dla Sb_2O_4 są azot i wodór.

Zawartość pary wodnej w gazie nośnym (szczególnie w azocie) wpływa w bardzo znacznym stopniu na wielkość uzyskiwanych koncentracji antymonu.

Na rysunku 30 pokazano przebieg koncentracji powierzchniowej antymonu w zależności od temperatury wody w urządzeniu nasycającym gaz



Rys. 30. Wpływ zawartości pary wodnej w azocie na wielkość koncentracji powierzchniowej antymonu

nośny. Jak widać z rysunku, w zakresie temperatur wody od 0 °C do 70 °C można uzyskać prawie dziesięciokrotną zmianę wielkości koncentracji powierzchniowej w przypadku, gdy gazem nośnym jest wilgotny azot. Zmiany koncentracji powierzchniowej w funkcji temperatury H_2O są znacznie mniejsze w przypadku dyfuzji w wodorze, który w tym samym zakresie temperatur wody daje zaledwie trzykrotną zmianę wartości N_s.

W tablicy 3 zamieszczono kilka dalszych przykładów ilustrujących zachowanie się antymonu jako domieszki w krzemie w przypadku dyfuzji w próżni i w gazie.

2.7.3. *Dyfuzja fosforu*

Fosfor jest najszybciej dyfundującym donorem w krzemie. Jego stała dyfuzji posiada identyczną wartość jak stała dyfuzji boru, przewyższając jednocześnie stałą dyfuzji indu i talu. Wielkości uzyskiwanych koncentracji powierzchniowych zależą ogólnie od tych samych czynników i w podobny sposób, jak ma to miejsce w przypadku arsenu i antymonu. Szczególnie duże znaczenie dla dyfuzji fosforu ma rodzaj użytej domieszki.

Najwyższe koncentracje powierzchniowe, bo rzędu 10^{20} at/cm^3, otrzymuje się w przypadku użycia pięciotlenku fosforu P_2O_5. Koncentracje powierzchniowe otrzymywane przy użyciu fosforu czerwonego są mniej więcej tego samego rzędu, co koncentracje uzyskiwane w przypadku arsenu i antymonu.

Przy użyciu pięciotlenku fosforu do dyfuzji w strumieniu gazu należy używać tylko gazów suchych, bez pary wodnej. Mogą to być gazy obojętne lub utleniające, takie jak np. argon lub tlen.

Dla fosforu czerwonego jako domieszki do dyfuzji w gazie nośnym stosuje się z kolei gazy obojętne lub redukujące.

Jak wynika z danych zamieszczonych w tablicy 3, fosfor, podobnie jak arsen i antymon, jest częściowo maskowany przez warstwę SiO_2. Dzięki tej właściwości oraz dzięki łatwości uzyskiwania zarówno dużych, jak i małych koncentracji powierzchniowych jest on bardzo chętnie stosowany nie tylko w technologii diod dyfuzyjnych, lecz także w technologii tranzystorów dyfuzyjnych. W tranzystorach typu *n-p-n* jest on wykorzystywany do wytworzenia obszaru emitera (używa się wtedy P_2O_5, dającego wysokie koncentracje niezbędne dla uzyskania dużej wydajności emitera).

W technologii tranzystorów dyfuzyjnych typu *p-n-p* używa się fosforu do wytworzenia bazy. Ponieważ koncentracja donorów przy powierzchni płytki nie może być większa niż 10^{18} at/cm^3 (w obszarze tym trzeba jeszcze wytworzyć emiter przy pomocy dyfuzji akceptorów), do wykonania bazy tranzystora używa się zazwyczaj fosforu czerwonego, który łatwo spełnia postawiony wyżej warunek.

2.7.4. *Dyfuzja glinu*

Glin (Al) jest stosunkowo najrzadziej używaną domieszką akceptorową w krzemie. Stosuje się go jedynie w przypadku dyfuzji w próżni. Jako domieszka bywa używany w postaci czystego metalu lub też w postaci stopu z krzemem.

Dane dotyczące wyników dyfuzji glinu w krzem są podane w tablicy 4 wraz z danymi dla dyfuzji boru i galu.

Tablica 4

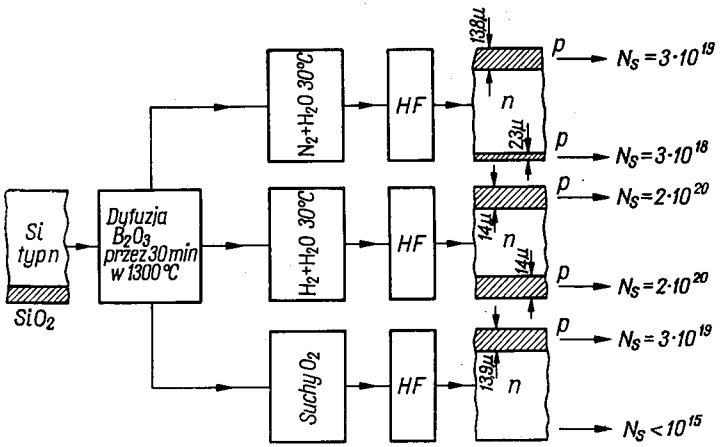
Dane dotyczące dyfuzji akceptorów w krzemie zaczerpnięte z pracy C. S. Fullera i J. A. Ditzenbergera (dyfuzja w próżni) oraz C. J. Froscha i L. Dericka (dyfuzja w gazie)

Dyfuzja w próżni Domieszki w postaci tlenków umieszczono wraz z krzemem w zalutowanych ampułkach kwarcowych			Dyfuzja w gazach nośnych					
Akceptor	Temperatura dyfuzji [°C]	N_s [at/cm^3]	Akceptor	Rodzaj gazu	Rodzaj i temperatura domieszki	Temperatura dyfuzji [°C]	N_s [at/cm^3]	N_s pod warstwą SiO_2 o grubości 2500 Å [at/cm^3]
Bor B	950	$8 \cdot 10^{21}$	Bor	$N_2 + H_2O$ 30°C	B_2O_3 1250°C	1250	$7 \cdot 10^{18}$	$< 10^{15}$
	1020	$8 \cdot 10^{21}$			B_2O_3 1300°C	1300	$3 \cdot 10^{19}$	$3 \cdot 10^{18}$
	1122	10^{22}						
Gl:n Al	1(85	$3 \cdot 10^{16}$		$H_2 + H_2O$ 30°C	B_2O_3 1300°C	1300	$2 \cdot 10^{20}$	$2 \cdot 10^{20}$
	1130	$2 \cdot 10^{16}$						
	1300	$5 \cdot 10^{16}$		O_2 (suchy)	B_2O_3 1300°C	1300	$3 \cdot 10^{19}$	$< 10^{15}$
Gal G	1105	$5{,}8 \cdot 10^{19}$	Gal	$H_2 + H_2O$ 30°C	Ga_2O_3 900°C	1200	$6 \cdot 10^{16}$	$6 \cdot 10^{16}$
	1255	$1{,}5 \cdot 10^{19}$						
	1320	$6{,}7 \cdot 10^{19}$				1300	$1 \cdot 10^{17}$	$1 \cdot 10^{17}$

2.7.5. *Dyfuzja boru*

Charakterystyczną cechą boru jako domieszki akceptorowej w krzemie jest możliwość uzyskiwania przy jego pomocy warstw dyfuzyjnych o koncentracjach przekraczających 10^{20} atomów domieszki/cm^3. Drugą cechą charakterystyczną jest bardzo dobra efektywność maskowania boru przez warstwę SiO_2 w gazach obojętnych i utleniających. Jako domieszka jest używany w postaci chlorku boru (BCl_3) lub tlenku boru (B_2O_3).

Najczęściej stosowaną metodą dyfuzji boru jest dyfuzja w strumieniu gazu nośnego. W przypadku użycia jako domieszki BCl_3 dozowanie jej polega na wprowadzeniu do gazu nośnego w sposób ciągły około kilkudziesięciu cm^3/min BCl_3 (gaz) dla szybkości przepływu gazu nośnego zawierającej się w granicach 1000 cm^3/min. B_2O_3 jako domieszka jest umieszczony w oddzielnej strefie grzejnej lub też w tej samej strefie grzejnej, w której znajduje się krzem. Stosowany zakres temperatur dla B_2O_3 zawiera się w granicach od 800 do 1300 °C.



Rys .31. Wpływ rodzaju gazu nośnego na wielkość koncentracji powierzchniowej boru oraz efektywność maskowania przez warstwę SiO_2

Na wielkość uzyskiwanych koncentracji powierzchniowych boru ma wpływ nie tylko temperatura domieszki, lecz także rodzaj gazu nośnego. Ten ostatni czynnik ma bardzo duże znaczenie w przypadku, gdy interesuje nas maskujące działanie warstwy SiO_2.

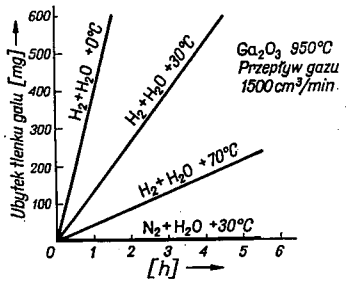
Na rysunku 31 przedstawiono wpływ rodzaju gazu nośnego na wielkość uzyskiwanych koncentracji powierzchniowych oraz skuteczność maskowania przez warstwę SiO_2. Jak widać, efekt maskowania jest najbardziej widoczny w suchym tlenie, natomiast jest on całkowicie niezauważalny w przypadku dyfuzji w wilgotnym wodorze.

Ponieważ koncentracja donorów w poddanym dyfuzji krzemie wynosiła około 10^{15} at/cm^3, więc w przypadku braku zmiany rodzaju przewod-

nictwa pod warstwą SiO_2 należy uważać, że warstwa ta spowodowała obniżenie koncentracji boru o ponad 4 rzędy wielkości w stosunku do powierzchni nie utlenionej.

2.7.6. Dyfuzja galu

Gal, obok boru, jest obecnie najbardziej rozpowszechnioną domieszką typu akceptorowego w krzemie. W przeciwieństwie do boru nie podlega on maskującemu działaniu warstwy SiO_2 podczas dyfuzji, co w wielu przypadkach stanowi zaletę galu i decyduje o jego szerokim użyciu. Najczęściej stosowaną postacią domieszki jest tlenek galu Ga_2O_3, rzadziej — czysty gal. Zakres temperatur domieszki zawiera się w granicach od 900 do 1000 °C.



Rys. 32. Zależność ubytku wagowego Ga_2O_3 od czasu grzania w różnych gazach nośnych

Do dyfuzji galu w strumieniu gazu nośnego używa się prawie wyłącznie wodoru (czasami bywa używany do celów badawczych tlenek węgla CO) i to w odniesieniu zarówno do domieszki w postaci czystego galu, jak i Ga_2O_3. Użycie innych gazów nośnych, takich jak np. argon lub azot jest niecelowe, ponieważ lotność galu lub jego związku w tych gazach jest bardzo mała i nie wystarcza do wytworzenia wymaganej koncentracji domieszek w samym gazie nośnym. O lotności domieszki można sądzić z ubytku jej wagi w funkcji czasu.

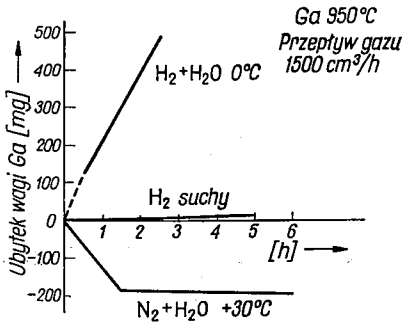
Na rysunku 32 przedstawiono zależność ubytku wagowego Ga_2O_3 od czasu grzania tej domieszki w różnych gazach nośnych. Największy ubytek wagowy ma miejsce — jak widać — w przypadku użycia wodoru o małej zawartości pary wodnej (temperatura H_2O = 0 °C). Ubytek ten, a więc w konsekwencji i lotność domieszki, maleje ze wzrostem zawartości pary wodnej w wodorze, natomiast dla azotu, argonu i helu jest on, praktycznie biorąc, równy zeru, niezależnie od zawartej w nich ilości H_2O lub O_2. Większe ubytki wagowe Ga_2O_3 w gazach redukujących w porównaniu z ubytkami Ga_2O_3 w gazach obojętnych lub utleniających

sugerują zachodzenie redukcji Ga_2O_3 na bardziej lotny związek Ga_2O. Reakcja ta może być przedstawiona przy pomocy następujących równań:

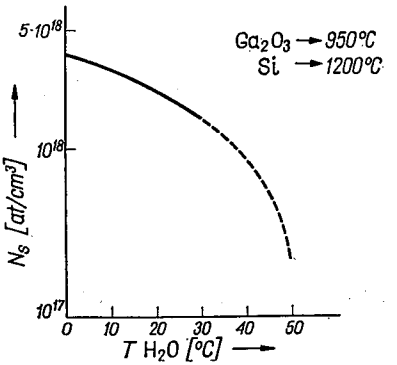
$$Ga_2O_3 + 2H_2 \rightleftarrows Ga_2O + 2H_2O,$$
$$Ga_2O_3 + 2CO \rightleftarrows Ga_2O + 2CO_2.$$

Powyższe równania wyjaśniają jednocześnie przyczynę zmniejszania się ubytku wagowego Ga_2O_3 przy wzroście zawartości pary wodnej w wodorze lub CO_2 w CO jako gazie redukującym.

Zupełnie inaczej przedstawia się zagadnienie lotności czystego galu w różnych gazach nośnych. Ubytek wagowy galu jest bardzo duży w wilgotnym wodorze, natomiast w suchym wodorze jest on minimalny (rys. 33). W przypadku wilgotnego azotu ma miejsce wzrost wagi domieszki wynikający z utlenienia galu na Ga_2O_3. Wpływ rodzaju gazu na lotność Ga_2O_3 i czystego galu znajduje również swoje odwzorowanie w wielkości



Rys. 33. Wpływ czasu grzania i rodzaju gazu nośnego na ubytek wagowy galu



Rys. 34. Wpływ zawartości pary wodnej w wodorze na wielkość koncentracji powierzchniowej galu dyfundowanego ze źródła w postaci Ga_2O_3

koncentracji powierzchniowej. Na rys. 34 przedstawiono wyniki pomiarów koncentracji powierzchniowych dla dyfuzji galu ze źródła w postaci Ga_2O_3. Jak należało się spodziewać na podstawie wpływu wilgotności gazu na lotność domieszki, wzrost zawartości pary wodnej w wodorze powoduje zmniejszanie się koncentracji galu na powierzchni krzemu.

Koncentracje powierzchniowe galu przy dyfuzji w strumieniu takich gazów nośnych, jak A, He, N_2 czy O_2 są bardzo niskie. Są one na pewno niższe od wartości 10^{15} at/cm^3, ponieważ nie powodują one zmiany rodzaju przewodnictwa próbek krzemu typu n, o oporności 5 Ωcm. Koncentracje powierzchniowe galu przy dyfuzji w tych gazach nośnych są zawsze mniejsze od wartości 10^{15} at/cm^3, niezależnie od stopnia nasycenia ich parą wodną. Jak widać, powyższe zjawisko ogranicza nam dyfuzję galu do dwóch gazów nośnych, ponieważ pozostałe gazy techniczne nie pozwalają

na uzyskanie dostatecznie wysokiej wartości koncentracji powierzchniowej.

Selektywność dyfuzji galu w odniesieniu do rodzaju gazu nośnego okazuje się bardzo cenną rzeczą, ponieważ pozwala ona na wprowadzenie znacznych uproszczeń w technologii wielokrotnych przejść *p-n*. Uproszczenia te wynikają z możliwości wykonywania następujących na przemian po sobie dyfuzji antymonu i galu lub arsenu i galu jedynie przez zmianę rodzaju gazu nośnego. Oczywiście musi to być zmiana któregokolwiek z gazów obojętnych lub utleniających na wodór lub ewentualnie tlenek węgla, gdy chcemy prowadzić dyfuzję galu (jednocześnie dyfunduje również i domieszka donorowa) lub zmianę wodoru na gaz obojętny, gdy dyfuzję galu chcemy przerwać.

W następnym rozdziale omówimy zastosowanie tego rodzaju metody do wykonywania tranzystorów dyfuzyjnych typu *n-p-n*. W metodzie tej jest wykorzystana również wspomniana uprzednio właściwość galu, mianowicie brak widocznego wpływu warstwy SiO_2 na wielkość koncentracji powierzchniowej i związana z tym możliwość dyfundowania galu na jednakową głębokość, niezależnie od stopnia utlenienia powierzchni płytki.

2.8. Dyfuzja dwóch domieszek

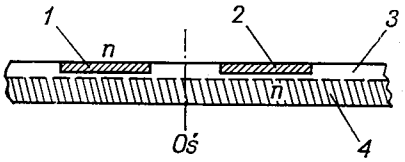
W rozdziale 1.3 omówiono zasady otrzymywania wielokrotnych przejść *p-n* przy pomocy dyfuzji dwóch domieszek, wykorzystując do tego celu metodę sukcesywnej albo jednoczesnej ich dyfuzji. Ponieważ w przypadku dyfuzji dwóch domieszek, donorowej i akceptorowej, można traktować dyfuzję każdej z nich oddzielnie, nie będziemy zajmowali się szerzej tym zagadnieniem, bowiem charakterystyczne cechy dyfuzji większości domieszek w krzemie zostały wyczerpująco podane w rozdziale 2.7. Zajmiemy się natomiast omówieniem szczególnie ciekawej metody dyfuzji w gazie, która w pewnym sensie jest połączeniem metody sukcesywnej oraz jednoczesnej dyfuzji.

Metoda dyfuzji w gazie jest stosowana zazwyczaj przy dyfuzji galu jako akceptora oraz dowolnego w zasadzie donora. Wykorzystuje ona bardzo krytyczną zależność koncentracji powierzchniowej galu od rodzaju gazu nośnego, dzięki czemu możliwe jest rozpoczęcie dyfuzji galu lub jej przerwanie przez samą zmianę gazu nośnego. Zaletę tej metody najlepiej będzie można ocenić na przykładzie techniki wykonywania złącz tranzystora typu *n-p-n*. Strukturę tego typu, przedstawioną na rys. 35, można uzyskać przy pomocy omówionych uprzednio metod lub też znacznie prościej i szybciej w następujący sposób:

1) Płytkę krzemową typu *n*, z usuniętą warstwą SiO_2 z tych obszarów, które mają być w efekcie obszarami typu *n*, poddaje się dyfuzji wolno dyfundującej domieszki donorowej, takiej jak np. arsen lub antymon.

2) W czasie tej dyfuzji zarówno domieszka typu donorowego, jak również domieszka akceptorowa są podgrzewane do takiej temperatury, jaka jest wymagana dla uzyskania właściwej koncentracji powierzchniowej. Dla uniknięcia dyfuzji galu w tym etapie wystarczy, by gaz nośny był gazem obojętnym lub utleniającym.

3) Po osiągnięciu przez warstwę typu n wymaganej głębokości zmienia się gaz nośny na wodór, wskutek czego oprócz dyfuzji domieszki donorowej następuje także dyfuzja galu.



Rys. 35. Struktury tranzystorowe typu n-p-n uzyskane za pomocą dyfuzji przy wykorzystaniu selektywnego maskowania domieszki donorowej przez warstwę SiO_2 *1, 2* — emitery typu n uzyskane dzięki dyfuzji arsenu (antymonu), *3* — baza typu p uzyskana przez dyfuzję galu, *4* — materiał wyjściowy

Ponieważ gal jest domieszką szybko dyfundującą i nie ulega maskującemu działaniu warstwy tlenków, w stosunkowo krótkim czasie wytworzy on warstwę typu p o żądanej głębokości, niezależnej od stopnia utlenienia płytki.

O ile dobierze się odpowiednio warunki procesu, miejsca, z których przed rozpoczęciem dyfuzji usunięto warstewkę SiO_2, pozostaną obszarami typu n ze względu na możliwość uzyskania wysokiej koncentracji powierzchniowej donorów, zaś miejsca leżące pod warstwą SiO_2 zmienią swój typ przewodnictwa na dziurowy, ponieważ warstwa SiO_2 powoduje obniżenie koncentracji dyfundowanych donorów znacznie poniżej koncentracji wdyfundowanych atomów galu.

Należy zwrócić uwagę na fakt, iż w metodzie tej dyfuzja donorów trwa przez cały czas procesu, również i podczas dyfuzji galu, dzięki czemu możliwe jest uniknięcie ewentualnego spadku koncentracji donorów, który mógłby mieć miejsce w przypadku, gdyby dyfuzję galu przeprowadzono bez jednoczesnej dyfuzji donorów. Ma to specjalnie duże znaczenie w przypadkach, w których zależy nam szczególnie na wysokich wartościach koncentracji powierzchniowych.

W tablicy 5 zamieszczono szereg wyników uzyskanych przez Froscha i Dericka dla dyfuzji galu oraz antymonu i arsenu w krzem typu n o oporności 5 Ω cm. Dyfuzja była przeprowadzona w ten sposób, że najpierw dyfundowano antymon lub arsen w wilgotnym (H_2O w temp. 30 °C) azocie, a następnie przez zamianę azotu na wilgotny wodór (H_2O również w temperaturze 30 °C) uzyskiwano dyfuzję galu. Jak widać z zamieszczonych danych, czas grzania w wilgotnym azocie przy dyfuzji antymonu wzrastał progresywnie, podczas gdy czas grzania w wilgotnym wodorze utrzymany był na stałej wartości.

We wszystkich obserwowanych przypadkach grubość warstwy typu n zwiększała się wraz ze wzrostem czasu grzania w azocie, natomiast suma grubości obu warstw, $n + p$, odpowiadająca głębokości dyfuzji galu miała wartość stałą, co jest zgodne z uprzednio podanymi właściwościami dyfuzji galu.

Tablica 5

Wyniki kombinowanej dyfuzji pierwiastka donorowego i akceptorowego

Temperatura dyfuzji [°C]	Czas dyfuzji [min]			Grubość warstwy [μ]		
	N_2	H_2	całkowity	n	p	$n + p$
			Sb_2O_4 900°C			
1300	30	90	120	5,1	13	18,1
1300	60	90	150	5,6	13,4	19
1300	150	90	240	7,1	11,4	18,5
1300	210	90	300	8,4	10,1	18,5
1300	330	90	420	9,7	8,1	17,8
			As_2O_3 220°C			
1200	180	60	240	3,3	1,5	4,8
1200	120	120	240	3,6	4,6	8,2
1350	5	10	15	3,6	6,1	9,7
1350	9	6	15	3,8	3,3	7,1

W przypadku dyfuzji arsenu wykonano pomiary dla różnych czasów grzania w azocie i wodorze z tym jednak, że całkowity czas grzania w azocie i wodorze był dla danej temperatury dyfuzji wielkością stałą.

Otrzymane wyniki pokazują duże zmiany grubości warstwy typu p przy zachowaniu prawie stałej grubości warstwy typu n. Jest rzeczą zrozumiałą, że dane zamieszczone w tablicy 5 nie wyczerpują wszystkich zalet i możliwości ewentualnego zastosowania omawianej metody.

Przy pomocy wielokrotnych, następujących po sobie zmian gazu nośnego z azotu na wodór przy jednoczesnej kontroli koncentracji powierzchniowych dyfundowanych domieszek, dzięki użyciu suchych lub wilgotnych gazów nośnych możliwe jest również uzyskanie wielokrotnych przejść typu n-p-n-p w czasie jednego procesu termicznego, bez potrzeby wyjmowania płytek z aparatury i zmiany rodzaju domieszki.

2.9. Przykłady zastosowań techniki dyfuzyjnej w technologii elementów półprzewodnikowych

W poprzednim rozdziale pokazano sposób otrzymywania dyfuzyjnych przejść n-p-n pozwalających na wykonanie tranzystora o strukturze,

w której uzyskanie kontaktu z obszarem bazy nie przedstawia problemu. Dodatkową korzyścią płynącą z tego typu struktury jest fakt, że kontakt bazy jest wykonywany na powierzchni silnie domieszkowanej, co pozwala na znaczne zredukowanie jego oporności jednostkowej. Sprawa ta ma szczególnie duże znaczenie, zwłaszcza w przypadku kontaktów o bardzo małych powierzchniach stosowanych w tranzystorach wielkiej częstotliwości.

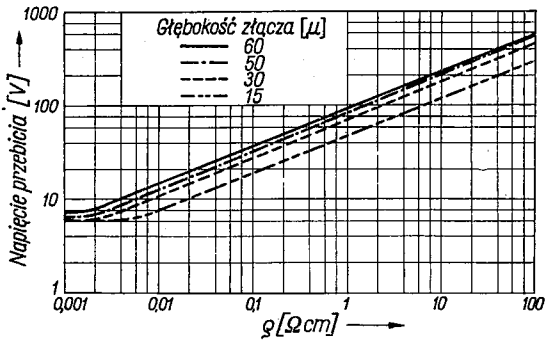
Obecnie podamy kilka innych przykładów zastosowania techniki dyfuzyjnej w technologii elementów półprzewodnikowych, a mianowicie w technologii diod Zennera na krzemie typu p i technologii obwodów monolitycznych.

Technologia diod Zennera wykonywanych przy pomocy dyfuzji nie różni się zasadniczo od technologii diod prostowniczych.

Krzem typu p o oporności właściwej zależnej od wymaganej wartości napięcia Zennera (rys. 36) poddaje się dyfuzji fosforu w celu uzyskania przejścia p-n na głębokości określonej przez dodatkowe wymagania dotyczące wielkości gradientu koncentracji domieszek w samym przejściu p-n.

Dyfuzję fosforu zapewniającą uzyskanie koncentracji powierzchniowej rzędu 10^{20} at/cm^3 przeprowadza się w suchym azocie jako gazie nośnym, zawierającym pary pięciotlenku fosforu.

Zakres stosowanych temperatur dla krzemu zawiera się w granicach 1100—1350 °C, zaś zakres temperatur P_2O_5 mieści się zazwyczaj w granicach 200 ÷ 220 °C.

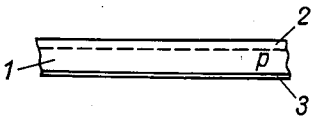


Rys. 36. Zależność napięcia przebicia od wartości oporu właściwego krzemu dla diod wykonanych przy pomocy dyfuzji fosforu

Po przeprowadzeniu dyfuzji fosforu jedną powierzchnię płytki krzemowej szlifuje się w celu usunięcia wdyfundowanej z tej strony warstwy typu n oraz uzyskania pożądanej grubości płytki. Dla zapewnienia niskiej oporności kontaktu z krzemem typu p tak przygotowaną płytkę poddaje się dyfuzji w gazie zawierającym pary BCl_3. Chlorek boru reaguje tylko z odsłoniętą powierzchnią krzemu nie pokrytą tlenkami, dzięki czemu

wolne atomy boru powstaną tylko na powierzchni zeszlifowanej i tylko z tej powierzchni będzie miała miejsce dyfuzja boru. Ponieważ dyfuzja boru ma tu jedynie na celu utworzenie wysoko domieszkowanej warstwy powierzchniowej, może być ona przeprowadzona na mniejszą głębokość niż dyfuzja fosforu.

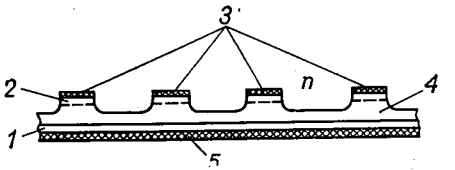
Po dyfuzji boru, gdy płytka krzemowa posiada już strukturę przedstawioną na rys. 37 można przystąpić do wykonania kontaktów. Mogą to



Rys. 37. Struktura płytki krzemowej przeznaczonej do wykonania diod po procesie dyfuzji fosforu i boru

1 — materiał wyjściowy typu *p*, *2* — *n* (dyfuzja fosforu), *3* — *p*+ (po dyfuzji boru)

być np. kontakty niklowe, uzyskane przez nałożenie na powierzchnię krzemu cienkiej warstwy niklu przy użyciu specjalnego roztworu, nie wymagającego zastosowania do tego celu prądu elektrycznego. Kontakty niklowe wykonywane tą metodą, mimo że zawierają domieszki innych pierwiastków chemicznych, mogą być zastosowane zarówno do górnej, jak i do dolnej powierzchni płytki, ponieważ obie te powierzchnie są wysoko domieszkowane i nie grozi nam powstanie na nich niepożądanych przejść *p-n* nawet w przypadku, gdyby materiał kontaktu był domieszką innego typu niż typ przewodnictwa tej powierzchni.



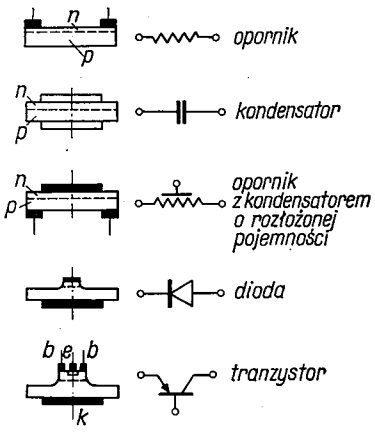
Rys. 38. Diody typu MESA

1 — obszar *p*+, *2* — obszar typu *n*, *3* — kontakty niklowe, *4* — typ *p*, *5* — kontakt niklowy

Przy użyciu metody fotolitograficznej można uzyskać właściwe wymiary przejścia *p-n* wykonywanej diody lub też wielu diod, o ile pozwala na to wymiar płytki krzemowej. Metoda ta polega na pokryciu górnej powierzchni płytki (o ile jest ona tak usytuowana, jak to pokazano na rys. 37) warstwą światłoczułej, a jednocześnie kwasoodpornej emulsji, naświetleniu jej przez odpowiednie szablony, usunięciu emulsji z nienaświetlonych obszarów, a następnie wytrawieniu tych obszarów w mieszaninie HF i HNO_3. Osłonięte przez emulsję obszary utworzą po wytrawieniu płytki szereg płaskich wzgórków, z których każdy tworzy oddzielną diodę (rys. 38). Od tych płaskich wzgórków pochodzi angielska nazwa opisywanej metody, zwanej metodą „mesa".

Jak już podano na wstępie, technika dyfuzji zajmuje również bardzo ważne miejsce w technologii obwodów monolitycznych, stanowiących

prawdziwą rewelację we współczesnej elektronice. Rewelacja ta polega przede wszystkim na niespotykanej dotychczas możliwości zagęszczenia elementów układu, pozwalającej na daleko idącą miniaturyzację sprzętu. Dość wspomnieć, że gęstość elementów w niektórych układach, takich jak np. multiwibratory, odpowiada wartości rzędu kilku $\times 10^5$ elementów/dcm^3. Tak daleko posunięta mikrominiaturyzacja możliwa jest tylko dzięki zastosowaniu w technologii obwodów monolitycznych techniki dyfuzji w połączeniu z techniką fotolitograficzną i techniką naparowywania



Rys. 39. Przykład zastosowania techniki dyfuzji do wykonania części składowych obwodów monolitycznych

cienkich warstw metalicznych. Technika dyfuzji pozwala w omawianym przypadku na wykonanie zarówno aktywnych, jak i passywnych elementów konstruowanego układu elektronicznego. Przykład takiego zastosowania dyfuzji podaje rys. 39, przedstawiający schematycznie wykonany na płytce krzemowej opornik, kondensator wykorzystujący pojemność przejścia *p-n*, opornik z kondensatorem o rozłożonej pojemności oraz tranzystor i diodę.

Na zakończenie należy wspomnieć o jeszcze jednej możliwości zastosowania dyfuzji, mianowicie do domieszkowania pałeczek krzemu przeznaczonego do monokrystalizacji w metodzie beztyglowej. Przeprowadzone w Zakładzie Elektroniki IPPT-PAN próby wykazały, że istnieje możliwość kontrolowanego domieszkowania monokryształu krzemu typu *p* pierwiastkami donorowymi w celu otrzymania z niego, po przetopieniu, monokryształu typu *n* o żądanej oporności.

Autor dziękuje mgrowi A. Kobendzy za sporządzenie nomogramu, przedstawionego na rys. 4.

Instytut Podstawowych Problemów Techniki PAN
Zakład Elektroniki

WYKAZ LITERATURY

1. Aschner J. F., Bittmann C. A.: A Double Diffused Silicon High-Frequency Switching Transistor Produced by Oxide Masking Techniques. — Journ. Electrochem. Soc. vol. 106:1959, s. 415—417.
2. Aschner J. F., Bittmann C. A., Hare W. F. J., Kleimack J. J.: A Silicon Medium-Power Transistor for High-Current High Speed Switching Application. — IRE Transactions on Electron Devices. ED-7:1960, s. 251—262.
3. Backenstoss G.: Evaluation of the Surface Concentration of Diffused Layers in Silicon. — The Bell System Technic. Journ. vol. 37:1958, s. 699—710.
4. Buck T. M., Mc Kim F. S.: Effects of Certain Chemical Treatments and Ambient Atmospheres on Surface Properties of Silicon. — Journ. Electrochem. Soc. vol. 105: 1958, s. 709—714.
5. Cline J. E., Seed R. G.: Diffusion into Silicon from Glassy Layers. — Journ. Electrochem. Soc. vol. 105: 1958, s. 700—701.
6. Crank J.: The Mathematic of Diffusion. Oxford 1956.
7. Daburlos K. E., Patterson H. J.: Oxide Masking. — Bell Laborat. Rec. vol. 38: 1960, s. 417—420.
8. Fast J. D., Van Bueren H. G.: La diffusion dans les métaux. Eindhoven 1957.
9. Frosch C. J., Derick L.: Surface Protection and Selective Masking During Diffusion in Silicon. — Journ. Electrochem. Soc. vol. 104: 1957, s. 547—552.
10. Frosch C. J., Derick L.: Diffusion Control in Silicon by Carrier Gas Composition. — Journ. Electrochem. Soc. vol. 105: 1958, s. 695—699.
11. Fuller C. S., Ditzenberger J. A.: Diffusion of Donor and Acceptor Elements in Silicon. — Journ. Appl. Physics. vol. 27: 1956, s. 544—553.
12. Hannay N. B.: Semiconductors. New York 1959.
13. Henkels H. W., Nowalk T. P.: High-Gain Silicon Transistors. Semiconductors Products. vol. 3: 1960, s. 39—42.
14. Jost W.: Diffusion in Solids, Liquids, Gases. New York 1952.
15. Kurtz A. D., Yee R.: Diffusion of Boron into Silicon. — Journ. Appl. Physics. vol. 31: 1960, s. 303—305.
16. Lamort M. F.: Calculation of Concentration Profiles and Surface Concentration from Sheet-Conductance Measurements of Diffused Layers. — Solid State Electronics. vol. 1: 1960, s. 164—171.
17. Lathrop Jay W., Lee R. E., Phipps C. H.: Semiconductor Networks for Microelectronics. — Electronics. vol. 33: 1960, s. 69—78.
18. Lawrence H., Warner Jr. R. M.: Diffused Junction Depletion Layer Calculation. — The Bell System Technic. Journ. vol. 39: 1960, s. 389—403.
19. Lesk I. A., Gonzalez R. E.: Selective Electrolitic Etching of Germanium and Silicon Junction Transistor Structures. — Journ. Electrochem. Soc. vol. 105: 1958, s. 469—472.
20. Little W. A.: A *P-N-P* High-Frequency Transistors. — Journ. Electrochem. Soc. vol. 107: 1960, s. 789—791.
21. Michelet P., Vareine I.: Application des procédés de photogravure au dépot de contacts métaliques sur les semicunducteurs. — L'Onde électrique. vol. 39: 1959, s. 858—862.
22. Pei Wang: Etching of Germanium and Silicon. — The Sylvania Technologist. vol. 11: 1958, s. 50—58.
23. Prince M. B.: Drift Mobilities in Semiconductors. — II. Silicon. — Physical Rev. vol. 93: 1954, s. 1206.
24. Prugh T. A., Nall J. R., Doctor N. J.: The DOFL Microelectronics Program. — Proceedings of the IRE. vol. 47: 1958, s. 882—894.

25. Robbins H., Schwartz B.: Chemical Eatching of Silicon — Journ. Electrochem. Soc. vol. 106: 1959, s. 505—508.
26. Rohan J. J., Pickering N. E. Kennedy J.: Diffusion of Radioactive Antimony in Silicon. — Journ. Electrochem. Soc. vol. 106: 1959, s. 705—709.
27. Rosenberg A. J.: Separation of the Defect Concentration and Diffusion Coefficient in Diffusion Limited Scaling Reactions. — Journ. Electrochem. Soc. vol. 107: 1960, s. 795—799.
28. Seith W.: Diffusion in Metallen. Berlin 1955.
29. Semiconductor Solid Circuitry. — Electronics. vol. 32: 1959, s. 82.
30. Silverman S. J., Singleton J. B.: Technique for Preserving Lifetime in Diffused Silicon. — Journ. Electrochem. Soc. vol. 105: 1958, s. 591—593.
31. Smits F. M.: Measurement of Sheet Resistivities with Four-Point Probe. — Bell System Technic. Journ. vol. 37: 1958, s. 711—718.
32. Swalin R. A.: Model for Solute Diffusion in Crystals with the Diamond Structure. — Journ. Appl. Physics. vol. 29: 1958, s. 670—674.
33. Theuerer H. C., Kleimack J. J.: Epitaxial Diffused Transistors. — Proceedings of the IRE, vol. 48: 1960, s. 1642—1643.
34. Turner D. R.: Junction Delineation on Silicon in Electrochemical Displacement Solutions. — Journ. Electrochem. Soc. vol. 106: 1959, s. 701—705.
35. Turner D. R.: On the Mechanism of Chemically Etching of Germanium and Silicon. — Journ. Electrochem. Soc. vol. 107: 1960, s. 810—816.
36. Veloric H. S., Prince M. B., Eder M. J.: Avalanche Breakdown Voltage in Silicon Diffused p-n Junctions as a Function of Impurity Gradient. — Journ. Appl. Physics. vol. 27: 1956, s. 895—899.
37. Voprosy mikrominiatiurizacji wojennoj radioelektronnoj apparatury. — Zarubieżna Elektronika. 1960, 8, s. 145—156.
38. Whoriskey P. J.: Two Chemical Stains Marking p-n Junction in Silicon. — Journ. Appl. Physics. vol. 29: 1958, s. 867—868.
39. Warden Waring: Diffusion Concentration Profiles by Analog Computation. — Journ. Electrochem. Soc. vol. 105: 1958, s. 702—705.

T. JANICKI

DIFFUSION OF DONOR AND ACCEPTOR ELEMENTS IN SILICON

Summary

High degree of precision when performing flat and parallel p-n junctions, with simultaneous lack of faults characteristic for alloying techniques has made the diffusion the most progressive method in semiconductor devices manufacture.

The masking effect of silicon dioxide layer upon the diffusion of some impurities allows for its limitation to the selected regions of silicon surface, deprived of the cover of SiO_2 layer, which permits to obtain p-n junctions of required configuration.

The possibility of diffusing impurities to silicon in carrier gases simplifies considerably the apparatus and the dependence of volatility of impurities upon the kind of carrier gas facilitates the control of surface concentration.

In the article, besides the laws governing the diffusion process, the ways of measurement and computations of fundamental parameters were presented. In the technological part, the article gives a description of the apparatus, technology of the diffusion of a series of donor and acceptor elements, and application of the diffusion for production of single and multiple p-n junctions.

T. JANICKI

DIFFUSION DES DONNEURS ET DES ACCEPTEURS DANS SILICIUM

Résumé

Une grande précision de la réalisation des jonctions p-n, plates et paralleles, liée avec une absence des défauts caractéristiques de la technique d'alliage ont fait de la diffusion une technique dominante dans la production des transistors, surtout dans le cas du silicium où les défauts de la technique de l'alliage se présentent plus clairement, que dans le cas de germanium. L'effet de l'action de protection sur la diffusion des quelques impuretés par une couche de SiO_2 permet de la limiter à l'étendue choisie sur la surface du silicium, qui n'est pas couverte par cette couche de SiO_2, et d'obtenir à la fin, la jonction p-n de configuration désirée. La posibilité de la diffusion des impuretés dans le silicium en courant gazeux simplifie beaucoup l'équipement, et puis, la relation existante entre l'évoparabilité des impuretés et la nature du courant gazeux facilite la régulation de la concentration superficielle des impuretés.

Dans cet article, à côté des lois déterminant le procédé de la diffusion, on présente également des méthodes pour mesurer les paramètres les plus importants de ce procédé ainsi que des méthodes de calcul.

Dans la partie consacrée à la technologie on décrit l'équipement, la technologie de la diffusion de quelques impuretés et son application à la réalisation des jonctions p-n uniques et multiples.

T. JANICKI

DIFFUSION DER DONATOREN- UND AKZEPTORENELEMENTE IM SILIZIUM

Zusammenfassung

Die grosse Präzision, welche bei der Herstellung von platten und parallelen p-n Übergängen bei gleichzeitigem Ausbleiben (Fehlen) der in der Legierungstechnik auftretenden Mängel charakteristisch ist, machten aus der Diffusion eine hervorragende Methode für die Herstellung von Halbleiterbauelementen (bezw. Apparaten).

Die Erscheinung der schützenden (maskierenden) Wirkung einer Siliziumoxydschicht auf die Diffusion einiger Dotierungselemente, erlaubt diese auf bestimmte Oberflächenabschnitte des Siliziums zu beschränken, die die aufgetragene Schicht des SiO_2 verlieren, was letzten Endes den Übergang p-n der geforderten Konfiguration zulässt.

Die Möglichkeit einer Durchführung des Diffusionsprozesses von Silizium in Trägergasen der Dotierungssubstanzen, vereinfacht bedeutend die Apparatur, und die Abhängigkeit der Verdampfung der Dotierungssubstanzen von der Art des Trägergases erleichtert die Regulierung der Oberflächenkonzentration.

In diesem Artikel wurden neben Gesetzen, welche den Diffusionsprozess leiten, die Messweise und die Anleitung für die Grundberechnungen der Diffusionsparameter angegeben. Im technologischen Teil wurde die Beschreibung der Apparatur eingelegt, die Technologie der Diffusion (die Diffusionstechnologie) der Donatoren- und Akzeptorenelementenreihe, sowie die Anwendung der Diffusion bei der Herstellung einzelner und vielmaliger p-n Übergänge angegeben.

Т. ЯНИЦКИ

ДИФФУЗИЯ ДОНОРНЫХ И АКЦЕПТОРНЫХ ПРИМЕСЕЙ В КРЕМНИИ

Резюме

Благодаря большой точности при изготовлении плоских и параллельных p-n переходов, при одновременном устранении недостатков характерных для сплавной техники, диффузия стала передовым методом изготовления полупроводниковых приборов, в частности в случае кремния, где недостатки сплавной техники выступают более резко, чем в случае германии.

Маскирующее действие окисной плёнки на диффузию некоторых примесей позволяет ограничивать её только до областей лишенных маскирующего покрытия пленкой SiO_2 и, в результате, получить p-n переходы желаемой конфигурации.

Возможность проведения процесса диффузии в токе несущего газа значительно упрощает аппаратуру, а зависимость испаримости примесей от природы несущего газа — облегчает регуляцию поверхностной концентрации примесей.

В статье рядом с законами диффузии приведены методы измерения и получения основных параметров диффузии.

В технологической части помещено описание аппаратуры, приведена технология диффузии основных донорных и акцепторных примесей, а также её применение для изготовления одиночных и многократных p-n переходов.

621.315.592

JERZY BARANOWSKI

Powielanie lawinowe w elementach półprzewodnikowych

Rękopis dostarczono 21.5.1962

Praca niniejsza przedstawia zasadnicze zjawiska fizyczne towarzyszące procesowi powielania lawinowego w elementach półprzewodnikowych i jest adresowana przede wszystkim do stosujących praktycznie takie elementy. W pracy omówiono kolejno powielanie lawinowe w jednolitej próbce półprzewodnika, w złączu *p-n* oraz właściwości tranzystora warstwowego z powielaniem lawinowym w barierze kolektora. Stosunkowo najwięcej uwagi poświęcono powielaniu lawinowemu w tranzystorze, omówiono tu nie tylko bezpośredni wpływ powielania lawinowego na charakterystyki statyczne, lecz również zjawiska towarzyszące, takie jak anomalie charakterystyk oraz zwiększenie częstotliwości granicznej tranzystora. Na zakończenie przedstawione zostały dziedziny możliwych zastosowań technicznych elementów z powielaniem lawinowym.

1. WSTĘP

Zainteresowanie zjawiskiem jonizacji lawinowej w półprzewodnikach datuje się od czasu stwierdzenia, że w wielu praktycznych przypadkach, a zwłaszcza w tranzystorach, właśnie to zjawisko, a nie przebicia typu Zenera, warunkuje w złączach *p-n* ograniczenie maksymalnego napięcia. Obecnie lawinowe elementy półprzewodnikowe znajdują szerokie zastosowanie, przede wszystkim w dziedzinie generacji krótkich impulsów w zakresie nanosekundowym.

Jakkolwiek konstrukcje elementów półprzewodnikowych wykorzystujące celowo powielanie lawinowe znajdują się dopiero w stadium rozwojowym, należy jednak przypuszczać, że stosowanie takich elementów utrwali się.

Praca niniejsza zawiera omówienie podstawowych właściwości jonizacji lawinowej oraz omówienie zakresu występowania tego zjawiska w diodach i tranzystorach.

2. RUCH NOŚNIKÓW PRĄDU POD WPŁYWEM POLA ELEKTRYCZNEGO

W półprzewodniku nie poddanym działaniu żadnego zewnętrznego pola elektrycznego w stanie równowagi termicznej ruch nośników prądu odbywa się w sposób chaotyczny. Wobec licznych zderzeń nośników z ato-

mami siatki krystalicznej drogi poszczególnych nośników prądu tworzą linie łamane, przy tym średnia szybkość nośników pomiędzy kolejnymi zderzeniami jest zdeterminowana temperaturą półprzewodnika.

Wytworzenie w półprzewodniku pola elektrycznego porządkuje ruch nośników powodując powstanie składowej szybkości skierowanej zgodnie z kierunkiem natężenia pola elektrycznego. Średnia szybkość nośników w kierunku pola elektrycznego, zwana szybkością dryftu, powoduje powstanie prądu o gęstości

$$j = q\varrho v_\alpha, \tag{1}$$

gdzie:

q — ładunek elektronu,
ϱ — koncentracja nośników,
v_α — szybkość dryftu.

W przypadku gdy nośnikami prądu są dziury i elektrony, gęstość prądu określa zależność

$$j = q(nu_n + pu_p)K, \tag{2}$$

gdzie:

n, p — koncentracja elektronów i dziur,
u_n, u_p — tzw. ruchliwość elektronów lub dziur, określona jako stosunek szybkości dryftu do wartości natężenia pola elektrycznego,
K — natężenie pola elektrycznego.

Wzór (2) może być traktowany jako prawo Ohma dopóty, dopóki koncentracje nośników i ruchliwości mogą być uważane za wielkości niezależne od natężenia pola elektrycznego. W rzeczywistości przybliżenie takie można zrobić jedynie w zakresie słabych pól elektrycznych, natomiast w zakresie silnych pól elektrycznych przede wszystkim ruchliwości, a także koncentracje nośników są w półprzewodniku funkcjami pola.

Ustaleniu praw przewodzenia prądu w półprzewodnikach poświęcono szereg opracowań teoretycznych i doświadczalnych [1], [2], [3], [4], [5], [6]. Według Gunna [5] w zależności $j(K)$ można wyróżnić kilka charakterystycznych zakresów:

(1) zakres stałych pól, gdzie spełnione jest prawo Ohma, tzn. $u = \text{const}$,
(2) zakres średnich pól i małych zmian ruchliwości nośników, w przybliżeniu według zależności

$$u = u_0(1 + \beta K^2)$$

(dla przewodników o małej koncentracji domieszek $\beta < 0$, przy bardzo dużej gęstości domieszek może być $\beta > 0$),

(3) zakres dużych pól i dużych zmian ruchliwości, w przybliżeniu według zależności

$$u \sim K^{-1/2}$$

(niezależnie od gęstości domieszek); zakres ten przechodzi płynnie w zakres (4), tzw. nasycenia szybkości dryftu, dla którego

$$v_d = u \cdot K = \text{const}.$$

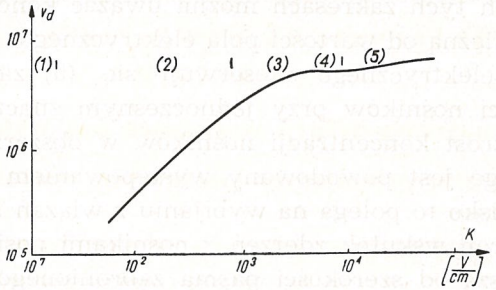
We wszystkich tych zakresach można uważać koncentrację nośników za wielkość niezależną od wartości pola elektrycznego. Przy dalszym podwyższeniu pola elektrycznego odbserwuje się (5) zakres nieznacznego wzrostu szybkości nośników przy jednoczesnym znacznym wzroście ich koncentracji. Wzrost koncentracji nośników w obszarze dużych natężeń pola elektrycznego jest powodowany występowaniem zjawiska jonizacji lawinowej. Zjawisko to polega na wybijaniu z wiązań siatki krystalicznej par dziura-elektron wskutek zderzeń z nośnikami posiadającymi energię kinetyczną większą od szerokości pasma zabronionego. Wytworzone dodatkowe nośniki same następnie mogą jonizować dalsze pary nośników, w ten sposób zjawisko nabiera cech lawiny.

Przyrost ilości nośników, jaki zachodzi pod wpływem jonizacji lawinowej, jest przy małych natężeniach pola ograniczany zjawiskiem rekombinacji. Jednakże szybkość rekombinacji nośników maleje ze wzrostem natężenia pola elektrycznego, natomiast zdolność jonizująca nośników rośnie bardzo szybko ze wzrostem pola. Wskutek tego przy pewnej krytycznej wartości natężenia pola elektrycznego przyrost nośników spowodowany jonizacją lawinową równoważy ubytek nośników spowodowany rekombinacją. Po przekroczeniu krytycznej wartości pola natężenie prądu płynącego w półprzewodniku rośnie do nieskończoności (następuje tzw. przebicie lawinowe półprzewodnika). Dla przypadku jednorodnej nieskończenie długiej próbki półprzewodnika krytyczna wartość natężenia pola elektrycznego wyliczona została w pracy [2], wynosi ona dla Ge około $8 \cdot 10^4 \frac{\text{V}}{\text{cm}}$, a dla Si około $2 \cdot 10^5 \frac{\text{V}}{\text{cm}}$. W przypadku złączy p-n, w których duże natężenia pola elektrycznego występuje w ograniczonym obszarze bariery potencjału, krytyczne natężenie pola elektrycznego osiąga wartości znacznie większe od przytoczonych wyżej. Większe wartości natężenia pola, przy których zachodzi lawinowe przebicie złączy p-n, wynikają ze zmniejszającego się prawdopodobieństwa zderzeń jonizujących przy zmniejszaniu obszaru dryftu.

Na rys. 1 jest pokazany przebieg $v_d(K)$ dla germanu, z zaznaczeniem poszczególnych zakresów zmian ruchliwości.

Należy podkreślić znaczne trudności występujące przy eksperymentalnym wyznaczaniu zależności $j(K)$. Trudności te wynikają m. in. z dużej mocy wydzielanej w próbce podczas pomiarów i związanym z tym wzrostem temperatury. Dlatego często stosuje się tu specjalne metody po-

miarowe przy zdejmowaniu różnych części charakterystyki. W pracy [5] do pomiarów w zakresie średnich i dużych wartości pola elektrycznego zastosowano metodę mostkową przy użyciu impulsów o czasie trwania < 100 ns. Zastosowanie tak krótkich impulsów usunęło praktycznie wpływ zmiany temperatury próbki w czasie pomiarów oraz dodatkowo wpływ modulacji przewodności wskutek wstrzykiwania nośników ze styków pomiarowych. W pracy [4] zastosowano nadzwyczaj czułą metodę pomiaru



Rys. 1. Szybkość dryftu nośników dla germanu typu n w zależności od natężenia pola elektrycznego ($T = 300°$)

ruchliwości nośników w zakresie słabych pól elektrycznych, wykorzystującą zjawisko zmiany absorpcji mikrofal w próbce zasilanej zmiennym polem elektrycznym. Stosując fale o długości 8 mm wykazano m. in., że odchylenia od prawa Ohma występują już przy natężeniu pola $K = 10 \frac{V}{cm}$.

Współczesna teoria stosunkowo dobrze wyjaśnia przebieg $j(K)$, przy tym najlepszą zgodność między wynikami teoretycznymi a praktycznymi uzyskuje się w zakresie (3) i (4).

Dla przykładu w tablicy 1 podano zestawienie wartości szybkości dryftu przy nasyceniu v_{ds}, uzyskane na drodze obliczeń i na drodze pomiarów [3], [5].

Tablica 1

Materiał, rodzaj nośników		Ge elekrony	Ge dziury	Si elektrony
v_{ds} $\left[\frac{cm}{sek}\right]$	teoria [5]	$6.1 \cdot 10^6$	$6.1 \cdot 10^6$	$8.2 \cdot 10^6$
	pomiary [5]	$5.2 \cdot 10^6$	—	—
	pomiary [3]	$6.2 \cdot 10^6$	$6.0 \cdot 10^6$	$8.5 \cdot 10^6$

W innych zakresach zależności $j(K)$ rozbieżności między wynikami teoretycznymi i pomiarowymi są większe, mogą one jednak pochodzić z licznych uproszczeń oraz nieuwzględniania pewnych zjawisk przy wyprowadzaniu wzorów teoretycznych.

Oprócz jonizacji lawinowej w półprzewodnikach może mieć miejsce generacja dodatkowych nośników prądu wskutek przejścia tunelowego

elektronów z pasma walencyjnego do pasma przewodnictwa (efekt Zenera, zwany również emisją polową). Jak wykazują dane eksperymentalne [7], emisja polowa ma miejsce w germanie przy większym natężeniu pola, przekraczającym wartość $2{,}8 \cdot 10^5 \frac{\mathrm{V}}{\mathrm{cm}}$.

3. POWIELANIE LAWINOWE W ZŁĄCZACH *P-N*

3.1. Charakterystyka powielania lawinowego w złączu *p-n*

Z teorii mechanizmu dyfuzji w złączach *p-n* wynika, że przy napięciach zwrotnych przewyższających kilkakrotnie wartość $\frac{kT}{q}$ $\left(\frac{kT}{q} = 0{,}026\ \mathrm{V}\right.$ przy $\left.T = 300°\right)$ prąd płynący przez złącze jest uwarunkowany generacją cieplną nośników mniejszościowych i teoretycznie nie zależy od wielkości napięcia na złączu. W rzeczywistych złączach przy dostatecznie dużych napięciach zwrotnych obserwuje się silny wzrost prądu, prowadzący w krańcowym przypadku do przebicia złącza i jego zniszczenia wskutek nieodwracalnych zmian zachodzących pod wpływem wydzielania się dużych ilości ciepła.

Jeżeli wykluczyć z rozważań wpływ wzrostu temperatury oraz wpływ nieliniowych upływności, wzrost prądu zwrotnego przy wzroście napięcia może być spowodowany zarówno jonizacją lawinową w barierze złącza, jak i zjawiskiem Zenera, bądź obu tymi zjawiskami łącznie.

Przy występowaniu powielania lawinowego łatwo mierzalną wielkością jest tzw. współczynnik powielania nośników prądu (oznaczany zwykle symbolem *M*). Jest to stosunek ilości nośników tego samego znaku opuszczających barierę złącza z jednej strony, do ilości nośników tego znaku wnikających do złącza z drugiej strony, w tej samej jednostce czasu. Współczynnik powielania może być wyrażony przez stosunek odpowiednich prądów nośników tego samego znaku — prądu wychodzącego ze złącza i prądu wejściowego. Pierwszą empiryczną zależność współczynnika powielania od napięcia na złączu ustalił Miller, w postaci

$$M = \frac{I}{I_0} = \frac{1}{1 - \left(\frac{U}{U_B}\right)^n}, \tag{3}$$

gdzie:

I — prąd złącza przy napięciu zwrotnym U,

I_0 — prąd złącza przy $U \approx 0$ (tzw. prąd zerowy złącza),

U_B — napięcie przebicia złącza (przy $U \to U_B$, $I \to \infty$),

n — eksperymentalnie dobrana stała liczbowa, zawarta praktycznie w zakresie od 2 do 6.

Wzór (3) ustalony przez Millera na drodze doświadczalnej przy założeniu jednakowej zdolności jonizacji dziur i elektronów, został później uogólniony w pracach Wuła i Szotowa [8], [9], [10], [11].

Dla identyfikacji zjawiska powodującego przebicie złącza wykonuje się pomiary współczynnika powielania nośników mniejszościowych wprowadzanych zewnętrznie do bariery, przy tym najdogodniejszym sposobem jest naświetlenie odpowiedniego miejsca złącza wąskim strumieniem światła z zakresu widma krótkofalowego lub napromieniowanie cząstkami α. Wytwarzaniu się nośników towarzyszy przyrost prądu zerowego złącza spolaryzowanego zwrotnie. Jeżeli przyrost prądu jest przy tym uzależniony od wielkości napięcia polaryzującego złącze, oznacza to, że w barierze zachodzi zjawisko powielania lawinowego nośników. Przy badaniach tego typu określono, że czas tworzenia się lawiny przy napromieniowaniu impulsowym leży znacznie poniżej przeciętnej zdolności rozdzielczej współczesnych oscylografów (< 20 ns). Dokonanie takiego pomiaru może mieć znaczenie wówczas, gdy zachodzi podejrzenie, że powielanie nośników zachodzi wskutek tworzenia się warstw wielokrotnych p-n-p-n, dla których odpowiednie czasy tworzenia się lawiny są znacznie większe.

Opisana powyżej metoda pozwala ustalić z całą pewnością charakter lawinowy przebicia złącza. Dodatkowych informacji może dostarczyć fakt, że przebicie Zenera dla danego półprzewodnika występuje zawsze przy tej samej wartości natężenia pola elektrycznego K_z (o ile gęstość domieszek nie jest zbyt wysoka). Praktycznie można rozgraniczyć zakres występowania obu zjawisk nie znając wartości K_z przez dokonanie serii pomiarów złączy wykonanych z materiałów o różnych zawartościach domieszek.

3.2. Zależność napięcia przebicia złącza od koncentracji domieszek

Wśród wytwarzanych obecnie złączy dominują dwa zasadnicze typy przejść między obszarami p i n. Pierwszy rodzaj stanowią złącza „skokowe", np. typu stopowego, w których oba obszary są jednorodne, a gęstość domieszek w przejściu między nimi zmienia się w przybliżeniu skokowo. W złączach tego typu jeden z obszarów ma zwykle znacznie większą koncentrację domieszek niż drugi, np. baza w tranzystorach, bariera potencjału wnika wówczas głównie w obszar o mniejszej koncentracji domieszek (np. do bazy). Przyjmując jednowymiarowość złącza oraz zaniedbując pole wytworzone przez nośniki prądu wobec pola wytwarzanego przez zjonizowane atomy domieszek, przebieg natężenia pola elektrycznego w barierze można w tym przypadku określić jako

$$K = K_{max}\left(1 + \frac{x}{d}\right) \qquad 0 \leqslant x \leqslant d, \tag{4}$$

$$K_{max} = \frac{2(U+\Phi)}{d} \approx \frac{2U}{d}, \tag{5}$$

$$d = \left[\frac{2\varepsilon\varepsilon_0(U+\Phi)}{qN}\right]^{1/2} \approx \left(\frac{2\varepsilon\varepsilon_0 U}{qN}\right)^{1/2}, \tag{6}$$

gdzie:

U — napięcie zwrotne na złączu,
Φ — potencjał bariery w warunkach równowagi,
d — szerokość bariery,
N — koncentracja domieszek w bazie,
$\varepsilon\varepsilon_0$ — stała dielektryczna,
x — współrzędna odległości mierzonej od płaszczyzny rozdziału obszarów p i n.

Drugi rodzaj tworzą złącza „stopniowe", np. typu dyfuzyjnego, dla których gęstość domieszek w przejściu między obszarami zmienia się w przybliżeniu liniowo. Rozkład natężenia pola w takich złączach jest kwadratowy, zaś wzory analogiczne do przybliżonych wzorów (4), (5), (6) mają postać

$$K = K_{max}\left[1 - \left(\frac{2x}{d}\right)^2\right] \qquad (x) \leqslant \frac{d}{2}, \tag{7}$$

$$K_{max} \approx \frac{3}{2}\frac{U}{d}, \tag{8}$$

$$d \approx \left(\frac{12\varepsilon\varepsilon_0 U}{qa}\right)^{1/3}, \tag{9}$$

gdzie:

a — gradient koncentracji domieszek w obszarze bariery.

Ze wzorów (5) i (6) oraz (8) i (9) wynika, że w przypadku przebicia typu Zenera (K_z = const) iloczyn napięcia przebicia złącza U_z i koncentrancji N powinien być stały, tzn.

$$U_z \sim \frac{1}{N} \tag{10}$$

(w przypadku złącza „stopniowego" przyjęto $a \cdot d$ = N).

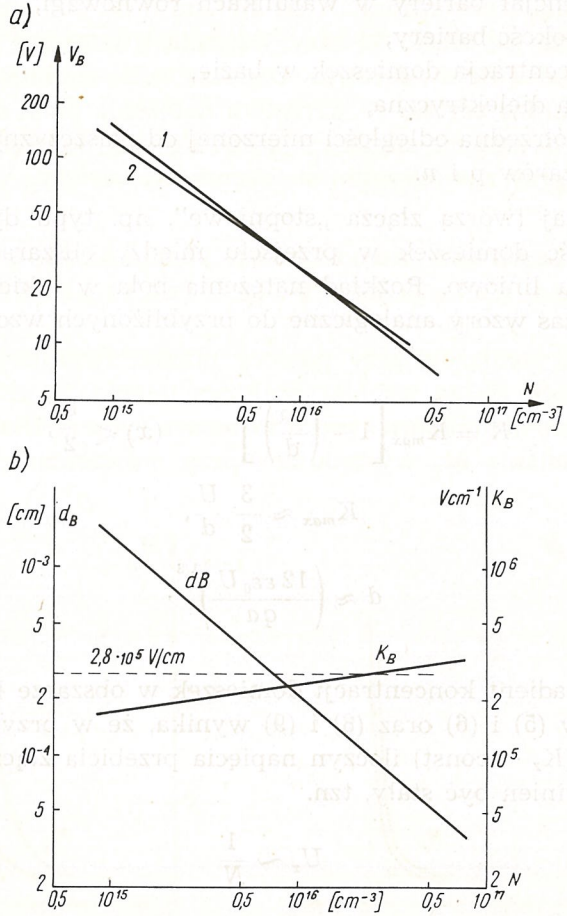
Na rys. 2a przedstawiono zależność napięć przebicia lawinowego U_B w funkcji gęstości domieszek N, uzyskaną dla skokowych złączy germanowych na drodze pomiarowej przez Millera oraz przez Wuła i Szotowa. Na wykresach tych zależność $U_B(N)$ przedstawiona w skali podwójnie logarytmicznej reprezentuje linię prostą, której nachylenie zdecydowanie różni się jednak od wartości -1, jaka wynika z (10) dla przebicia typu Zenera. Z wykresu wynika, że nawet przy znacznych koncentracjach domieszek, zbliżających się do 10^{17} cm^{-3} (co odpowiada opornościom właściwym poniżej 0,1 Ω cm) przebicie złącza jest raczej typu lawinowego.

Z wyników pomiarów określono następującą zależność pomiędzy napięciem przebicia lawinowego U_B a koncentrancją domieszek N

$$U_B = C \cdot N^{-r}, \tag{11}$$

gdzie C i r są stałymi, wynoszącymi (według Wuła i Szotowa) dla skokowych złączy germanowych:

$$C = 2{,}8 \cdot 10^{12} \qquad r = 0{,}69,$$



Rys. 2. a) Napięcie przebicia złącza germanowego w zależności od gęstości domieszek w bazie (złącze skokowe); *1* — wg Millera, *2* — wg Szotowa i Wuła; b) Szerokość bariery oraz maksymalne natężenie pola elektrycznego w złączu germanowym przy przebiciu lawinowym (złącze skokowe)

a dla skokowych złączy krzemowych:

$$C \approx 3 \cdot 10^{13} \qquad r \approx 0{,}75.$$

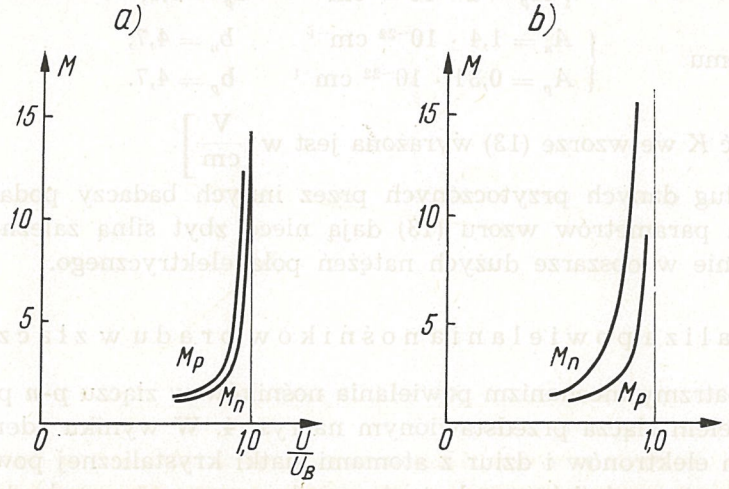
Na rys. 2b przedstawiono wyliczone według wzorów (5), (6), (11) wartości K_{max} i α, jakie odpowiadają warunkom przebicia skokowych złączy germanowych, w funkcji koncentracji domieszek N.

Na wykresie zaznaczono wartość natężenia pola K_z podaną dla przebicia typu Zenera w pracy [7]. Jak widać, pomimo przekroczenia przez natężenie pola zaznaczonego poziomu napięcie przebicia (rys. 2a) przy wartościach $N \to 10^{17}\ \mathrm{cm}^{-3}$ spełnia jeszcze zależność (11). Oznacza to, że w rzeczywistości wartość natężenia pola wywołująca w germanie emisję polową jest większa niż $2{,}8 \cdot 10^5 \frac{\mathrm{V}}{\mathrm{cm}}$.

Przy zmianie koncentracji N od $10^{15}\ \mathrm{cm}^{-3}$ do $10^{17}\ \mathrm{cm}^{-3}$ szerokość bariery zmniejsza się blisko 50 razy, natomiast natężenie pola zmienia się nieznacznie, zaledwie dwukrotnie. Oznacza to, że zdolność jonizacji nośników rośnie bardzo szybko ze wzrostem natężenia pola, dzięki czemu jest możliwe zrównoważenie zmniejszającego się prawdopodobieństwa zderzeń jonizujących przy zmniejszaniu się szerokości bariery.

3.3. Zależność zdolności jonizacji nośników od natężenia pola elektrycznego

Dokonując pomiarów przy pomocy plamki świetlnej stwierdzono, że dla tego samego złącza współczynnik powielania M jest różny w przypadku naświetlenia obszaru typu n i obszaru typu p. Dla germanu współczynnik powielania M_p występujący przy naświetlaniu obszaru typu n



Rys. 3. Przebieg współczynników powielania M_n i M_p w funkcji stosunku $\frac{U}{U_B}$ dla skokowych złączy: a) germanowych, b) krzemowych

(generacja dziur jako nośników mniejszościowych) jest zawsze większy od współczynnika powielania M_n występującego przy naświetlaniu obszaru typu p (generacja elektronów jako nośników mniejszościowych), dla krzemu natomiast zależność jest odwrotna (rys. 3a i b). Różnice między M_p i M_n są widoczne szczególnie w przypadku złączy krzemowych, dla

których ponadto obserwuje się znaczne odstępstwa rzeczywistego przebiegu $M(U)$ w stosunku do zależności (3).

Różnice między M_p i M_n są spowodowane różną zdolnością jonizacji dziur i elektronów. Zdolność ta jest wyrażona zwykle przez ilość zderzeń jonizujących zachodzących przy poruszaniu się nośnika na drodze jednostkowej. Ustalono, że zdolność jonizacji dla dziur (α_p) i zdolność jonizacji dla elektronów (α_n) pozostają w stałym stosunku (podobnie jak to ma miejsce w gazach), tj.

$$\alpha_p = k \cdot \alpha_n, \tag{12}$$

przy tym $k = 2$ dla germanu, dla krzemu natomiast $k \approx \frac{1}{4{,}5}$.

Zdolność jonizacji jest szybko rosnącą funkcją natężenia pola elektrycznego. Wuł i Szotow wykazali [10], [11], że jeżeli przyjąć za podstawę rozważań zależności (11) i (12), to zdolność jonizacji powinna wyrażać się funkcją potęgową pola elektrycznego

$$\alpha = A \cdot K^b. \tag{13}$$

Przeprowadzając szereg pomiarów, ustalili oni następujące wartości na A i b:

dla germanu $\begin{cases} A_n = 10^{-25}\ \mathrm{cm}^{-1} & b_n = 5{,}5, \\ A_p = 2 \cdot 10^{-25}\ \mathrm{cm}^{-1} & b_p = 5{,}5; \end{cases}$

dla krzemu $\begin{cases} A_n = 1{,}4 \cdot 10^{-22}\ \mathrm{cm}^{-1} & b_n = 4{,}7, \\ A_p = 0{,}31 \cdot 10^{-22}\ \mathrm{cm}^{-1} & b_p = 4{,}7. \end{cases}$

[wartość K we wzorze (13) wyrażona jest w $\frac{\mathrm{V}}{\mathrm{cm}}$].

Według danych przytoczonych przez innych badaczy podane wyżej wartości parametrów wzoru (13) dają nieco zbyt silną zależność $\alpha(K)$, szczególnie w obszarze dużych natężeń pola elektrycznego.

3.4. Analiza powielania nośników prądu w złączach p-n

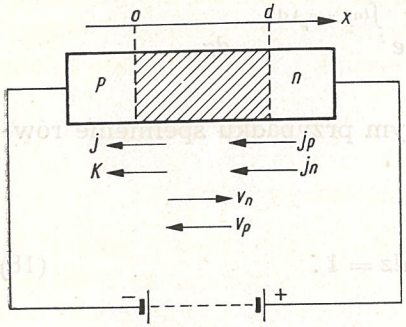
Rozpatrzmy mechanizm powielania nośników w złączu p-n posługując się modelem złącza przedstawionym na rys. 4. W wyniku zderzeń jonizujących elektronów i dziur z atomami siatki krystalicznej powstają dodatkowe pary nośników prądu i odpowiednie przyrosty prądu dziurowego i elektronowego. Posługując się definicją zdolności jonizacji dziur i elektronów można ustalić następujące równania na przyrosty gęstości prądu elektronowego j_n i prądu dziurowego j_p w postaci:

$$\frac{d_{jn}}{dx} = -\frac{d_{jp}}{dx} = \alpha_n \cdot j_n + \alpha_p \cdot j_p, \tag{14}$$

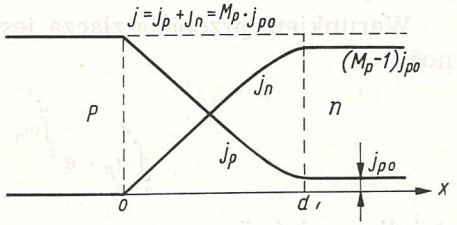
gdzie:

α_n i α_p są poprzez natężenie pola funkcjami zmiennej x.

Jeżeli koncentracja domieszek po stronie n jest znacznie mniejsza niż po stronie p, czynnikiem pierwotnie jonizującym jest prąd dziurowy j_p płynący z obszaru n do obszaru p. Pod wpływem jonizacji atomów siatki krystalicznej, wskutek wybijania z nich dodatkowych par nośników, rośnie prąd całkowity, przy tym elektrony i dziury poruszając się w przeciwnych kierunkach powodują, że wzrost prądu j_p następuje w kierunku $-x$, a wzrost prądu j_n następuje w kierunku $+x$ (rys. 5). Dla przypadku $p \gg n$ ustalamy następuje warunki brzegowe:



Rys. 4. Model złącza p-n z powielaniem lawinowym nośników



Rys. 5. Rozkład prądów wzdłuż złącza przedstawionego na rys. 4

1) $x = 0 \qquad j_n = 0,$

2) $x = d \qquad j_p = j_{p0}.$

Dla każdej wartości x zachodzi

3) $j_n + j_p = j = \text{const}.$

Rozwiązanie układu równań (14) przy przyjętych warunkach brzegowych pozwala obliczyć współczynnik powielania lawinowego nośników, w danym przypadku nośników dziurowych. Ma on postać

$$M_p = \frac{j_p(0)}{j_p(d)} = \frac{j}{j_{p0}} = \frac{e^{\int_0^d (\alpha_p - \alpha_n)\,dx}}{\int_0^d \alpha_n \cdot e^{\int_0^z (\alpha_p - \alpha_n)\,dx} \cdot dz - 1}. \tag{15}$$

Jak widać, warunkiem przebicia złącza ($M_p = \infty$) jest spełnienie równości

$$\int_0^d \alpha_n \cdot e^{\int_0^z (\alpha_p - \alpha_n)\,dx} \cdot dz = 1. \tag{16}$$

Gdy czynnikiem pierwotnie jonizującym są elektrony (gdy $n \gg p$) przechodzące z obszaru p do obszaru n, analogiczny współczynnik powie-

lania elektronów wyraża się następującą zależnością

$$M_n = \frac{j_n(d)}{j_n(0)} = \frac{j}{j_{n0}} = \frac{e^{\int_0^d (\alpha_n - \alpha_p)\,dx}}{1 - \int_0^d \alpha_p \cdot e^{\int_z^d (\alpha_n - \alpha_p)\,dx} \cdot dz}. \tag{17}$$

Warunkiem przebicia złącza jest w danym przypadku spełnienie równości

$$\int_0^d \alpha_p \cdot e^{\int_z^d (\alpha_n - \alpha_p)\,dx} \cdot dz = 1. \tag{18}$$

Jeżeli przyjąć, że

$$\frac{\alpha_p}{\alpha_n} = k = \text{const},$$

wzory (15), (18) znacznie uproszczają się i przyjmują postać

$$M_p = \frac{(k-1)z}{k-z}, \tag{19}$$

$$M_n = \frac{k-1}{k-z}, \tag{20}$$

gdzie:

$$z = e^{(k-1)\int_0^d \alpha_n\,dx}. \tag{21}$$

Przebicie złącza następuje, gdy zachodzi równość

$$z = k.$$

Warunek ten nie zależy od rodzaju nośników, które powodują pierwotną jonizację.

Przeprowadzając całkowanie według (21) przy użyciu wzorów (4), (9) oraz (13), można przedstawić współczynniki powielania M_p i M_n bezpośrednio jako funkcje stosunku napięcia na złączu do napięcia przebicia złącza $\frac{U}{U_B}$. Funkcje te mają postać

$$M_p = \frac{(k-1)k^{\rho}}{k-k^{\varsigma}}, \tag{22}$$

$$M_n = \frac{k-1}{k-k^{\rho}}, \tag{23}$$

przy tym dla złączy skokowych (stopowych)

$$\varrho = \left(\frac{U}{U_B}\right)^{\frac{b+1}{2}}, \tag{24}$$

a dla złączy stopniowych (dyfuzyjnych)

$$\varrho = \left(\frac{U}{U_B}\right)^{\frac{2b+1}{2}}. \tag{25}$$

Jak z powyższych wzorów wynika, współczynniki M_n i M_p zależą od napięcia w sposób różny, co spowodowane jest różną zdolnością jonizacji dziur i elektronów. Dla germanu $k = 2$ i wobec tego $M_p > M_n$, dla krzemu natomiast $k \approx \frac{1}{4{,}5}$ i $M_n > M_p$. Wobec większej różnicy α_n i α_p, w złączach krzemowych obserwuje się większą różnicę pomiędzy M_p i M_n niż w złączach germanowych (rys. 3).

Wzory (22) i (23) są często zbyt skomplikowane dla obliczeń, dlatego niekiedy do przedstawienia tych zależności wykorzystuje się wcześniejszy historycznie wzór Millera (3). Zależność (3) można uzyskać bezpośrednio ze wzorów (22) i (23), kładąc bowiem $\alpha_p = \alpha_n$, otrzymujemy

$$M_p = M_n = M = \frac{1}{1-\varrho}. \tag{26}$$

Powyższy wzór dzięki swojej prostocie jest szeroko stosowany obok dokładniejszych wzorów (22) i (23), przy tym w zależności

$$\varrho = \left(\frac{U}{U_B}\right)^n$$

wartość n ustala się bądź eksperymentalnie, bądź zgodnie ze wzorami (24), (25).

3.5. Wstrzykiwanie lawinowe [5], [13]

Rozpatrując mechanizm powielania lawinowego w złączach p-n, zakładaliśmy dotychczas, że proces ten zachodzi pod wpływem pola elektrycznego pochodzącego od niezrównoważonych ładunków atomów domieszek w obszarze bariery i że wpływ pola elektrycznego wytworzonego przez same nośniki prądu jest pomijalny.

Takie pominięcie staje się jednak niesłuszne przy dużej koncentracji nośników prądu, tj. przy dużych gęstościach prądu. Za krytyczną wartość gęstości prądu, powyżej której pole wytwarzane przez nośniki prądu staje się czynnikiem istotnym, można przyjąć wartość

$$j_0 = qNv_{ds},$$

gdzie:

N — koncentracja domieszek,

v_{ds} — szybkość dryftu nośników w nasyceniu, po osiągnięciu której dalszy wzrost prądu następuje przez zwiększenie się koncentracji nośników.

Krytyczną wartość gęstości prądu osiąga się tym łatwiej, im bardziej dany materiał jest zbliżony do półprzewodnika samoistnego (mała wartość N). Proces powielania lawinowego, jaki może powstać po przekroczeniu gęstości krytycznej prądu, posiada cechy różniące go dość istotnie od procesu powielania przedstawionego uprzednio. Omówimy go pokrótce, stosując uproszczone rozważania i model złącza przedstawiony na rys. 4, w którym strona p jest silnie domieszkowana, a strona n jest zbliżona do półprzewodnika samoistnego. Przy tych założeniach obszar powielania znajduje się po stronie n i równanie Poissona przyjmuje postać

$$\frac{dK}{dx} = \frac{q\varrho}{\varepsilon\varepsilon_0} = \frac{q}{\varepsilon\varepsilon_0}(p - n). \tag{27}$$

Przyjmując dla uproszczenia rozważań, że zdolność jonizacji dziur i elektronów jest jednakowa ($\alpha_p = \alpha_n = \alpha$) oraz jednakowe i stałe są wartości modułów szybkości nasycenia dryftu ($|v_p| = |v_n| = v_{ds}$), po zróżniczkowaniu równania (27) i podstawieniu do niego

$$\frac{dj_p}{dx} = -\frac{dj_n}{dx} = \alpha \cdot j$$

uzyskujemy równanie

$$\frac{d^2K}{dx^2} = \frac{1}{\varepsilon\varepsilon_0 v_{ds}}\left(\frac{dj_p}{dx} - \frac{dj_n}{dx}\right) = \frac{2}{\varepsilon\varepsilon v_{ds}}\alpha \cdot j. \tag{28}$$

Załóżmy, że lewa granica obszaru powielania pokrywa się z $x = 0$, przyjmijmy również zwroty j, K oraz v_{ds} skierowane w stronę $-x$ (złącze spolaryzowane zaporowo). W takim przypadku w miejscu $x = 0$ prąd elektronowy jest równy zeru i w ogólnym bilansie nośników w kierunku $+x$ ilość dziur zmniejsza się, a ilość elektronów powiększa się, a zatem z równania (27) wynika, że wartość bezwzględna K maleje i dla pewnej wartości $x = d_0$, dla której $j_n = j_p$, staje się równa zeru. Stosując tak ustalone warunki brzegowe przy rozwiązaniu równania

$$\frac{dK}{dx} = \left(\frac{4Aj}{\varepsilon\varepsilon_0 v_{ds}(b+1)}\right)^{1/2} K^{\frac{b+1}{2}} \tag{29}$$

uzyskanego przez podstawienie zależności (13) do równania (28), otrzymujemy

$$K = \left[\left(\frac{Aj}{\varepsilon\varepsilon_0 v_{ds}(b+1)}\right)^{1/2}(1-b)\right]^{\frac{2}{1-b}}(x - d_0)^{\frac{2}{1-b}}. \tag{30}$$

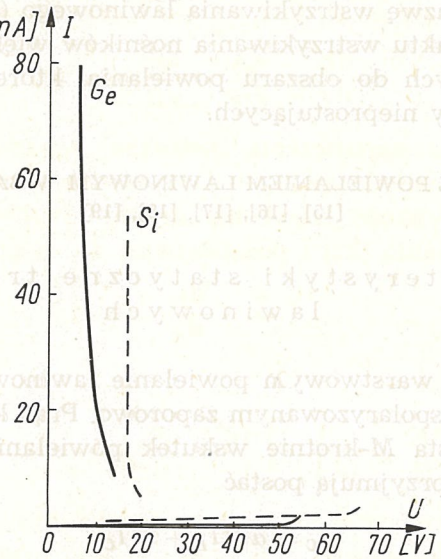
Występująca w tym wzorze szerokość obszaru powielania d_0 jest uzależniona od gęstości prądu

$$d_0 = \left(\frac{1}{4\varepsilon\varepsilon_0 v_{ds}}\right)^{\frac{-b}{1+b}} \frac{1}{2(b-1)} \left(\frac{A}{1+b}\right)^{\frac{-1}{1+b}} j^{\frac{-b}{1+b}}.$$

Ze wzoru tego wynika, że szerokość obszaru powielania lawinowego zmniejsza się ze wzrostem gęstości prądu [1].

Całkując wzór (30) w granicach $x = 0$, $x = d_0$, uzyskujemy wartość napięcia panującego na złączu

$$U_a = \int_0^{d_0} K\,dx = \left(\frac{1}{4\varepsilon\varepsilon_0 v_{ds}}\right)^{\frac{1-b}{1+b}} \frac{1}{2(b-3)} \left(\frac{1+b}{A}\right)^{\frac{2}{1+b}} j^{\frac{1-b}{1+b}}. \tag{31}$$



Rys. 6. Typowe charakterystyki diod punktowych z powielaniem lawinowym w obszarze dużych prądów

Ze wzoru tego wynika możliwość powstania oporności ujemnej przy dostatecznie silnej zależności $\alpha(K)$ (przy $b > 3$). Oporność ujemna dla złącza o powierzchni S ma wartość

$$r_a = \frac{dU_a}{S\,dj} = \frac{1}{S} \frac{1-b}{2(b-3)} \left(\frac{1+b}{4\varepsilon\varepsilon_0 v_{ds}}\right)^{\frac{1-b}{1+b}} (Aj)^{-\frac{2b}{1+b}}. \tag{32}$$

Charakterystyki diod doświadczalnych potwierdzając możliwość uzyskania oporności ujemnych dają zwykle nieco mniej silną zależność

[1] Można konstrukcyjnie uniezależnić obszar powielania od prądu umieszczając w diodzie pomiędzy obszarami p i n cienką warstwę półprzewodnika samoistnego, w granicach której odbywa się proces powielania (dioda p^+-i-n^+).

$U_a(j)$. Jest to prawdopodobnie związane z tym, że wzór (13) dla bardzo dużych wartości natężeń pola elektrycznego podaje zbyt silną zależność $\alpha(K)$. Różnice te mogą być również związane z nieco odmienną budową diod doświadczalnych (układ niepłaski).

Zjawisko powielania omawianego typu uzyskano również przy złączach nieprostujących, typu nn^+ lub pp^+. Złącza takie wykonano [5], [13] przez wtopienie kulki z materiału n^+ (lub p^+) w objętości półprzewodnika o małej koncentracji domieszek, odpowiednio typu n lub p. Tak wykonane złącze półkuliste umożliwia uzyskanie dużych gęstości prądu w półprzewodniku samoistnym, na granicy obu obszarów. Charakterystyki statyczne takich złączy przy zastosowaniu materiałów typu n przedstawione są na rys. 6.

Przedstawione zjawisko powielania lawinowego przy dużych gęstościach prądu nosi nazwę wstrzykiwania lawinowego (avalanche injection). Nazwa wynika z faktu wstrzykiwania nośników większościowych w głąb obszarów przyległych do obszaru powielania, które ma miejsce nawet w przypadku złączy nieprostujących.

4. TRANZYSTORY Z POWIELANIEM LAWINOWYM W BARIERZE KOLEKTORA [15], [16], [17], [18], [19]

4.1. Charakterystyki statyczne tranzystorów lawinowych

W tranzystorze warstwowym powielanie lawinowe występuje w złączu kolektorowym spolaryzowanym zaporowo. Prąd kolektorowy w takim tranzystorze wzrasta M-krotnie wskutek powielania nośników, i równania tranzystora przyjmują postać

$$I_C = \alpha M I_E + M I_{C0} \tag{33}$$

dla układu ze wspólną bazą,
oraz

$$I_C = \frac{\alpha M}{1 - \alpha M} I_B + \frac{M}{1 - \alpha M} I_{C0} \tag{34}$$

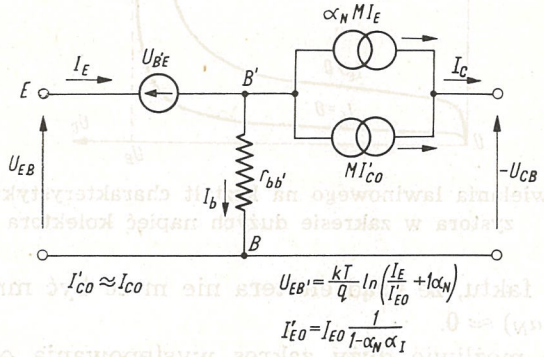
dla układu ze wspólnym emiterem.

Powielanie lawinowe daje więc wartość wzmocnienia prądowego równą iloczynowi αM w układzie ze wspólną bazą. Schemat zastępczy tranzystora lawinowego dla prądu stałego przedstawiony jest na rys. 7. Na schemacie tym zastosowano następujące oznaczenia:

I_{C0}, I_{E0} — prądy zerowe złączy kolektorowego i emiterowego,
α_N, α_I — wzmocnienia zwarciowe prądowe tranzystora przy normalnym i inwersyjnym połączeniu tranzystora (przy zamianie emitera i kolektora) [15].

Występowanie mnożnika M w równaniach (33) i (34) powoduje, że prąd I_C rośnie nieograniczenie przy wzroście napięcia kolektora. Przyjmując uproszczoną zależność $M\left(\frac{U_C}{U_B}\right)$ w postaci wzoru (3) lub (26) równanie (34) może być zapisane w postaci

$$I_C = \frac{\beta I'_B + \frac{\beta}{\alpha} I_{C0}}{1 - \left(\frac{U_C}{U_\alpha}\right)^n}, \tag{34a}$$



Rys. 7. Schemat zastępczy tranzystora lawinowego dla dużych sygnałów

gdzie:

$$\beta = \frac{\alpha}{1-\alpha},$$

$$U_\alpha = U_B (1-\alpha)^{\frac{1}{n}}, \tag{35}$$

U_C — napięcie na złączu kolektorowym, równe w przybliżeniu napięciu U_{CE}.

Charakterystyki statyczne tranzystora odpowiadające równaniom (33) i (34a) przedstawione są na rys. 8 i 9. W układzie ze wspólną bazą prąd kolektora rośnie monotonicznie ze wzrostem napięcia, maksymalne napięcie kolektora, przy którym prąd kolektora dąży do nieskończoności jest równe U_B. Przy $I_E = 0$ prąd kolektora zmienia się wg wzoru

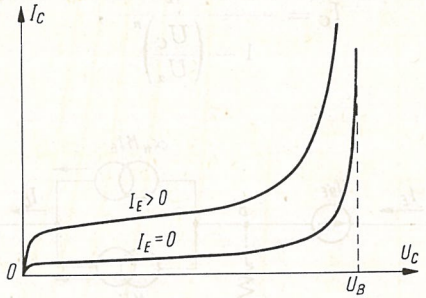
$$I_C = I_{C0} \cdot M. \tag{36}$$

Na charakterystykach statycznych tranzystora lawinowego w układzie wspólnego emitera (rys. 9) można wyróżnić obszar dodatnich oporności kolektora oraz obszar ujemnych oporności kolektora, oba te obszary rozgranicza wartość napięcia kolektora $U_C = U_\alpha$, dla której wypadkowe wzmocnienie prądowe $\alpha M(U_C)$ jest równe jedności. Obszar oporności

ujemnej kolektora istnieje jedynie dla ujemnych prądów bazy spełniających warunek

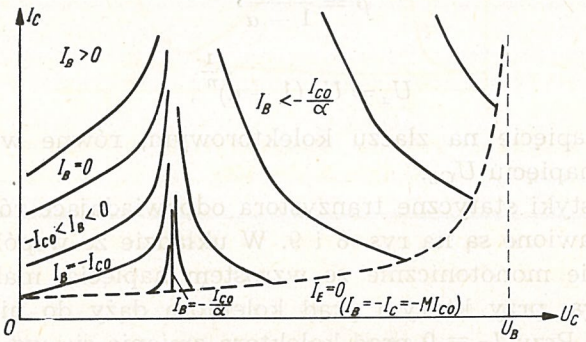
$$I_B < -\frac{I_{C0}}{\alpha}. \tag{37}$$

Przebieg charakterystyk statycznych jest ograniczony najmniejszą możliwą wartością prądu kolektora wyrażoną zależnością (36), ograniczenie to wynika z faktu, że prąd emitera nie może być mniejszy od wartości $-I'_{E0}(1-\alpha_N) \approx 0$.



Rys. 8. Wpływ powielania lawinowego na kształt charakterystyki statycznej tranzystora w zakresie dużych napięć kolektora

Aby uzyskać możliwie duży zakres występowania oporności ujemnych $\left(\text{mały stosunek } \frac{U_\alpha}{U_B}\right)$, co jest korzystne przy stosowaniu tranzystorów lawinowych w układach regeneracyjnych, zgodnie ze wzorem (35) należy dążyć do uzyskania możliwie dużej wartości α i małej wartości n.

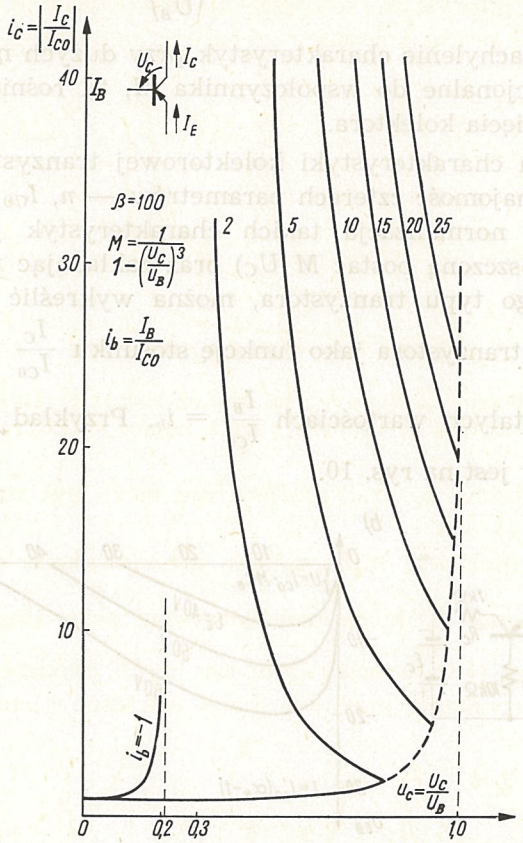


Rys. 9. Schematyczne przedstawienie rodziny kolektorowych charakterystyk tranzystora warstwowego

Z tego względu stosowanie np. germanowych tranzystorów typu p-n-p jest korzystniejsze od stosowania analogicznych tranzystorów typu n-p-n ($M_p > M_n$). Odpowiednio, dla stopowych tranzystorów krzemowych zależność jest odwrotna (co jednak nie ma znaczenia praktycznego, gdyż tranzystory krzemowe nie są na ogół wykonywane techniką stopową).

Przewodność dynamiczną kolektora można znaleźć przez zróżniczkowanie wzoru (34)

$$g_C = \frac{dI_C}{dU_C} = \frac{I_C}{M(1-M\alpha)} \frac{dM}{dU_C} \qquad (I_B = \text{const},\ U_{EB'} \approx 0). \tag{38}$$



Rys. 10. Znormalizowana rodzina charakterystyk kolektorowych tranzystora ($n = 3$, $\beta = 100$)

Bezwzględna wartość przewodności dynamicznej kolektora jest proporcjonalna do prądu kolektora. Stosując zależność (3) uzyskujemy wzór

$$g_C = \frac{M-1}{1-M\alpha} n \frac{I_C}{U_C}, \tag{39}$$

który dla $M \gg 1$ i $\alpha \approx 1$ można uprościć do postaci

$$g_C \approx -n \frac{I_C}{U_C}. \tag{40}$$

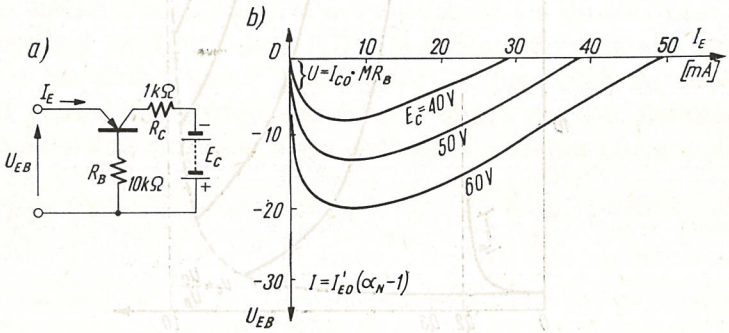
Jest to wzór słuszny przy napięciach kolektora zbliżonych do U_B, tj. $U_C \approx U_B$. Podstawiając do tego wzoru zależność (36) można wyznaczyć

początkowe nachylenie charakterystyk oporności ujemnej g_{C0} (przy $I_E \approx 0$);

$$g_{C0} \approx -\frac{I_{C0}}{U_B} \frac{n}{1-\left(\frac{U}{U_B}\right)^n}. \tag{41}$$

Początkowe nachylenie charakterystyk przy dużych napięciach kolektora jest proporcjonalne do współczynnika M, tj. rośnie bardzo szybko ze wzrostem napięcia kolektora.

Do określenia charakterystyki kolektorowej tranzystora lawinowego potrzebna jest znajomość czterech parametrów — n, I_{C0}, U_B, U_α (lub α), z tego względu normalizacja takich charakterystyk jest utrudniona. Przyjmując uproszczoną postać $M(U_C)$ oraz zakładając ustaloną wartość α i n dla danego typu tranzystora, można wykreślić znormalizowaną charakterystykę tranzystora jako funkcję stosunku $\frac{I_C}{I_{C0}} = i_c$ od stosunku $\frac{U_C}{U_B} = u_c$, przy stałych wartościach $\frac{I_B}{I_{C0}} = i_b$. Przykład takiej charakterystyki podany jest na rys. 10.



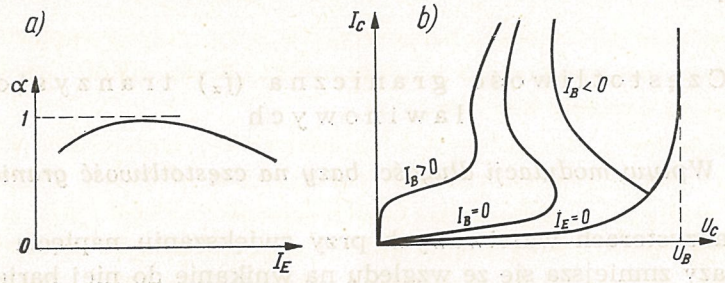
Rys. 11. Układ lawinowy z opornością ujemną w obwodzie emitera

Charakterystyki oporności ujemnej mogą być otrzymane również w obwodzie emitera przez wykorzystanie dodatniego sprzężenia zwrotnego na oporności w obwodzie bazy. Układ tego typu wraz z odpowiadającymi mu charakterystykami oporności ujemnej przedstawiony jest na rys. 11. Dla uzyskania oporności ujemnej niezbędne jest, aby napięcie zasilające w obwodzie kolektora było większe od U_α[1].

Charakterystyki rzeczywistych tranzystorów lawinowych mogą w pewnych przypadkach wykazywać anomalie, różniące je w dość istotny sposób od omówionych charakterystyk idealnych. Największy wpływ na

[1] Do obliczenia charakterystyk wykorzystano schemat zastępczy tranzystora lawinowego przedstawiony na rys. 7. Przy dokonywaniu obliczeń zaleca się wyrazić oddzielne wielkości I_E i U_{EB} jako funkcje prądu kolektora I_C.

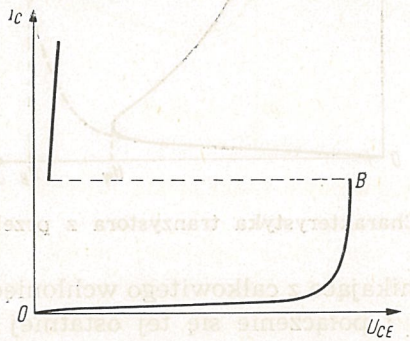
przebieg charakterystyk wywiera zależność α od prądu emitera, która powoduje zależność napięcia U_α od prądu emitera, zgodnie ze wzorem (35). W wyniku tej zależności charakterystyki kolektorowe mogą przybrać kształt jak na rys. 12. Jak widać, wykorzystując zakres początkowego wzrostu α przy wzroście prądu emitera można uzyskać charakte-



Rys. 12. Wpływ silnej zależności $\alpha(I_E)$ na kształt charakterystyk kolektorowych tranzystora lawinowego

rystyki oporności ujemnej w zakresie zerowych, a nawet dodatnich wartości prądu bazy. Zjawisko to jest wykorzystywane praktycznie do realizacji dwójników z opornością ujemną o małym napięciu w punkcie pracy, w tym celu stosuje się tranzystory o większej powierzchni emitera niż kolektora, dla których zależność $\alpha(I_E)$ jest wyrażona szczególnie silnie ze względu na duży wpływ rekombinacji powierzchniowej.

W zakresie dużych prądów kolektora oprócz powielania lawinowego uzależnionego od napięcia kolektora może wystąpić powielanie lawinowe



Rys. 13. Kształt charakterystyki kolektorowej tranzystora z powielaniem lawinowym uzależnionym od prądu (wstrzykiwanie lawinowe)

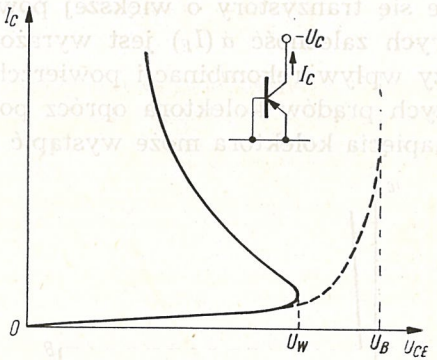
uzależnione od prądu, przy którym duże znaczenie ma pole ładunku przestrzennego nośników większościowych wprowadzanych do bazy w wyniku tzw. wstrzykiwania lawinowego. Typowe charakterystyki kolektorowe dla tego przypadku przedstawione są na rys. 13 [19]. W zakresie napięć od zera do p.B. występuje powielanie lawinowe napięciowe, po przekroczeniu p.B. czynnikiem decydującym staje się powielanie la-

winowe prądowe, w wyniku działania którego napięcie ustala się na bardzo małym poziomie (rzędu 1 V). Podobny kształt charakterystyk otrzymuje się w przypadku struktur wielokrotnych typu *p-n-p-n*, m. in. w zwykłych tranzystorach przy niedoskonałej konstrukcji złączy doprowadzających. Mechanizm tych zjawisk nie jest jeszcze dokładnie wyjaśniony.

4.2. Częstotliwość graniczna (f_α) tranzystorów lawinowych

4.2.1. Wpływ modulacji długości bazy na częstotliwość graniczną

W tranzystorach warstwowych przy zwiększaniu napięcia kolektora obszar bazy zmniejsza się ze względu na wnikanie do niej bariery złącza kolektorowego. Zjawisko to ma szczególne znaczenie w przypadku tranzystorów o jednorodnej bazie, natomiast w przypadku tranzystorów o gęstości domieszek rosnącej w miarę zbliżania się do emitera wpływ ten jest w dużej mierze pomijalny. Zmniejszanie się skutecznej długości bazy może powodować w tranzystorach stopowych tzw. przebicie skrośne



Rys. 14. Charakterystyka tranzystora z przebiciem bazy

emiter-kolektor, wynikające z całkowitego wchłonięcia obszaru bazy przez barierę kolektorową i połączenie się tej ostatniej z barierą emiterową. W stanie przebicia skrośnego bazy tranzystor charakteryzuje się nadzwyczaj małą opornością między emiterem i kolektorem oraz utratą możliwości kontrolowania prądów poprzez napięcie bazy wskutek gwałtownego wzrostu oporności wewnętrznej bazy r_{bb}'. Ponieważ przebiciu bazy towarzyszy zwykle silnie wyrażone powielanie lawinowe nośników w barierze kolektorowej, wypadkowa charakterystyka statyczna tranzystora z przebiciem bazy może mieć nachylenie ujemne (rys. 14). Ten rodzaj charakterystyki nie wyróżnia jednak w sposób szczególny tranzystorów z przebiciem bazy, gdyż może on być osiągnięty również w zwykłych

tranzystorach lawinowych bez przebicia bazy, przez włączenie do obwodu bazy dodatkowej oporności (ewentualnie bez dodatkowej oporności, gdy oporność własna bazy jest dostatecznie duża). Istotnie nowymi cechami tranzystorów pracujących w pobliżu przebicia bazy jest wspomniana utrata kontroli nad prądem kolektora poprzez napięcie bazy, a przede wszystkim znaczne zwiększenie częstotliwości granicznej tranzystora f_α wskutek zwężenia bazy. Jak wiadomo, częstotliwość graniczna f_α tranzystorów z bazą jednorodną wyraża wzór

$$f_\alpha = 2{,}44\,\frac{D}{W^2}, \tag{42}$$

gdzie:

D — stała dyfuzja nośników mniejszościowych w bazie,

$W = W_0 - d$ — skuteczna długość bazy (zmniejszona w stosunku do nominalnej długości bazy W_0 wskutek wnikania bariery kolektorowej).

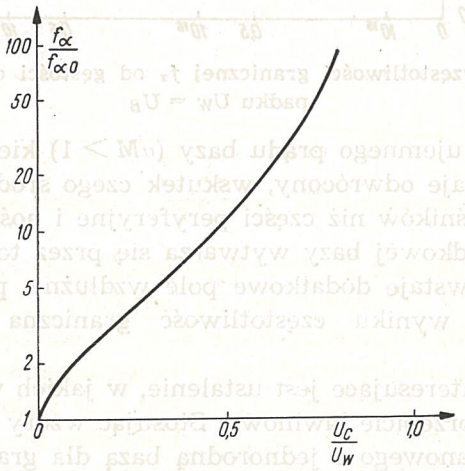
Stosując wzory (42) i (5) można ustalić dla złączy skokowych następującą zależność częstotliwości f_α od napięcia kolektora U_C

$$f_\alpha = \frac{f_{\alpha 0}}{\left[1 - \left(\frac{U_C}{U_W}\right)^{1/2}\right]^2}, \tag{43}$$

gdzie:

$f_{\alpha 0}$ — częstotliwość odpowiadająca szerokości bazy W_0,

U_W — napięcie kolektorowe, przy którym następuje przebicie bazy (tj. dla którego $W_0 = d$).

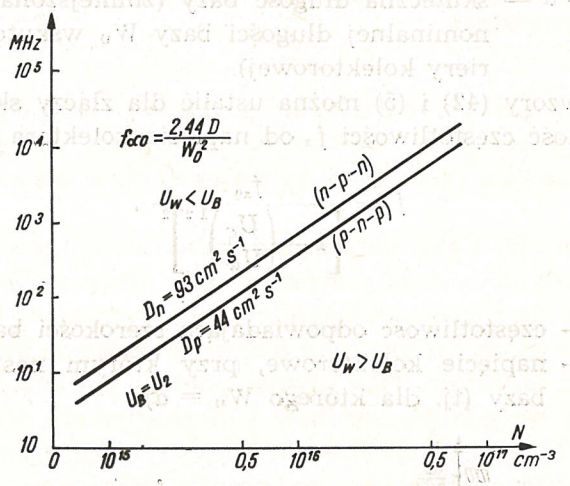


Rys. 15. Zależność częstotliwości granicznej f_α od napięcia kolektora

Wzór powyższy ilustruje wykres przedstawiony na rys. 15.

Należy wskazać na pewne czynniki, które mogą wpływać na niezgodność wyników praktycznych ze wzorem (43). Pierwszym takim czynni-

kiem jest fakt, że przebicie bazy nie następuje jednocześnie dla całej powierzchni rozgraniczającej obszar emitera od obszaru bazy, wskutek czego poszczególnym partiom tej powierzchni odpowiadają różne częstotliwości graniczne. Wypadkowa częstotliwość graniczna jest więc pewną średnią wartością, zależną w sposób decydujący od geometrii tranzystora. Drugim czynnikiem jest istnienie w bazie pola elektrycznego skierowanego poprzecznie, pole to wynika z przepływu prądu bazy i powoduje zakłócenia w ruchu nośników zmierzających od emitera do kolektora oraz nierównomierności emisji poszczególnych partii powierzchni emitera. W przypadku dodatniego prądu bazy ($\alpha M < 1$) części peryferyjne emitera wysyłają więcej nośników niż części środkowe, droga nośników ulega przy tym zakrzywieniu na zewnątrz i wydłużeniu, przez co średnia częstotliwość graniczna tranzystora zmniejsza się.



Rys. 16. Zależność częstotliwości granicznej f_α od gęstości domieszek N dla przypadku $U_W = U_B$

W przypadku ujemnego prądu bazy ($\alpha M > 1$) kierunek poprzecznego pola w bazie zostaje odwrócony, wskutek czego środkowa część emitera wysyła więcej nośników niż części peryferyjne i nośniki ulegają skupieniu. W części środkowej bazy wytwarza się przez to obszar o dużej gęstości prądu i powstaje dodatkowe pole wzdłużne, przyspieszające ruch nośników, a w wyniku częstotliwość graniczna tranzystora ulega zwiększeniu.

Praktycznie interesujące jest ustalenie, w jakich warunkach przebicie bazy wyprzedza przebicie lawinowe. Stosując wzory (5) i (11) można dla tranzystora germanowego z jednorodną bazą dla granicznego przypadku $U_W = U_B$ uzyskać następującą zależność:

$$W_0^2 N^{1{,}69} = 0{,}5 \cdot 10^{20}. \tag{44}$$

Zależność ta w zmodyfikowanej postaci [zastąpienie W_0^2 przez $f_{\alpha 0}$ według wzoru (42)] przedstawiona jest na rys. 16. Uwzględniając fakt,

że częstotliwości graniczne tranzystorów germanowych o jednorodnej bazie nie przekraczają obecnie kilkudziesięciu MHz, można wysnuć wniosek, że mniejsza wartość napięcia U_W od napięcia U_B może mieć miejsce jedynie dla tranzystorów o dużej wartości $f_{\alpha 0}$ (małe W_0) przy stosowaniu dużej oporności właściwej bazy.

Zależność (44) może być przekształcona przy pomocy wzorów (41) i (42) do postaci

$$f_{\alpha 0} U_B^{2,45} = 1{,}1 \cdot 10^{12} \tag{45}$$

przydatnej, gdy zachodzi konieczność zidentyfikowania rodzaju przebicia w tranzystorach o nieznanej koncentracji domieszek N w bazie. W przypadku gdy iloczyn $f_{\alpha 0} U_B^{2,45}$ ma wartość większą niż prawa strona zależności (45) ma miejsce przebicie bazy, w odwrotnym przypadku przebicie jest typu lawinowego.

4.2. Wpływ pola ładunku przestrzennego nośników większościowych wstrzykiwanych do bazy na częstotliwość graniczną tranzystora lawinowego. [17]

Zależności (42) i (43) określające częstotliwość graniczną tranzystora są słuszne jedynie w tym przypadku, gdy przenoszenie ładunków przez bazę odbywa się ruchem dyfuzyjnym. Dla tranzystora lawinowego przybliżenie takie może być przyjęte jedynie w zakresie bardzo małych prądów i napięć kolektora. Gdy gęstość prądu i powielanie lawinowe stają się tak duże, że gęstość nośników większościowych wstrzykiwanych do bazy w wyniku powielania staje się porównywalna z gęstością domieszek, wytwarza się pole przyspieszające nośniki, którego wpływ na przenoszenie ładunków w bazie może być decydujący. W tym przypadku wzór (42) przestaje być słuszny.

Aby określić częstotliwość graniczną tranzystora lawinowego przy dużych gęstościach prądu, spróbujemy wyznaczyć rozkład pola i rozkład gęstości nośników w bazie dla najprostszego przypadku tranzystora jednowymiarowego o jednorodnej bazie. Można tego dokonać przez rozwiązanie układu równań określających prądy dziurowy i elektronowy płynące przez bazę jako sumy prądów dyfuzyjnych i dryftowych

$$j_p = -qD_p \frac{dp}{dx} + qpu_p K, \tag{46}$$

$$j_n = qD_n \frac{dn}{dx} + qnu_n K. \tag{47}$$

Wobec dużej sprawności emitera prąd dziurowy j_p może być traktowany jako prąd emitera. Całkowity prąd płynący wzdłuż bazy $j_p + j_n$ jest M-krotnie większy od prądu j_n, zakładamy przy tym, że wpływ re-

kombinacji jest do pominięcia. Założymy również, że prąd elektronowy jest odprowadzany na zewnątrz przez wyprowadzenie bazy tuż przed emiterem, bez wywoływania poprzecznej składowej pola w bazie. W rzeczywistości to ostatnie założenie jest spełnione najgorzej, przyjmiemy je jednak dla uproszczenia rozważań.

Kompletując równania (46) i (47) warunkiem

$$j_n = (M-1)\, j_p \tag{48}$$

oraz warunkiem neutralizacji ładunku w obszarze bazy

$$n = p + N$$

uzyskujemy po ich rozwiązaniu

$$\frac{dp}{dx} = -\frac{j_p}{qD_p}\,\frac{N + p\left[1-(M-1)\dfrac{u_p}{u_n}\right]}{N_d + 2p}, \tag{49}$$

$$K = \frac{M-1+\dfrac{u_p}{u_n}}{qu_n(N+2p)}\, j_p. \tag{50}$$

Podstawiając

$$C = 1-(M-1)\frac{u_p}{u_n}$$

i całkując (49) przy warunku brzegowym

$$p(W) = 0$$

uzyskujemy

$$W - x = \frac{qD_pN}{j_pC^2}\left[2C\frac{p}{N} - (2-C)\ln\left(1+C\frac{p}{N}\right)\right]. \tag{51}$$

Rozkład gęstości nośników, jaki wynika z powyższej zależności, przedstawiony jest na rys. 17; rozkład ten przy dużych prądach j_p zbliża się do prostokątnego. Przy $M \gg 1$ (ujemna wartość C) największa możliwa wartość p wynosi

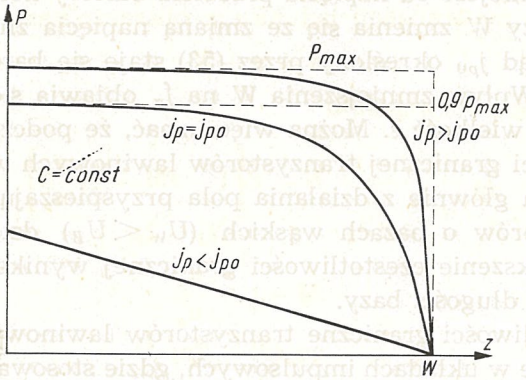
$$p_{max} = -\frac{N}{C}. \tag{52}$$

Gdy gęstość p zbliża się do p_{max}, wartość pochodnej $\dfrac{dp}{dx}$ w całym zakresie bazy (z wyjątkiem $x = W$) staje się równa zeru, tj. gęstość nośników osiąga wartość stałą. Dla tego przypadku, jak wynika ze wzoru (50), pole w bazie jest również stałe, co niezmiernie upraszcza rozważania, gdyż wówczas wartość częstotliwości granicznej tranzystora jest znana.

Łatwo jednakże zauważyć, że dla osiągnięcia gęstości p równej p_{max} przy skończenie dużych wartościach C i W wymagana gęstość prądu jest

równa nieskończoności. Rozwiązując zagadnienie w sposób praktyczny uznamy rozkład nośników za prostokątny, gdy gęstość nośników w punkcie $x = 0$ przekroczy wartość 0,9 p_{max}. Odpowiadający temu warunkowi prąd j_{p0} wynosi

$$j_{p0} = \frac{qD_pN}{W}\,\frac{2{,}8 - 2{,}3\,C}{C^2}. \tag{53}$$



Rys. 17. Rozkład gęstości nośników w bazie tranzystora lawinowego dla różnych gęstości prądu dziurowego przy stałej wartości napięcia kolektora

Dla $j_p \gg j_{p0}$ można przyjąć, że wartość K jest stała i równa

$$K = -\frac{j_pC}{qu_pN} = \frac{kT}{q}\,\frac{j_p}{qD_pp_{max}} \tag{54}$$

$\left(\text{gdzie } \frac{kT}{q} = 0{,}026 \text{ przy } T = 300^0\right)$.

Przy stałym polu w bazie wartość f_α wynosi [20]

$$f_\alpha = \varepsilon^{3/2} f_{\alpha 0}, \tag{55}$$

gdzie:

$$\varepsilon = \frac{q}{2kT} KW,$$

$f_{\alpha 0}$ — częstotliwość graniczna określona wzorem (42).

Działanie pola wytworzonego przez nośniki większościowe może powodować, że częstotliwość f_α staje się wielokrotnie większa od częstotliwości $f_{\alpha 0}$. Dla przykładu, przy $M = 5$, $N_d = 10^{15}$ cm^{-3}, $W = 2{,}5 \cdot 10^{-3}$ cm, dla tranzystora germanowego p-n-p przy gęstości prądu $j_p + j_n = 500$ A cm^{-2} ($j_p = 100$ A cm^{-2}) uzyskujemy $f_\alpha \approx 75\, f_{\alpha 0}$.

Częstotliwość graniczna tranzystora lawinowego zależy w równym stopniu od j_p jak i od C, gdyż obie te wielkości występują w zależności (54) w iloczynie. Wpływ napięcia kolektora na f_α może być stosunkowo

większy niż wpływ prądu kolektora, gdyż C jest szybko rosnącą funkcją napięcia kolektora. Wzrost napięcia kolektora i towarzyszący mu wzrost C powoduje przy tym zmniejszenie j_{p0} [zależność (53)], przez co warunek dostatecznie dużej gęstości prądu ($j_p \gg j_{p0}$) może być spełniony również przy umiarkowanie dużych prądach kolektora.

Można jednak wskazać warunki, w których wzrost f_α w stosunku do $f_{\alpha 0}$ nie jest duży. Tak jest w przypadku, gdy napięcie przebicia bazy U_W jest wyraźnie mniejsze od napięcia przebicia bariery kolektora U_B. Wówczas długość bazy W zmienia się ze zmianą napięcia znacznie silniej niż wartość C^2 i prąd j_{p0} określony przez (53) staje się bardzo duży, trudny do osiągnięcia. Wpływ zmniejszenia W na f_α objawia się również bezpośrednio poprzez wielkość ε. Można więc uznać, że podczas gdy zwiększenie częstotliwości granicznej tranzystorów lawinowych w przypadku długich baz wynika głównie z działania pola przyspieszającego, to w przypadku tranzystorów o bazach wąskich ($U_W < U_B$) działanie pola jest mniejsze i zwiększenie częstotliwości granicznej wynika głównie ze zjawiska modulacji długości bazy.

Duże częstotliwości graniczne tranzystorów lawinowych są najpełniej wykorzystywane w układach impulsowych, gdzie stosowanie dużych chwilowych gęstości prądu i dużych napięć nie prowadzi do zniszczenia tranzystora wskutek przekroczenia jego mocy admisyjnej, gdyż średnie wartości wydzielanej mocy są małe.

5. ZASTOSOWANIE DIOD I TRANZYSTORÓW LAWINOWYCH

Półprzewodnikowe elementy lawinowe mogą być wykorzystane w układach:

a) stabilizatorów napięcia,

b) wzmacniaczy oraz generatorów przebiegów sinusoidalnych, pracujących na zasadzie ujemnej oporności,

c) generatorów przebiegów impulsowych.

Omówimy pokrótce możliwości zastosowania diod i tranzystorów w każdej z wymienionych dziedzin.

a) Diody i tranzystory lawinowe mogą z powodzeniem zastępować napięciowe stabilizatory jonowe, różniąc się od nich korzystnie brakiem histerezy między punktem zapłonu a punktem gaśnięcia. Stałość temperaturowa napięcia przebicia U_B jest wysoka i wynosi około 0,1% na jeden stopień przyrostu temperatury. W przypadku tranzystorów lawinowych możliwe jest uzyskanie regulowanego napięcia stabilizacji (przy pomocy prądu bazy), uzyskuje się wówczas wobec ujemnej oporności kolektora przyrosty napięcia stabilizowanego przeciwnego znaku w stosunku do przyrostów napięcia zasilającego.

b) Wytwarzanie się w pewnych warunkach procesu lawinowego oporności ujemnej stwarza możliwości wzmacniania i generacji przy pomocy

elementów lawinowych. Możliwości te nie są dotychczas w pełni wykorzystywane i w zasadzie nie wyszły z kręgu zagadnień doświadczalnych, na co składa się szereg przyczyn. Do najważniejszych należą: konieczność rozpraszania dużych mocy w punkcie pracy oraz silna konkurencja ze strony diody tunelowej. Pomimo to zastosowanie elementów lawinowych w tej dziedzinie wydaje się obiecujące [1].

c) Zalety tranzystorów lawinowych znalazły jak dotychczas najpełniejsze wykorzystanie w układach impulsowych. Przy pomocy tranzystorów lawinowych o nominalnej mocy admisyjnej 0,1 W możliwe jest uzyskiwanie impulsów prądowych o wartości szczytowej rzędu kilku amperów. Ze względu na dużą częstotliwość graniczną tranzystorów lawinowych czas narastania takich impulsów jest bardzo mały i w pewnych przypadkach sięga 10^{-9} sek. Możliwe jest również uzyskiwanie dużych częstotliwości powtarzania impulsów, rzędu 10^7 Hz. Ze względu na wymienione cechy tranzystory lawinowe najchętniej są stosowane w układach wyzwalających i kluczujących o szybkim działaniu.

Politechnika Warszawska
Katedra Podstaw Telekomunikacji

WYKAZ LITERATURY

1. Szałyt S. S.: Elektroprowodnost' połuprowodnikow, w książce Połuprowodniki w naukie i tiechnikie, t. I, AN ZSSR, 1957.
2. Czujenkow W. A.: Tieorija elektriczeskogo proboja połuprowodnikow, cz. I i II, w książce Fizika twiordogo tieła, t. II, AN ZSSR, 1959.
3. Ryder E. J.: Phys. Rev. t. 90, nr 5, 1953.
4. Arthur J. B., Gibson A. F., Granwill J. W.: J. of Electronics, t. 2, nr 2, 1956.
5. Gunn J. B.: High Electric Field Effects in Semiconductors, w książce Progress in Semiconductors, t. 2, Heywood, 1957.
6. Stratton R.: On the Theoretical Variation of Electron Mobility with Applied Electric Field in Semiconductors, w książce Solid State Physics, t. 1, cz. 1, Acad. Press, London — New York 1960.
7. Ilisawskij J. W.: Ławinnyje triody, w książce Połuprowodnikowyje pribory w naukie i tiechnikie, t. II, AN ZSSR, 1958.
8. Wuł B. M.: O proboje pieriechodnych słojow w połuprowodnikach. — ŻTF, t. 26, nr 11, 1956.
9. Szotow A. P.: Proboj elektronno-dyrocznych pieriechodow na udarnom napriażenii. — ŻTF, t. 26, nr 8, 1956.
10. Wuł B. M., Szotow A. P.: Ob udarnoj jonizacii w kriemniewych *p-n* pieriechodach, w książce Fizika twiordogo tieła, t. I, AN ZSSR, 1959.

[1] Należy tu odnotować niezmiernie interesującą propozycję [21] wykorzystania diody do generacji mocy rzędu kilkudziesięciu W w zakresie od 1 do 50 GHz. Proponuje się w tym celu specjalną konstrukcję diody lawinowej, w której oporność ujemna powstaje w wyniku opóźnienia wprowadzanego przez nośniki dryfujące w obszarze bariery.

11. Wuł B. M., Szotow A. P.: Multiplication of Electron and Holes in p-n Junctions, w książce Solid State Physics, t. 1, cz. 1, Academ. Press, London — New York 1960.
12. Lade R. W., Schlax T. R.: A High Field Effect Two Terminal Oscillator. — PIRE, 48, May 1960.
13. Gibson A. F.: The Theory of Avalanche Injection Diodes. Inst. of Electr. Eng., Paper No 3098E, Jan. 1960.
14. Lade R. W.: Theory and Application of Avalanche Diodes. — Proc. of the Not. El. Conf., vol. XVI, 1960.
15. Miller S. L., Ebers J. J.: Alloyed Junction Avalanche Transistors. — BSTJ, 34, 1955.
16. Kidd M. G.: Delayed Collector Conduction, a New Effect in Junction Transistors. — RCA Review, 16, 1955.
17. Shenkel H., Statz H.: Junction Transistors with Alpha Greater than Unity. — PIRE, 44, March 1956.
18. Beale J. R. A., Stephenson W. L., Wolfendale E.: A Study of High-Speed Avalanche Transistors. — PIEE, p. B, 104, July 1957.
19. Thornton G. G., Simmons C. D.: A New High Current Mode of Transistor Operation. — Trans. IRE on E. D., v. Ed-5, nr 1, 1958.
20. Krömer H.: Zur Theorie des Diffusions und des Drift Transistors. — Archiv. Elektr. Übertrg. nr 5, 8, 11, 1954.
21. Read W. T.: A Proposed High — Frequency Negative Distance — Resistance Diode. — BSTJ, 1958, v. XXXVII, nr. 2.
22. Hamilton D. J., Gibbons J. F., Shockley W.: Physical Principles of Avalanche Transistor Pulse Circuits. — PIRE, 1959, v. 47, nr 6.

J. BARANOWSKI

LAWINENARTIGE IONISATIONSVERVIELFÄLTIGUNG IN HALBLEITERELEMENTEN

Zusammenfassung

In diesem Aufsatz wurden grundsätzliche physikalische Erscheinungen angegeben, welche beim lawinenartigen Ionisations-Vervielfältigungsprozess als Begleiterscheinungen auftreten. Dieses betrifft vor allen Dingen diejenigen Instanzen, die in der Praxis solche Elemente anwenden. In diesem Aufsatz wurden in der Reihenfolge die lawinenartige Ionisations-Vervielfältigung der einheitlichen Halbleiterprobe, die p-n Verbindung, die Beschaffenheit des Schichtentransistors mit lawinenartiger Ionisations-Vervielfältigung in der Kollektorenschranke, besprochen.

Eine verhältnismässig grosse Beachtung wurde der lawinenartigen Transistorenionisation gewidmet. Hier wurde nicht nur allein der direkte Einfluss der lawinenartigen Ionisations-Vervielfältigung auf die statischen Charakteristiken besprochen, aber auch die Begleiterscheinungen, wie Abnormitäten der Charakteristiken und die Grenzfrequenzvergrösserung des Transistors.

Zuletzt wurden die Bereichsmöglichkeiten für die Anwendung der technischen Elemente mit lawinenartiger Ionisations-Vervielfältigung angegeben.

Е. БАРАНОВСКИ

ЛАВИННОЕ УМНОЖЕНИЕ В ПОЛУПРОВОДНИКОВЫХ ПРИБОРАХ

Резюме

Данная работа представляет основные физические явления связанные с процессом лавинного умножения в полупроводниковых приборах, и есть адресованная преимущественно к практическим употребителям таких приборов. В работе очередно оговорены лавинное умножение в однородном полупроводниковом образце, в электронно-дырочном переходе, и свойства плоскоемного транзистора с лавинном умножением в коллекторном переходе. Найболее внимания посвящено лавинному умножению в транзисторе, оговорено не только прямое влияние лавинного умножения на статическую характеристику, но тоже присутствующие явления, такие как искажения характеристик и явления повышения граничной частоты. В заключении оговариваются области техники в которых применение полупроводниковых приборов с лавинным умножением может иметь место.

621.396.677.7

MARIA RZEPECKA

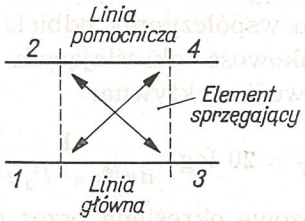
Krzyżowe sprzęgacze kierunkowe

Rękopis dostarczono 27.2.1963

Podano przegląd istniejących rozwiązań sprzęgaczy kierunkowych ze szczególnym uwzględnieniem sprzęgaczy krzyżowych. Na wstępie krótko omówiono parametry, zastosowanie oraz porównano właściwości różnych typów sprzęgaczy. Przeanalizowano działanie sprzęgaczy z aperturami sprzęgającymi o małych wymiarach w porównaniu z długością fali i na tej podstawie podano wzory do projektowania sprzęgaczy o krótkich i wąskich szczelinach oraz wyprowadzono wzory dla krzyżowych sprzęgaczy z otworami kołowymi. Szczególną uwagę poświęcono sprzęgaczom krzyżowym z trzema otworami kołowymi ze względu na szereg zalet takiego rozwiązania oraz brak dokładnego omówienia teoretycznego tego typu sprzęgacza w literaturze. Na zakończenie podano wyniki prac eksperymentalnych oraz parametry sprzęgaczy 10, 20 i 30 dB na pasmo 3 cm.

1. PARAMETRY SPRZĘGACZY KIERUNKOWYCH

Sprzęgacz kierunkowy jest biernym, liniowym układem ośmiozaciskowym (rys. 1), w którym moc doprowadzona na wejście linii głównej



Rys. 1. Sprzęgacz kierunkowy — oznaczenia

(*1*) rozchodzi się w linii głównej (w kierunku *3*) oraz jedynie w jednym kierunku w linii pomocniczej (*4*). Linia główna jest to prowadnica falowa (linia współosiowa lub falowodowa) łącząca generator z obciążeniem. Element sprzęgający przenosi część energii z linii głównej do linii pomocniczej.

Sprzęgacze kierunkowe charakteryzuje się za pomocą następujących parametrów: sprzężenie, kierunkowość, izolacja oraz współczynniki fali stojącej (*WFS*) w linii głównej oraz pomocniczej. Czasem używa się do

określenia właściwości sprzęgacza współczynników sprzężenia i kierunkowości.

Sprzężenie oraz współczynnik sprzężenia określają stosunek mocy w linii pomocniczej rozchodzącej się w kierunku pożądanym do mocy na wejściu linii głównej.

$$\text{Współczynnik sprzężenia} = c = \sqrt{\frac{P_1}{P_i}} \tag{1}$$

$$\text{Sprzężenie} = C = 20 \log \frac{1}{c} \text{[dB]}, \tag{2}$$

gdzie:

P_i — oznacza moc na wejściu linii głównej,

P_4 — moc dochodząca do obciążenia *4*.

Kierunkowość oraz współczynnik kierunkowości określają stosunek mocy rozchodzącej się w kierunku niepożądanym do mocy rozchodzącej się w kierunku pożądanym w linii pomocniczej.

$$\text{Współczynnik kierunkowości} = d = \sqrt{\frac{P_2}{P_4}} \tag{3}$$

$$\text{Kierunkowość} = D = 20 \log \frac{1}{d} \text{[dB]}, \tag{4}$$

gdzie P_2 — moc dochodząca do obciążenia *2*.

Izolacja określa stosunek mocy rozchodzącej się w kierunku niepożądanym w linii pomocniczej i mocy na wejściu układu.

$$\text{Izolacja} = I = 20 \log \sqrt{\frac{P_i}{P_2}}. \tag{5}$$

Dla rzeczywistego sprzęgacza, tzn. gdy obciążenie *2* nie jest idealnie dopasowane i wprowadza współczynnik odbicia Γ_2, parametrem bardziej miarodajnie niż kierunkowość określającym właściwości kierunkowe sprzęgacza jest kierunkowość efektywna:

$$D_{ef} = 20 \log \frac{1}{|d e^{j\Phi_1} + \Gamma_2 e^{j\Phi_2}|}, \tag{6}$$

gdzie Φ_1, Φ_2 — kąty fazowe określone przez długość drogi propagacji wzdłuż sprzęgacza, kąt fazowy współczynnika odbicia oraz przesunięcie fazy w elemencie sprzęgającym.

WFS w linii głównej i pomocniczej zdeterminowany jest przez odbicia do elementu sprzęgającego, dopasowanie obciążeń oraz ewentualne nieciągłości w prowadnicy (strojniki, podpory itp.).

2. ZASTOSOWANIE ORAZ PORÓWNANIE WŁAŚCIWOŚCI RÓŻNYCH TYPÓW SPRZĘGACZY

Sprzęgacze kierunkowe znajdują szerokie zastosowanie w technice mikrofalowej. Są one stosowane do pomiarów mocy, pomiarów współ-

czynnika odbicia, jako kalibrowane tłumiki, dzielniki mocy, przesuwniki fazy, do rozdzielania i pomiaru dwóch fal rozchodzących się w prowadnicy w przeciwnych kierunkach oraz w wielu innych przypadkach. W zależności od celu, do jakiego jest przeznaczony sprzęgacz, różne są stawiane wymagania dotyczące poszczególnych parametrów, co z kolei pociąga za sobą różnorodne rozwiązania konstrukcyjne.

Skala potrzebnych sprzężeń jest szeroka i obejmuje zakres 0—60 dB. Wysokie wymagania dotyczące dużej kierunkowości (ponad 40 dB) oraz dopasowania ($WFS \leqslant 1{,}05$) w szerokim pasmie częstotliwości powodują konieczność konstrukcji sprzęgaczy o dużych gabarytach (sprzęgacze wielootworowe). W wielu przypadkach nie zachodzi potrzeba stosowania sprzęgaczy o tak wysokich wymaganiach dotyczących parametrów, natomiast istotną sprawą jest zapewnienie możliwie małych gabarytów. Szerokie zastosowanie w tych przypadkach znalazły sprzęgacze krzyżowe dla sprzężeń powyżej 10 dB oraz rozwidlone dla sprzężeń silniejszych [1] (3 dB sprzęgacz Ribleta).

Wszystkie rozważania przeprowadzone w pracy dotyczą klasy sprzęgaczy falowodowych.

Istnieje duża różnorodność zarówno w kształcie elementów sprzęgających, jak i w sposobie ich usytuowania względem osi falowodu. Zarówno jedno, jak i drugie ma decydujący wpływ na parametry sprzęgacza. Elementy sprzęgające mogą znajdować się w szerszej ściance falowodu lub w węższej, istnieją także sprzęgacze, w których szersza ścianka falowodu głównego jest połączona z węższą ścianką falowodu pomocniczego.

Falowody mogą być równoległe, skrzyżowane (tworzą wtedy kąt 90°) lub pod kątem (sprzęgacz Bethego [7]).

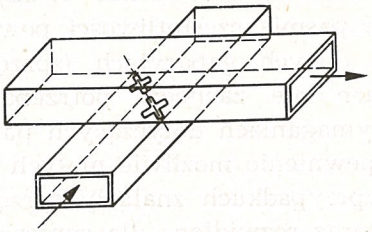
Jako elementy sprzęgające stosuje się rozmaite apertury. Najczęściej stosowane są otwory kołowe ze względu na takie zalety, jak łatwość wykonania, możliwość spełnienia wysokich wymagań dotyczących kierunkowości, dopasowania, szerokopasmowości (dla wielootworowych sprzęgaczy interferencyjnych) oraz dobra powtarzalność wyników.

W bardzo wielu sprzęgaczach apertury mają kształt wąskich i krótkich szczelin, czasem skrzyżowanych — sprzęgacz Moreno (rys. 2), czasem usytuowanych jak na rys. 3 (sprzęgacz Ribleta). Na rys. 4 został pokazany jeszcze jeden sprzęgacz szczelinowy Schwingera o przylegających do siebie różnych ściankach falowodów (szersza ścianka linii głównej przylega do węższej ścianki linii pomocniczej) [1].

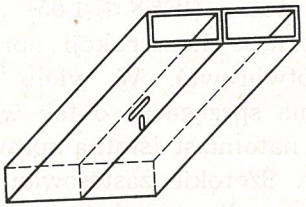
Element sprzęgający może mieć kształt długiej (kilka długości fali) i wąskiej szczeliny położonej wzdłuż osi falowodów. Przez wyjęcie części

[1] Przez sprzężenie silniejsze rozumie się większy współczynnik sprzężenia, co jest równoznaczne z mniejszym sprzężeniem wyrażonym w dB. Używane będzie zarówno określenie silniejsze sprzężenie, jak i mniejsze sprzężenie (w dB).

wspólnej ścianki przy przyleganiu węższych ścianek falowodów otrzymuje się sprzęgacze rozwidlone. Innym typem sprzęgacza kierunkowego jest sprzęgacz rozgałęziony, w którym sprzężenie uzyskuje się za pomocą ćwierćfalowych odcinków falowodu o impedancji charakterystycznej, na ogół różnej od impedancji obu sprzężonych falowodów (rys. 5). W zależności od żądanych parametrów można wybrać określoną konstrukcję sprzęgacza.

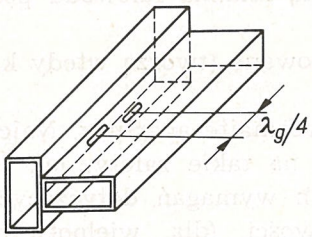


Rys. 2. Sprzęgacz Moreno

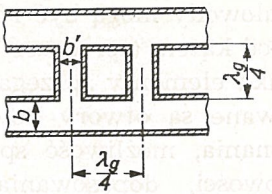


Rys. 3. Sprzęgacz Ribleta

Jak już powiedziano wyżej, dużą kierunkowość przy wartościach sprzężenia od 3—60 dB zapewniają wielootworowe sprzęgacze z otworami kołowymi. Sprzęgacz Bethego, krzyżowe sprzęgacze z otworami kołowymi oraz sprzęgacze z aperturami szczelinowymi pozwalają uzyskiwać sprzężenia od około 16 dB do 60 dB przy kierunkowości większej niż 15 dB, a w niektórych przypadkach około 20 dB. Krzyżowe sprzęgacze z otwo-



Rys. 4. Sprzęgacz Schwingera



Rys. 5. Sprzęgacz rozgałęziony

rami kołowymi przez zmianę impedancji charakterystycznej falowodów pozwalają uzyskać sprzężenie rzędu 10 dB. Sprzęgacze o długiej szczelinie, rozwidlone oraz rozgałęzione, stosowane są najczęściej, gdy zależy na małym sprzężeniu — mniejszym niż 10 dB przy kierunkowości około 15—20 dB. Jeszcze jednym ważnym parametrem, który należy wziąć pod uwagę przy wyborze typu sprzęgacza, są maksymalne dopuszczane zmiany sprzężenia w pasmie. Bliższe informacje na ten temat można znaleźć w literaturze [1], [4], [7].

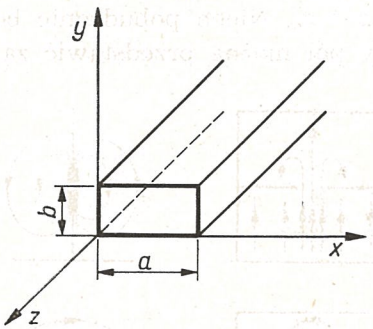
Omówione zostanie teraz zjawisko sprzężenia przez apertury o małych wymiarach w porównaniu z długością fali, co pozwoli na wyprowadzenie podstawowych zależności określających sprzężenie i kierunkowość.

Zjawiska występujące w sprzęgaczach rozwidlonych oraz rozgałęzionych mają inny charakter i w tej pracy nie będą omawiane.

3. TEORIA SPRZĘGACZY KIERUNKOWYCH

Stosowany będzie układ jednostek MKS oraz harmoniczna zmienność w czasie typu $e^{-j\omega t}$ (w dalszych wyprowadzeniach czynnik ten będzie pomijany, niemniej należy pamiętać o jego istnieniu).

Dla fali rozchodzącej się w falowodzie prostokątnym (rys. 6) można z równań Maxwella, uwzględniając warunki brzegowe, znaleźć składowe pól elektrycznego i magnetycznego.



Rys. 6. Falowód prostokątny

Dla rodzaju H_{10} rozchodzącego się w kierunku zgodnym ze zwrotem osi z otrzymuje się następujące znormalizowane (tzn. o amplitudach jednostkowych) składowe pól [9]:

$$\left.\begin{aligned} E_y &= \sin\frac{\pi x}{a}\, e^{-j\beta z} \\ H_x &= -\frac{\sqrt{1-(\lambda_0/\lambda_c)^2}}{\eta}\sin\frac{\pi x}{a}\, e^{-j\beta z} \\ H_z &= j\frac{\lambda_0}{\lambda_c}\frac{1}{\eta}\cos\frac{\pi x}{a}\, e^{-j\beta z} \end{aligned}\right\} \tag{7}$$

gdzie:

β — stała fazowa, $\beta = \dfrac{2\pi}{\lambda_g}$,

λ_0 — długość fali w wolnej przestrzeni,

λ_g — długość fali w falowodzie,

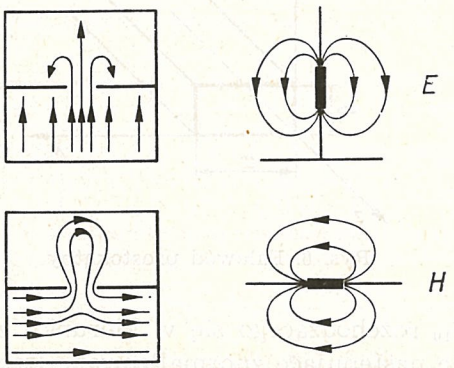
λ_c — graniczna długość fali,

$\eta = \sqrt{\dfrac{\mu}{\varepsilon}}$.

Dla propagacji w kierunku przeciwnym do zwrotu osi z:

$$\left.\begin{aligned} E_y &= \sin\frac{\pi x}{a}\, e^{j\beta z} \\ H_x &= -\frac{\sqrt{1-(\lambda_0/\lambda_c)^2}}{\eta}\sin\frac{\pi x}{a}\, e^{j\beta z} \\ H_z &= j\frac{\lambda_0}{\lambda_c}\frac{1}{\eta}\cos\frac{\pi x}{a}\, e^{j\beta z} \end{aligned}\right\} \tag{8}$$

Rozważmy teraz nieskończony falowód (ew. doskonale dopasowany na obu końcach) pobudzony w punkcie $z = z_o$. Rozkład pola w tym falowodzie określony będzie przez dwie fale rozchodzące się na prawo i na lewo od płaszczyzny $z = z_o$. Niech pobudzenie będzie rodzajem podstawowym H_{10}. Rozkłady pól można przedstawić za pomocą dipoli zastęp-



Rys. 7. Rozkłady pól w falowodzie oraz dipole zastępcze dla rodzaju H_{10}

czych (rys. 7) [5]. Sprzężenie przez pole elektryczne przedstawione zostało jako dipol o momencie PE_{oy} równoległym do pola wzbudzającego. Sprzężenie magnetyczne przedstawiono jako dipol magnetyczny o momentach M_1H_{ox} i M_2H_{oz} leżący w płaszczyźnie elementu sprzęgającego oraz mający zwrot przeciwny do pola wzbudzającego.

Indeks „o" oznacza pole wzbudzające — czyli pole w linii głównej. F, M_1, M_2 oznaczają zdolności polaryzacyjne otworu.

Zdolności polaryzacyjne są stałymi skupionymi charakteryzującymi własności elementu sprzęgającego (otworu), zależą one od wymiarów otworu, natomiast nie zależą od wielkości pola wzbudzającego, długości fali, położenia otworu w falowodzie. Dla elementu sprzęgającego w postaci otworu kołowego o promieniu r wynoszą one [3], [7]:

$$M_1 = M_2 = 2P = \frac{4}{3}r^3. \tag{9}$$

Dla szczeliny o długości L i szerokości w wyciętej prostopadle do osi z falowodu

$$\left.\begin{aligned} M_1 &= \frac{\pi}{3}\frac{L^3}{\ln\left(\frac{4L}{w}-1\right)} \\ M_2 &= P = \frac{\pi}{3}Lw^2 \end{aligned}\right\} \tag{10}$$

Dla bardzo wąskich szczelin można zaniedbać M_2 i P.

Bethe wykazał [3], że dla falowodu pobudzonego przez element sprzęgający o wymiarach małych w porównaniu z długością fali przy pobudzeniu rodzajem H_{10} falę rozchodzącą się w kierunku zgodnym z osią z (A_1) oraz w kierunku przeciwnym (B_1) można wyrazić następującymi wzorami.

Dla falowodów równoległych: [7]

$$\left.\begin{aligned} \hat{A}_1 &= \frac{j\beta}{2S_1}(M_1 H_{1x}H_{ox} + M_2 H_{1z}H_{oz} - PE_{1y}E_{oy}) \\ \hat{B}_1 &= \frac{j\beta}{2S_1}(-M_1 H_{1x}H_{ox} + M_2 H_{1z}H_{oz} - PE_{1y}E_{oy}) \end{aligned}\right\}. \tag{11}$$

Dla falowodów skrzyżowanych:

$$\left.\begin{aligned} \hat{A}_1 &= \frac{\beta}{2S_1}(M_1 H_{1x}H_{ol} + M_2 H_{1z}H_{ot} - jPE_{1y}E_{on}) \\ \hat{B}_1 &= \frac{\beta}{2S_1}(-M_1 H_{1x}H_{ol} + M_2 H_{1z}H_{ot} - jPE_{1y}E_{on}) \end{aligned}\right\}, \tag{12}$$

przy czym przyjęto następujące oznaczenia:
Indeks „o" oznacza linię główną
Indeks „1" oznacza linię pomocniczą
l, n, t — układ współrzędnych przy falowodach skrzyżowanych związany z linią główną.
x, y, z — układ współrzędnych związany z linią pomocniczą (rys. 12) (przy falowodach równoległych układy pokrywają się). Wielkości $\hat{A}_1$ i $\hat{B}_1$ są proporcjonalne do pierwiastków z mocy fal rozchodzących się w falowodach.

S_1 oznacza wektor proporcjonalny do wektora Poytinga i wynosi:

$$S_1 = \int_0^a \int_0^b \sqrt{1-(\lambda_0/\lambda_c)^2}\,\frac{1}{\eta}\sin^2\frac{\pi x}{a}\,dx\,dy = ab\,\frac{\lambda_g}{\lambda_o}. \tag{13}$$

Współczynniki $\hat{A}_1$ i $\hat{B}_1$ [równ. (11) i (12)] pozwalają bezpośrednio obliczyć sprzężenie i kierunkowość.

$$C = 20 \log \frac{1}{|\hat{A}_1|}, \tag{14}$$

$$D = 20 \log \left| \frac{\hat{A}_1}{\hat{B}_1} \right|. \tag{15}$$

4. ANALIZA WŁAŚCIWOŚCI NIEKTÓRYCH SPRZĘGACZY KIERUNKOWYCH

W zasadzie dla każdego ze sprzęgaczy o małych aperturach sprzężenie i kierunkowość można obliczyć z równań (11) lub (12). W wielu wypadkach jest kłopotliwe i można posługiwać się prostszą metodą. Na wielkość uzyskiwanego sprzężenia mają wpływ różne parametry prowadnicy oraz elementu sprzęgającego. Wyrażenie na sprzężenie jest więc sumą tłumień (wyrażonych w dB) wynikających z wpływu poszczególnych czynników. Poniżej zostaną podane czynniki wpływające na wielkość sprzężenia dla sprzęgaczy z aperturami szczelinowymi i kołowymi oraz sposób uzyskiwania wyrażeń na sprzężenie i kierunkowość dla większej liczby otworów.

4.1. Sprzęgacz szczelinowy

Jak już wspomniano, można zaniedbać zdolności polaryzacyjne P i M_2 dla bardzo wąskich szczelin. Wyrażenie na sprzężenie (w dB) można przedstawić w postaci szeregu składników związanych z wymiarami szczeliny, jej usytuowaniem w falowodzie, wymiarami falowodów, oraz długością fali w falowodzie [6]. Dla obranej konfiguracji należy wybrać czynniki, które w danym wypadku mają wpływ, i dla uzyskania pełnego wyrażenia na sprzężenie dodać odpowiednie składniki podane poniżej. Tłumienie związane z wymiarami szczeliny położonej w szerszej ściance falowodu a_o wyrażone jest zależnością:

$$a_o = 10 \log \left[\left(\frac{B}{2Y_o} \right)^2 + 1 \right] \quad [\text{dB}], \tag{16}$$

$$\frac{B}{Y_0} = \frac{ab\lambda_g}{4\pi M_1} \left(1 - \frac{\lambda_c}{\lambda_0} \right)^2,$$

gdzie:

λ_c — częstotliwość rezonansowa otworu,

$\lambda_c = 2L - 0{,}546\, w$.

Składnik tłumienia zależnego od częstotliwości:

$$a_1 = 20 \log \frac{2a}{\lambda_g} \quad [\text{dB}]. \tag{17}$$

Tłumienie związane ze skończoną grubością ścianki łączącej falowody:

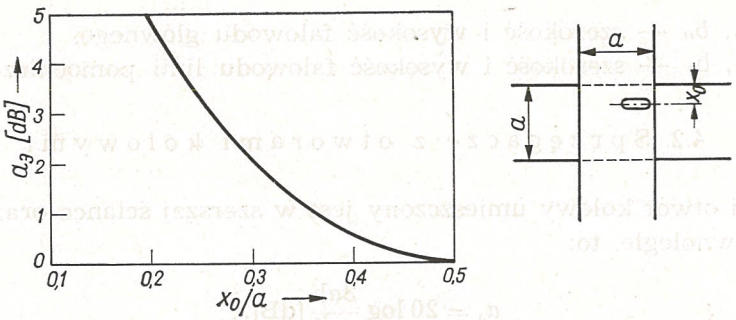
$$a_2 = 32\frac{t}{L}\sqrt{1-\left(\frac{1{,}71\,L}{\lambda_0}\right)^2} \quad [\text{dB}], \tag{18}$$

gdzie t — grubość ścianki.

Tłumienie wynikające z przesunięcia szczeliny na odległość x_o od brzegu falowodów:

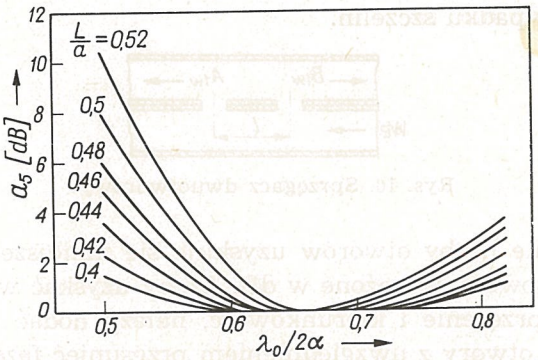
$$a_3 = 20\log\frac{1}{\sin\frac{\pi x_o}{a}} \quad [\text{dB}]. \tag{19}$$

Na rys. 8 podano wykres zależności (19).



Rys. 8. Tłumienie w funkcji przesunięcia szczeliny

Składnika tłumienia związanego z kątem, jaki tworzy szczelina z osią falowodu a_4, nie można wyrazić analitycznie. Dla kątów 15, 30, 45, 75 stopni zależność a_4 w funkcji $\frac{x_0}{a}$ przedstawiona została w artykule P. Sferrazzy [6]. Dla sprzęgaczy krzyżowych posiadających dwie szczeliny



Rys. 9. Tłumienie związane z przesunięciem fazy przez jedną ze szczelin w funkcji długości fali

umieszczone w przeciwległych rogach (rys. 10) oraz niektórych innych (np. sprzęgacz Schwingera) należy uwzględnić składnik związany z przesunięciem fazy o 180° przez jedną ze szczelin.

$$a_5 = 20 \log \frac{1}{\sin \beta l} \quad [\text{dB}], \tag{20}$$

gdzie l — odległość między środkami szczelin mierzona wzdłuż osi falowodu.

Zależność a_5 w funkcji $\lambda_0/2a$ pokazana została na rys. 9. Zależność od wymiarów falowodów wyraża tłumienie:

$$a_3 = 20 \log \frac{a_1 b_1}{b_0 a_1}, \tag{21}$$

gdzie:

a_0, b_0 — szerokość i wysokość falowodu głównego,
a_1, b_1 — szerokość i wysokość falowodu linii pomocniczej.

4.2. Sprzęgacze z otworami kołowymi

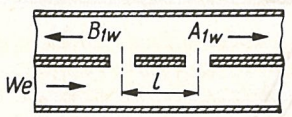
Jeżeli otwór kołowy umieszczony jest w szerszej ściance oraz falowody są równoległe, to:

$$a_0 = 20 \log \frac{3ab}{\pi d^3} \text{ [dB]}. \tag{22}$$

Dla otworu umieszczonego w węższej ściance:

$$a_0 = 20 \log \frac{6ab}{\pi d^3} \text{ [dB]}. \tag{23}$$

Pozostałe składniki wynikające ze zmian sprzężenia w funkcji częstotliwości (a_1), grubości ścianki (a_2 — zamiast L należy podstawić d), przesunięcia względem osi falowodu (a_3), wymiarów falowodu (a_6) są takie same jak w przypadku szczelin.



Rys. 10. Sprzęgacz dwuotworowy

Przez zwiększenie liczby otworów uzyskuje się mniejsze sprzężenie oraz większą kierunkowość, wyrażone w dB. Ażeby uzyskać w tym przypadku wyrażenia na sprzężenie i kierunkowość, należy dodać fale pobudzone przez wszystkie otwory z uwzględnieniem przesunięć fazowych. Na przykład dla sprzęgacza z dwoma otworami kołowymi umieszczonymi we wspólnej ściance równoległych falowodów (rys. 10) fala rozchodząca się

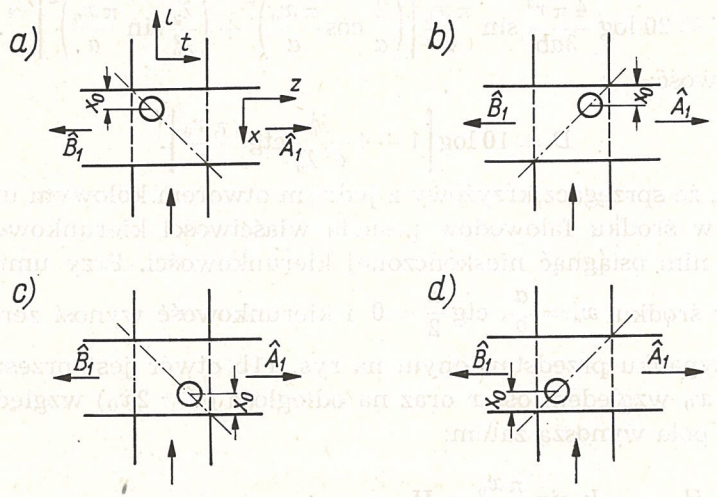
w prawo $\hat{A}_{1w}$) jest superpozycją dwu fal opóźnionych o βl, które są zatem w fazie i dodają się niezależnie od odległości l, natomiast fala rozchodząca się w lewo ($\hat{B}_{1w}$) jest superpozycją fali nie opóźnionej oraz opóźnionej o $2\beta l$; jeżeli $l = \lambda_{gl}$, fale te będą w przeciwfazie i odejmą się; uzyskuje się wtedy nieskończoną kierunkowość.

Zagadnienie to omówiono szczegółowo dla dwu i większej liczby otworów w literaturze [7], i nie będzie ono tutaj szerzej omawiane.

Dla sprzęgaczy krzyżowych z otworami kołowymi nie można znaleźć tak prostej metody projektowania. Sprzęgacze te są jednak chętnie i często stosowane w praktyce, gdyż poza prostotą wykonania posiadają stosunkowo dobre parametry elektryczne. Szczególnie dobre parametry ma sprzęgacz krzyżowy z trzema otworami odpowiednio rozmieszczonymi.

4.3. Sprzęgacz krzyżowy z otworami kołowymi

Na wstępie zostaną obliczone współczynniki określające sprzężenie i kierunkowość $\hat{A}_1$ i $\hat{B}_1$ [zob. równ. (12)] dla sprzęgacza z jednym otworem umieszczonym na przekątnej skrzyżowanych falowodów. Na rys. 11



Rys. 11. Możliwe położenia otworów sprzęgających o właściwościach kierunkowych w sprzęgaczu krzyżowym

pokazano możliwe położenie otworu (przypadek otworu położonego na środku skrzyżowania odrzucono ze względu na brak właściwości kierunkowych otworu tak umieszczonego).

Osie „l, t" związane są z linią główną, „x, y" z pomocniczą. Ponieważ otwory położone są na przekątnej falowodów, składowe pól są odpowiednio równe co do wartości bezwzględnej $|H_{1x}| = |H_{ot}|$ oraz $|H_{1z}| = |H_{ol}|$, natomiast mogą różnić się znakiem, w zależności od przesunięcia względem osi t i x.

Dla przypadku przedstawionego na rys. 11a składowe pól wynoszą:

$$H_{1x(a)} = H_{ot(a)} = k_1 \sin \frac{\pi x_0}{a},$$

$$H_{1z(a)} = H_{ol(a)} = k_2 \cos \frac{\pi x_0}{a},$$

gdzie: przez k_1 oznaczono: $k_1 = -\sqrt{1 - (\lambda_0/\lambda_c)^2}$,

przez k_2 : $k_2 = j \frac{\lambda_0}{\lambda_c}$.

Uwzględniając, że $M_1 = M_2 = 2P = \frac{4}{3} r^3$, można obliczyć współczynniki $\hat{A}_1$ i $\hat{B}_1$.

$$\left.\begin{aligned} \hat{A}_{1(a)} &= \frac{4\pi}{3} \frac{r^3}{ab} \left(\frac{2}{a} \sin \frac{\pi x_0}{a} \cos \frac{\pi x_0}{a} - j \frac{\lambda_g}{\lambda_0^2} \sin^2 \frac{\pi x_0}{a} \right) \\ \hat{B}_{1(a)} &= -j \frac{4\pi}{3} \frac{r^3}{ab} \frac{\lambda_g}{\lambda_0^2} \sin^2 \frac{\pi x_0}{a} \end{aligned}\right\} \quad (24a)$$

Stąd obliczone sprzężenie wynosi:

$$C = 20 \log \frac{4\pi r^3}{3ab} \sin \frac{\pi x_0}{a} \left[\left(\frac{2}{a} \cos \frac{\pi x_0}{a} \right)^2 + \left(\frac{\lambda_g}{\lambda_0^2} \sin \frac{\pi x_0}{a} \right)^2 \right]^{1/2}.$$

Kierunkowość:

$$D = 10 \log \left[1 + 4 \frac{\lambda_0^4}{a^2 \lambda_g^2} \operatorname{ctg}^2 \frac{\pi x_0}{a} \right].$$

Widać, że sprzęgacz krzyżowy z jednym otworem kołowym umieszczonym nie w środku falowodów posiada właściwości kierunkowe, ale nie można w nim osiągnąć nieskończonej kierunkowości. Przy umieszczeniu otworu w środku $x_0 = \frac{a}{2}$, $\operatorname{ctg} \frac{\pi}{2} = 0$ i kierunkowość wynosi zero.

W przypadku przedstawionym na rys. 11b otwór jest przesunięty na odległość x_0 względem osi x oraz na odległość $(a' - 2x_0)$ względem osi t, składowe pola wynoszą zatem:

$$H_{1x(b)} = k_1 \sin \frac{\pi x_0}{a} = H_{1x(a)},$$

$$H_{ot(b)} = k_1 \sin \pi \left(1 - \frac{x_0}{a} \right) = k_1 \sin \frac{\pi x_0}{a} = H_{ot(a)},$$

$$H_{ot(b)} = H_{1x(b)},$$

$$H_{1z(b)} = k_2 \cos \frac{\pi x_0}{a} = H_{1z(a)},$$

$$H_{ol(b)} = k_2 \cos \pi \left(1 - \frac{x_0}{a} \right) = -k_2 \cos \frac{\pi x_0}{a} = -H_{ol(a)},$$

$$H_{1z(b)} = -H_{ol(b)}.$$

Po podstawieniu do równania (12) otrzymuje się

$$\left.\begin{aligned} \hat{A}_{1(b)} &= -j\frac{4\pi r^3}{3ab}\frac{\lambda_g}{\lambda_0^2}\sin^2\frac{\pi x_0}{a}, \\ \hat{B}_{1(b)} &= \frac{4\pi r^3}{3ab}\left(\frac{2}{a}\sin\frac{\pi x_0}{a}\cos\frac{\pi x_0}{a} - j\frac{\lambda_g}{\lambda_0^2}\sin^2\frac{\pi x_0}{a}\right). \end{aligned}\right\} \tag{24b}$$

W przypadku przedstawionym na rys. 11c otwór przesunięty jest na odległość $(a\text{-}2x_0)$ względem obu osi x i t — składowe pola wynoszą:

$$H_{1x(c)} = k_1 \sin\pi\left(1-\frac{x_0}{a}\right) = k_1 \sin\pi\frac{x_0}{a} = H_{1x(a)},$$

$$H_{ot(c)} = k_1 \sin\pi\left(1-\frac{x_0}{a}\right) = k_1 \sin\frac{\pi x_0}{a} = H_{ot(a)},$$

$$H_{1z(c)} = k_2 \cos\pi\left(1-\frac{x_0}{a}\right) = -k_2\cos\frac{\pi x_0}{a} = -H_{1z(a)},$$

$$H_{ol(c)} = k_2 \cos\pi\left(1-\frac{x_0}{a}\right) = -k_2\cos\frac{\pi x_0}{a} = -H_{ol(a)},$$

$$H_{1z(c)} = H_{ol(c)} = -H_{ol(a)} = -H_{1z(a)}.$$

Zatem:

$$\left.\begin{aligned} \hat{A}_{1(c)} &= \frac{4\pi r^3}{3ab}\left(-\frac{2}{a}\sin\frac{\pi x_0}{a}\cos\frac{\pi x_0}{a} - j\frac{\lambda_g}{\lambda_0^2}\sin^2\frac{\pi x_0}{a}\right), \\ \hat{B}_{1(c)} &= \hat{B}_{1(a)}. \end{aligned}\right\} \tag{24c}$$

Przypadek pokazany na rys. 11d jest podobny do przypadku (b) z tym, że otwór jest przesunięty względem osi x o $(a\text{-}2x_0)$, a względem osi t o x_0. Wartości współczynników $\hat{A}_1$ i $\hat{B}_1$ wynoszą:

$$\left.\begin{aligned} \hat{A}_{1(d)} &= \hat{A}_{1(b)}, \\ \hat{B}_{1(d)} &= \frac{4\pi r^3}{3ab}\left(-\frac{2}{a}\sin\frac{\pi x_0}{a}\cos\frac{\pi x_0}{a} - j\frac{\lambda_g}{\lambda_0^2}\sin^2\frac{\pi x_0}{a}\right). \end{aligned}\right\} \tag{24d}$$

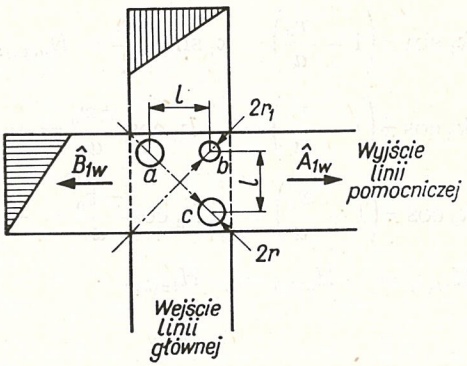
Sprzęgacze z pojedynczymi otworami umieszczonymi jak na rysunkach 11 b, c, d nie różnią się pod względem parametrów od sprzęgacza z rys. 11a; jedynie moc w sprzęgaczach z rys. 11b i 11d rozchodzi się w linii pomocniczej w przeciwnym kierunku niż w sprzęgaczach z rys. 11a i 11c.

Dla sprzęgacza z jednym otworem nie można uzyskać silnego sprzężenia przy dobrej kierunkowości. Powiększenie średnicy otworu, a tym samym konieczność przesuwania go w kierunku środka falowodów powoduje malenie kierunkowości sprzęgacza. Dlatego też bardziej skutecznym środkiem realizacji silniejszych sprzężeń jest zwiększenie liczby otworów. Dla sprzęgacza dwuotworowego (otwory a i c lub b i d) można

uzyskać większe sprzężenie przy kierunkowości takiej samej jak dla sprzęgacza jednootworowego. Zasadniczą poprawą kierunkowości (teoretyczna możliwość uzyskania nieskończonej kierunkowości) uzyskuje się przez dodanie trzeciego otworu odpowiednio umieszczonego i o odpowiedniej średnicy. Okazuje się, że korzystne jest wybranie otworów a, b, c, (rys. 11).

Obecnie zostaną wyprowadzone wzory na sprzężenie i kierunkowość sprzęgacza przedstawionego na rys. 12.

Fale dochodzące do otworów a i b są opóźnione względem fali dochodzącej do otworu c o kąt fazowy βl. W linii pomocniczej fala rozchodząca się w kierunku wyjścia pobudzona przez otwór a jest opóźniona o βl



Rys. 12. Trójotworowy sprzęgacz krzyżowy

względem fal pobudzanych przez otwory b i c, natomiast dla przeciwnego kierunku propagacji w linii pomocniczej opóźnione są fale pobudzane przez b i c. W celu znalezienia superpozycji należy dodać poszczególne fale z uwzględnieniem obu przesunięć fazowych (w linii głównej i pomocniczej). Niech $\hat{A}_{1w}$ oznacza wypadkową falę rozchodzącą się w kierunku wyjścia linii pomocniczej, a $\hat{B}_{1w}$ — wypadkową falę w kierunku obciążenia, wtedy:

$$\left.\begin{aligned} \hat{A}_{1w} &= \hat{A}_{1(a)} e^{-j2\beta l} + \hat{A}_{1(b)} e^{-j\beta l} + \hat{A}_{1(c)}, \\ \hat{B}_{1w} &= \hat{B}_{1(a)} e^{-j\beta l} + \hat{B}_{1(b)} e^{-j2\beta l} + \hat{B}_{1(c)} e^{-j\beta l}. \end{aligned}\right\} \tag{25}$$

Ażeby znaleźć sprzężenie i kierunkowość, trzeba obliczyć moduły wyrażeń $\hat{A}_{1w}$ i $\hat{B}_{1w}$. Najłatwiej i najbardziej obrazowo można to wykonać przez sumowanie wektorów przedstawiających odpowiednio współczynniki $\hat{A}_1$ i $\hat{B}_1$ na płaszczyźnie zmiennej zespolonej.

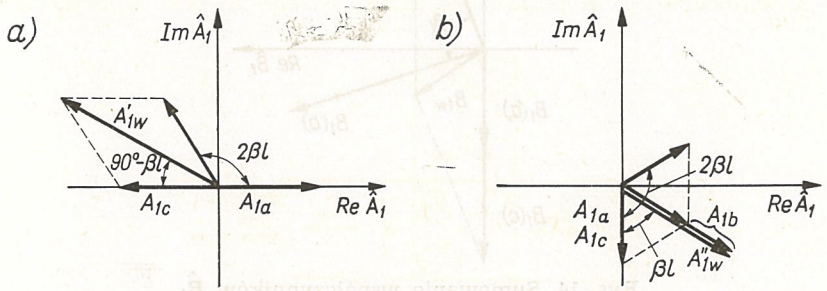
Wektory na rys. 13 przedstawiają odpowiednio części rzeczywiste i urojone współczynników $\hat{A}_1$ [wyrażonych wzorem (24)] oraz wektory wypadkowe powstałe przez zsumowanie składowych po uwzględnieniu ich przesunięć fazowych.

Otrzymuje się:

$$|\hat{A}'_{1w}| = 2\,Re\,\hat{A}_{1(a)} \sin\beta l,$$
$$|\hat{A}''_{1w}| = 2\,Im\,\hat{A}_{1(a)} \cos\beta l + |\hat{A}_{1(b)}|,$$

gdzie: $\hat{A}'_{1w}$ oznacza sumę części rzeczywistych, $\hat{A}''_{1w}$ sumę części urojonych. Wektory $\hat{A}'_{1w}$ i $\hat{A}''_{1w}$ leżą na jednej prostej i są przeciwnie skierowane, zatem moduł ich sumy wynosi:

$$|\hat{A}_{1w}| = |\,2\,Re\,\hat{A}_{1(a)} \sin\beta l - (2\,Im\,\hat{A}_{1(a)} \cos\beta l + |\hat{A}_{1(b)}|)|.$$



Rys. 13. Sumowanie współczynników $\hat{A}_1$

Podstawiając wartości z równań (24) otrzymuje się wyrażenie na sprzężenie dla trójotworowego sprzęgacza krzyżowego:

$$C = 20 \log \frac{3ab\,\lambda_g}{4\pi s \lambda_0^2} \cdot \frac{1}{2r^3(p - s\cos\beta l) - r_1^3 s}, \tag{26}$$

gdzie:

$$p = \frac{2}{a} \cos\frac{\pi x_0}{a} \sin\beta l,$$

$$s = \frac{\lambda_g}{\lambda_0^2} \sin\frac{\pi x_0}{a}.$$

Z równania (26) wynika dodatkowy wniosek, że trzeci otwór zwiększa sprzężenie, jednakże w praktyce, ponieważ $r_1 < r$ (co zostanie później uzasadnione), zmiana sprzężenia jest niewielka (około 0,4 dB).
Można wykazać, że najsilniejsze sprzężenie występuje dla

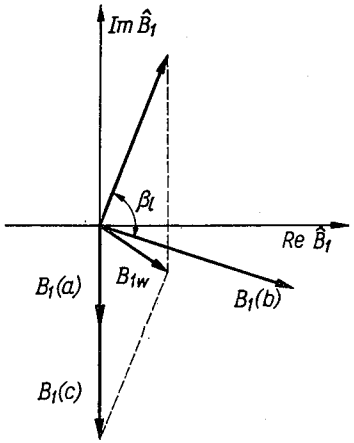
$$\beta l \approx \frac{2}{3}\pi,$$
$$l \approx \frac{\lambda_{g0}}{3}, \tag{27}$$

gdzie jako λ_{g0} przyjmuje się

$$\lambda_{g0} = \frac{2\lambda_{g1}\lambda_{g2}}{\lambda_{g1} + \lambda_{g2}}, \tag{28}$$

gdzie λ_{g1}, λ_{g2} są długościami fali w falowodzie dla skrajnych częstotliwości pracy sprzęgacza.

Kierunkowość można obliczyć na podstawie rys. 14. Aby uzyskać teoretyczną wartość kierunkowości nieskończoność, wektor B_1(b) powinien



Rys. 14. Sumowanie współczynników $\hat{B}_1$

być przeciwnie skierowany niż wektory $B_1(a)$ i $B_1(c)$ oraz równy ich sumie. Pozwala to na określenie wielkości promienia r_1 mniejszego otworu.

Z warunku równości wektorów:

$$\frac{8\pi r^3 \lambda_g}{3ab\lambda_0^2} \sin^2 \frac{\pi x_0}{a} = \frac{4\pi r_1^3}{3ab} \sin \frac{\pi x_0}{a} \left[\frac{4}{a^2} \cos^2 \frac{\pi x_0}{a} + \frac{\lambda_g^2}{\lambda_0^4} \sin \frac{\pi x_0}{a} \right]^{1/2}$$

przyjmując $\lambda_{g0} = 2a$, co odpowiada $\lambda_0 = \sqrt{2}\, a$, otrzymuje się po podstawieniu:

$$r = 1{,}54 r_1.$$

Można więc przyjąć dla prostoty $r_1 = 2/3\, r$.

$$|\hat{B}_{1w}| = 4|\hat{B}_{1(a)}| \cos \frac{\beta l + \operatorname{arctg}\left(\frac{2\lambda_0^2}{a\lambda_g} \operatorname{ctg} \frac{\pi x_0}{a}\right)}{2}, \tag{29}$$

$$|\hat{B}_{1w}| \cong 4|\hat{B}_{1(a)}| \cos\left(\frac{\pi}{6} - \frac{\beta l}{2}\right).$$

Warunek maksymalnej kierunkowości spełniony zostanie dla:

$$\frac{\beta l}{2} \approx \frac{\pi}{2} - \frac{\pi}{6}, \tag{30}$$

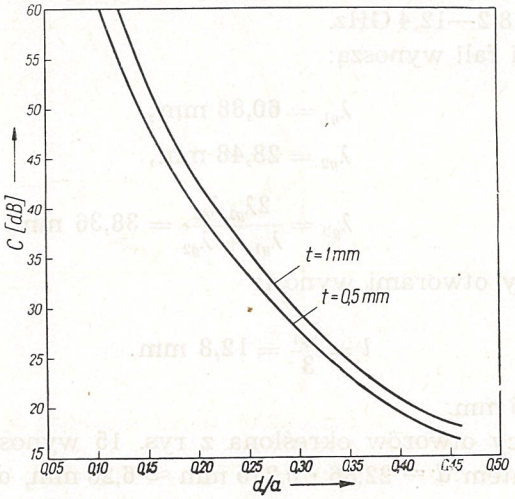
czyli

$$l \approx 0{,}42\lambda_g.$$

Warunek (30) jest spełniony dla wyższych częstotliwości pasma. Jeżeli odległość l wybrana zostanie zgodnie z warunkiem (27), co zapewni najmniejsze zmiany sprzężenia w funkcji częstotliwości, to należy się liczyć

z poprawą kierunkowości sprzęgacza trójotworowego w stosunku do dwuotworowego jedynie dla górnych częstotliwości pasma.

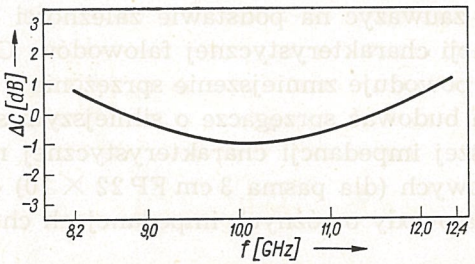
Przeprowadzone rozważania dotyczące warunku najsilniejszego sprzężenia oraz maksymalnej kierunkowości są przybliżone i mają charakter jedynie orientacyjny. Dokładna analiza prowadziłaby do bardzo skom-



Rys. 15. Sprzężenie w funkcji unormowanej średnicy otworów sprzęgających dla sprzęgacza krzyżowego

plikowanych zależności matematycznych, dla celów praktycznych wyprowadzone zależności są wystarczająco dokładne.

Na rys. 15 przedstawiono zależność sprzężenia od unormowanej średnicy otworu dla sprzęgaczy krzyżowych o trzech otworach, dla dwu grubości wspólnej ścianki [na podstawie (26)]. Ponieważ zależność została



Rys. 16. Zmiany sprzężenia w funkcji częstotliwości sprzęgaczy na pasmo 3 cm

podana dla unormowanej średnicy otworu, może służyć do projektowania sprzęgaczy na dowolne pasmo częstotliwości. Dokładność określenia sprzężenia na podstawie wykresu (przy zachowaniu odpowiednich tolerancji wykonania mechanicznego) wynosi $\pm 0{,}2$ dB. Zmiany sprzężenia w pasmie pracy sprzęgacza można znaleźć na podstawie zależności (26). Dla pasma 3 cm zależność ta została podana na rys. 16.

5. METODA PROJEKTOWANIA KRZYŻOWEGO SPRZĘGACZA O TRZECH OTWORACH KOŁOWYCH

W celu wykazania praktycznych możliwości wykorzystania przedstawionej powyżej analizy teoretycznej przedstawiono przykład liczbowy występujący przy projektowaniu sprzęgacza krzyżowego o sprzężeniu 30 dB na pasmo 8,2—12,4 GHz.
Skrajne długości fali wynoszą:

$$\lambda_{g1} = 60{,}88 \text{ mm},$$

$$\lambda_{g2} = 28{,}48 \text{ mm},$$

stąd

$$\lambda_{g0} = \frac{2\lambda_{g1}\lambda_{g2}}{\lambda_{g1} + \lambda_{g2}} = 38{,}36 \text{ mm}.$$

Odległość między otworami wynosi:

$$l = \frac{\lambda_{g0}}{3} = 12{,}8 \text{ mm}.$$

Wybrano $t = 0{,}5$ mm.
Wielkość średnicy otworów określona z rys. 15 wynosi: dla $C = 30$ dB, $d/a = 0{,}275$, a zatem $d = 22{,}86 \cdot 0{,}275$ mm $\approx 6{,}28$ mm, $d_1 \approx 4{,}18$ mm.

5.1. Zmodyfikowany sprzęgacz krzyżowy o sprzężeniu 10 dB

Jak widać z rys. 15, minimalne sprzężenie możliwe do uzyskania wynosi ok. 17 dB. Dla innego typu sprzęgaczy krzyżowych uzyskuje się tego samego rzędu minimalne sprzężenie. Natomiast w bardzo wielu zastosowaniach byłoby korzystne posiadanie krzyżowego sprzęgacza o sprzężeniu 10 dB. Można zauważyć na podstawie zależności (26), że sprzężenie zależy od impedancji charakterystycznej falowodów. Obniżenie obu falowodów do połowy powoduje zmniejszenie sprzężenia o 3 dB. Wynika stąd wniosek, że można budować sprzęgacze o silniejszym sprzężeniu stosując falowody o mniejszej impedancji charakterystycznej niż impedancja falowodów standartowych (dla pasma 3 cm FP 22 × 10) oraz transformatory dopasowujące falowody o różnych impedancjach charakterystycznych.

6. WYNIKI DOŚWIADCZALNE

Wykonano oraz zbadano sprzęgacze krzyżowe o trzech otworach kołowych zaprojektowane wg podanej metody. Na rys. 17 pokazano przebiegi sprzężenia w funkcji częstotliwości dla sprzęgaczy 30 dB i 20 dB, pracujących w paśmie 8,2—12,4 GHz. Linią przerywaną zaznaczono spodziewany przebieg sprzężenia na górze pasma, niemożliwy do doświad-

czalnego sprawdzenia ze względu na brak odpowiedniego wyposażenia. Jak widać, zmiany sprzężenia w paśmie wynoszą około ± 1 dB.

Pozostałe parametry tych sprzęgaczy:

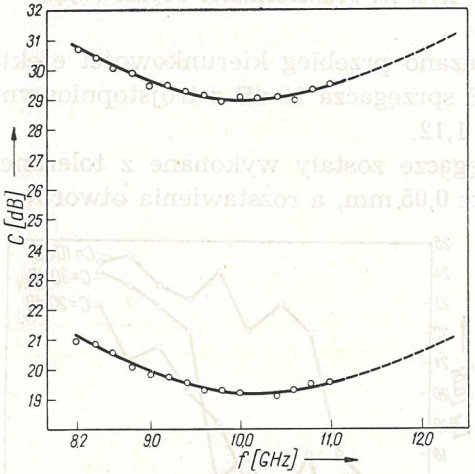
Kierunkowość powyżej 15 dB

WFS w linii głównej i pomocniczej $\leqslant 1{,}05$

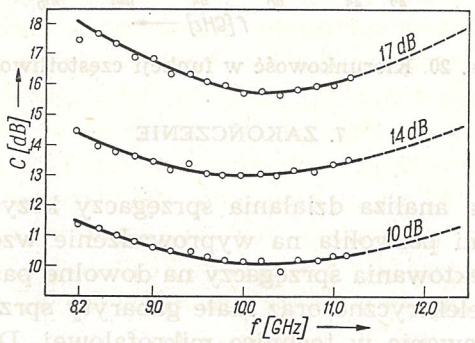
Dokładność określenia sprzężenia około ± 0,2 dB.

Wymiary zewnętrzne 80 × 80 mm.

Na rys. 18 pokazano przebieg sprzężenia sprzęgacza o maksymalnie dużych otworach (ponieważ $l = 12{,}8$ mm, $d_{max} = 10$ mm) na falowodzie



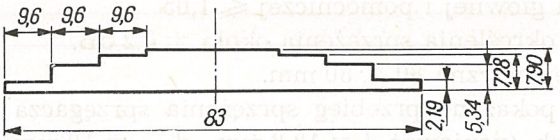
Rys. 17. Sprzężenie w funkcji częstotliwości sprzęgaczy 20 i 30 dB. Kółkami oznaczono wielkości zmierzone



Rys. 18. Sprzężenie w funkcji częstotliwości sprzęgaczy 17, 14 i 10 dB

standartowym oraz obniżonym do połowy. Okazuje się, że przy cztero i pół-krotnym obniżeniu obu falowodów uzyskuje się sprzężenie około 10 dB (trzecia krzywa na rys. 18). Zastosowano transformatory Czebyszewa dopasowujące falowody o różnych impedancjach charakterystycznych obliczone wg [8], szkic transformatora na pasmo 3 cm pokazano na rys. 19.

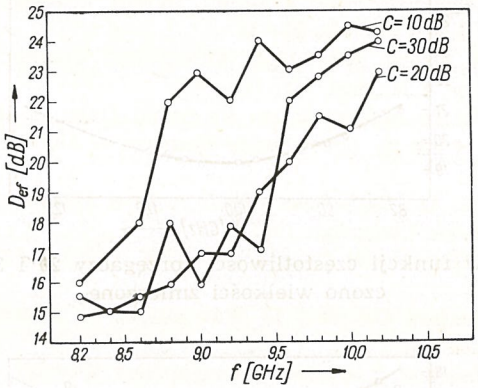
Dodatkową zaletą sprzęgacza na falowodzie obniżonym jest bardziej płaski przebieg sprzężenia w funkcji częstotliwości niż dla sprzęgacza na falowodzie standartowym. Zmiany sprzężenia wynoszą odpowiednio $(17 \pm {}^{0,8}_{1,2})$ dB, $(14 \pm {}^{0,5}_{1,0})$ dB oraz $(10 + 1{,}4)$ dB.



Rys. 19. Transformator dopasowujący

Na rys. 20 pokazano przebieg kierunkowości efektywnej w funkcji częstotliwości. WFS sprzęgacza 10 dB z trójstopniowym transformatorem dopasowującym $\leqslant 1{,}12$.

Wszystkie sprzęgacze zostały wykonane z tolerancjami dla otworów i grubości ścianki $\pm$ 0,05 mm, a rozstawienia otworów $\pm$ 0,1 mm.



Rys. 20. Kierunkowość w funkcji częstotliwości

7. ZAKOŃCZENIE

Przeprowadzona analiza działania sprzęgaczy krzyżowych z trzema otworami kołowymi pozwoliła na wyprowadzenie wzorów i wykonanie wykresów do projektowania sprzęgaczy na dowolne pasmo częstotliwości. Dobre parametry elektryczne oraz małe gabaryty sprzęgaczy zapewniają im szerokie zastosowanie w technice mikrofalowej. Dodatkowym rezultatem pracy jest skonstruowanie sprzęgacza krzyżowego o sprzężeniu 10 dB.

Badania doświadczalne przeprowadzone w pasmie 3 cm potwierdzają słuszność przeprowadzonych rozważań i wykazują dużą zgodność z wynikami obliczonymi teoretycznie.

Dodatkową ważną sprawą, na którą należy zwrócić uwagę przy konstrukcji sprzęgaczy, jest dokładne wykonanie mechaniczne oraz dobrze

dopasowane obciążenia w linii głównej i pomocniczej. Dla omawianego sprzęgacza zmiana stosunku d/a o 1% powoduje zmianę sprzężenia o około 0,35 dB. W pasmie 3 cm zmiana sprzężenia wynosi 4 dB/mm — na tej podstawie przyjęto konieczne tolerancje wykonania otworów (rozdział 6). Rozstawienie otworów ma o dwa rzędy wielkości mniejszy wpływ. Natomiast zmiana sprzężenia przy zmianie grubości ścianki wynosi od około 2,5 dB/mm dla dużych średnic otworów $\left(\frac{d}{a} > 0{,}4\right)$ do około 5 dB/mm dla małych średnic otworów $\left(\frac{d}{a} < 0{,}2\right)$.

Politechnika Warszawska
Katedra Radiolokacji

WYKAZ LITERATURY

1. Anderson T. E.: Directional Coupler Design Nomograms. — Microwave Journal, vol. 2, nr 5 V, 1959, str. 34—38.
2. Andrews P.: A Simple Waveguide Directional Coupler. — Journal of the British IRE, t. 15, nr 2 II, 1955, str. 112—116.
3. Bethe H. A.: Theory of Diffraction by Small Holes. — Physc. Rev., 66, 1944.
4. Gerlack R. Z.: 10 dB X_L Cross Guide Coupler. — IRE Trans. MTT-9, nr 6 XI, 1961, str. 571.
5. Montgomery C. G.: Technique of Microwave Measurements. — MIT, t. 11, Mc Graw — Hill Book Company, 1947.
6. Sferrazza P.: Directional Couplers. — IRE, Conv. Rec., vol. 2, part 8, 1954, str. 115—124.
7. Surdin M.: Directive Couplers in Waveguides. — Journal Instn. Elect. Eng., t. 93, Part III A, 1946, str. 725.
8. Young L.: Tables for Cascaded Homogeneous Quarter — Wave Transformers. — IRE Trans. MTT IV, 1959, str. 233.
9. M. Panecki, R. Litwin, L. Drozdowicz: Teoria i technika mikrofal. Wydawnictwa Naukowo-Techniczne, Warszawa.

M. RZEPECKA

CROSS GUIDE DIRECTIONAL COUPLERS

Summary

The paper presents the review of solutions of directional couplers, particularly cross guide couplers which are here under special consideration.

Parameters and applications of couplers are discussed. The comparison of properties of various units is made. The analysis of the operation of couplers having small coupling apertures compared with wavelenght is given. Formulas for designing slot couplers are given. Formulas for cross guide couplers with round holes are derived. The particular attention is paid to cross guide couplers with three

round holes because of their advantages. This type of coupler has not been discussed in detail up to this time in literature.

The calculated diagram allows to design such kinds of couplers for any frequency band.

The experiments in 3 cm band were done, the obtained results were in accordance with the theoretical assumption. The paper presents also the parameters of constructed, 10, 20, and 30 dB couplers. Their directivity over the whole 3 cm band is more than 15 dB, for 20 and 30 dB couplers VSWR is less than 1.05 and for 10 dB unit VSWR is less than 1.1. The changes of the coupling factor over the band are at most ± 1 dB.

M. RZEPECKA

LES COUPLEURS DIRECTIFS CROISÉS

Résumé

L'auteur présente dans cet article une revue des solutions de coupleurs directifs (en prenant particulièrement en consideration les coupleurs croisés), les performances, les applications ainsi que la comparaison des propriétés de différents types des coupleurs. On analyse le fonctionnement des coupleurs á ouverture de couplage de faibles dimensions par rapport à la longueur d'onde.

On donne aussi les formules principales pour le calcul d'un coupleur à fente courte et étroite, et celles d'un coupleur croisé à ouverture circulaire. Des coupleurs croisés à trois ouvertures circulaires sont considérés plus particulièrement. L'abaque ci-joint permet de calculer un coupleur de ce type d'une bande de fréquence quelconque.

Une expérience effectuée dans la bande de 3 cm est d'accord avec la théorie. On donne aussi les performances des coupleurs à 10, 20 et 30 dB. L'effet directif est supérieur à 15 dB dans la bande entière, taux d'ondes stationnaires pour les coupleurs á 20, 30 dB et 10 dB sont respectivement égaux á 1,05 et 1,1. Les variations de couplage dans la bande considerée sont inférieures á ± 1 dB.

M. RZEPECKA

RICHTKREUZKOPPLUNG

Zusammenfassung

In diesem Aufsatz wurde ein Überblick über bestehende Richtkopplungen mit besonderer Berücksichtigung der Kreuzkopplungen angegeben. Es wurden Parameter behandelt, ihre Anwendung erwogen und ein Vergleich der Eigenschaften verschiedener Kopplungstypen angestellt.

Hierbei wurde auch die Wirkung der Kopplungen samt Kopplungsapparatur kleiner Ausmessungen im Vergleich mit der Wellenlänge analysiert. Auf dieser Grundlage wurden Muster für die Projektierung von Kopplungen kurzer und schmaler Streifen angegeben. Ferner wurden Formaln für Kreuzkopplungen mit Kreisöffnungen ermittelt.

Eine besondere Beachtung widmete man Kreuzkopplungen mit drei Kreisöffnungen, und zwar mit Rücksicht auf eine Reihe von Vorzügen, die sich aus einer solchen Lösung ergeben, da es in der Fachlitteratur an einer diesbezüglichen Abhandlung fehlt.

Ein ausgearbeitetes Diagramm gestattet es, Kopplungen dieser Type für ein beliebiges Frequenzband zu projektieren. Die in 3-cm-Band durchgeführten Versuche bestätigen die theoretischen Voraussetzungen. Es wurden ferner Parameter für die Herstellung folgender Kopplungen, und zwar für 10, 20 und 30 dB angegeben. Ihre Richtlinie ist im ganzen Band grösser als 15 dB. Für Kopplungen 20 und 30 dB beträgt der Koeffizient der festen Welle $\leqslant 1.05$ im ganzen 3 cm Band, und für eine Kopplung von 10 dB $\leqslant 1.1$. Hierbei sind Kopplungsänderungen im ganzen Band nicht grösser als $\pm$ 1 dB.

М. ЖЕПЕЦКА

ПЕРЕКРЁСТНЫЕ НАПРАВЛЕННЫЕ ОТВЕТВИТЕЛИ

Резюме

В работе дан обзор существующих решений направленных ответвителей со специальным рассмотрением перекрёстных ответвителей.

Сначала коротко рассмотрены параметры, применения и сравнение свойств различных типов ответвителей. Дан анализ работы ответвителей со связывающими апертурами, размеры которых малы по сравнению с длиной волны, и на основании этого анализа даны формулы для проектирования ответвителей с короткими и узкими щелями, а также выведены формулы для перекрёстных ответвителей с круглыми отверствиями. Особенное внимание уделено перекрёстным ответвителям с тремя круглыми отверствиями, этот вариант связан с многочисленными преимуществами. В литературе нет точного анализа этой конструкции.

Результаты эксперимента в диапазоне 3 см подтвердили теоретические предвидения. Даны также параметры конкретных ответвителей 10, 20 и 30 дб. Их направленность больше чем 15 дб, а согласование в целом диапазоне для ответвителей 20 и 30 дб такое что К.С.В. $\leqslant$ 1,05, а для 10 дб К.С.В. $\leqslant$ 1,1. Изменение связи в диапазоне не больше чем $\pm$ 1 дб.

621.396.677.7

ANDRZEJ KRASZEWSKI

Sprzęgacz falowodowy o dużej kierunkowości

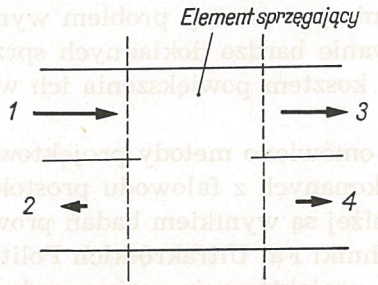
Rękopis dostarczono 3.12.1962

Opisano metodę projektowania wielootworowych sprzęgaczy kierunkowych o dużej kierunkowości, opartą na wykorzystaniu właściwości wielomianów Czebyszewa. Jeżeli współczynniki sprzężenia poszczególnych otworów są proporcjonalne do współczynników wielomianu Czebyszewa, można uzyskać maksymalną kierunkowość w maksymalnie szerokim pasmie częstotliwości. Omówiono sposoby nakładania szeregów sprzęgających.

Przykładem wykorzystania omówionych metod projektowania jest opisany sprzęgacz kierunkowy o sprzężeniu $10 \pm 0{,}5$ dB i kierunkowości ponad 40 dB w pasmie 8,4—11,5 GHz.

1. WPROWADZENIE

Mianem sprzęgacza kierunkowego określa się w technice mikrofalowej dopasowane połączenie czterech prowadnic falowych, charakteryzujące się tym, że energia dostarczana do jednej z prowadnic dzieli się pomiędzy



Rys. 1. Schemat sprzęgacza kierunkowego

dwie prowadnice, zaś czwarta prowadnica jest odizolowana. Można wykazać, że właściwość taką posiada każde dopasowane połączenie czterech dowolnych prowadnic falowych.

Parametrami charakteryzującymi sprzęgacz kierunkowy jest sprzężenie, kierunkowość oraz niedopasowanie wprowadzane przezeń do toru mikrofalowego. Na rys. 1 przedstawiono ogólny schemat sprzęgacza kie-

runkowego. Jeżeli układ jest zasilany od strony końcówek *1*, to sprzężenie można wyrazić jako

$$C = 10 \log \frac{P_1}{P_4} \text{ [dB]}, \tag{1}$$

kierunkowość

$$D = 10 \log \frac{P_4}{P_2} \text{ [dB]}, \tag{2}$$

zaś izolację pomiędzy przyległymi końcówkami *1* i *2* jako

$$I = 10 \log \frac{P_1}{P_2} = C + D \text{ [dB]}. \tag{3}$$

Miarą niedopasowania układu jest współczynnik odbicia w gałęzi *1* oraz gałęzi *4*, gdy pozostałe gałęzie są zakończone dopasowanymi obciążeniami. Prowadnicę falową pomiędzy gałęziami *1* i *3* nazywamy linią główną, zaś prowadnicę pomiędzy gałęziami *2* i *4* linią pomocniczą. Tak więc parametrami sprzęgacza jest niedopasowanie w linii głównej oraz w linii pomocniczej. W praktyce interesująca jest również zmiana wszystkich parametrów sprzęgacza w pasmie częstotliwości pracy jak również rozrzut wartości sprzężenia od sprzężenia nominalnego.

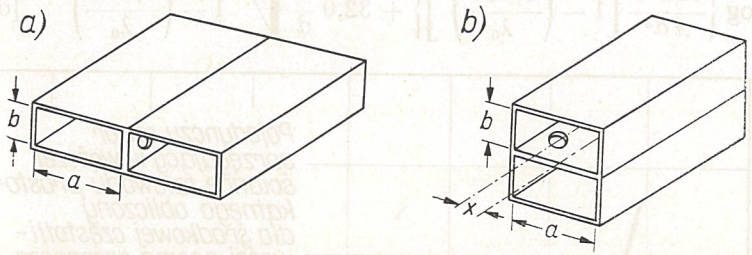
Zależnie od konkretnych wymagań pożądane są różne wartości liczbowe parametrów. W wielu przypadkach potrzebne są sprzęgacze o niewielkim współczynniku sprzężenia, w innych przypadkach najistotniejszym parametrem jest duża kierunkowość układu. W układach laboratoryjnych ważne jest poza tym, aby wartość sprzężenia nie zmieniała się zbyt znacznie przy dużych zmianach częstotliwości. Ponieważ w warunkach laboratoryjnych nie jest istotny problem wymiarów gabarytowych, możliwe jest wykonywanie bardzo dokładnych sprzęgaczy kierunkowych o dużej kierunkowości kosztem powiększenia ich wymiarów geometrycznych.

W pracy niniejszej omówiono metody projektowania sprzęgaczy o dużej kierunkowości, wykonanych z falowodu prostokątnego. Pewne uogólnienia przytoczone poniżej są wynikiem badań prowadzonych w tym kierunku w Katedrze Techniki Fal Ultrakrótkich Politechniki Warszawskiej. Przedstawione metody projektowania można wykorzystać przy budowie podobnych sprzęgaczy kierunkowych na dowolne pasma częstotliwości.

2. APERTURY SPRZĘGAJĄCE

Sprzężenie pomiędzy dwoma falowodami prostokątnymi można zrealizować za pomocą różnego rodzaju apertur sprzęgających umieszczonych albo w węższej, albo w szerszej ściance falowodu. Apertury te mogą mieć kształt szczelin o różnych proporcjach i różnie nachylonych do

osi falowodu. Jednakże najprostszym przypadkiem, zarówno do wykonania mechanicznego, jak i analizy teoretycznej, są otwory kołowe, które poza tym pozwalają prosto tworzyć skomplikowane szeregi sprzęgające wielootworowe jak i kombinacje tych szeregów. Z tego względu w pracy niniejszej ograniczono się wyłącznie do rozważenia sprzęgaczy z kołowymi otworami sprzęgającymi.



Rys. 2. Sprzężenie dwu falowodów prostokątnych za pomocą apertur kołowych a) otwór w węższej ściance, b) otwór w szerszej ściance

Na rys. 2 przedstawiono dwa możliwe sposoby sprzężenia dwu falowodów prostokątnych za pomocą apertur kołowych. Wyrażenie na sprzężenie uzyskiwane z jednym otworem umieszczonym we wspólnej ściance pomiędzy dwoma falowodami prostokątnymi połączonymi węższymi bokami ma postać:

$$C = 20 \log\left(\frac{6a^3 b}{\pi d^3 \lambda_g}\right) + 32{,}0\,\frac{t}{d}\sqrt{1-\left(\frac{1{,}71\,d}{\lambda_0}\right)^2} + 6{,}0\,[\mathrm{dB}], \qquad (4)$$

gdzie:

a — wewnętrzny szerszy wymiar falowodu,
b — wewnętrzny węższy wymiar falowodu,
d — średnica otworu sprzęgającego,
t — grubość ścianki w obszarze otworu,
λ_0 — długość fali w wolnej przestrzeni,
λ_g — długość fali w falowodzie.

Równanie to jest wyrażeniem na sprzężenie dwuotworowego sprzęgacza w węższej ściance, wyprowadzonym przez Bethego. Aby ograniczyć wyrażenie do jednego tylko otworu, dodano wartość 6 dB.

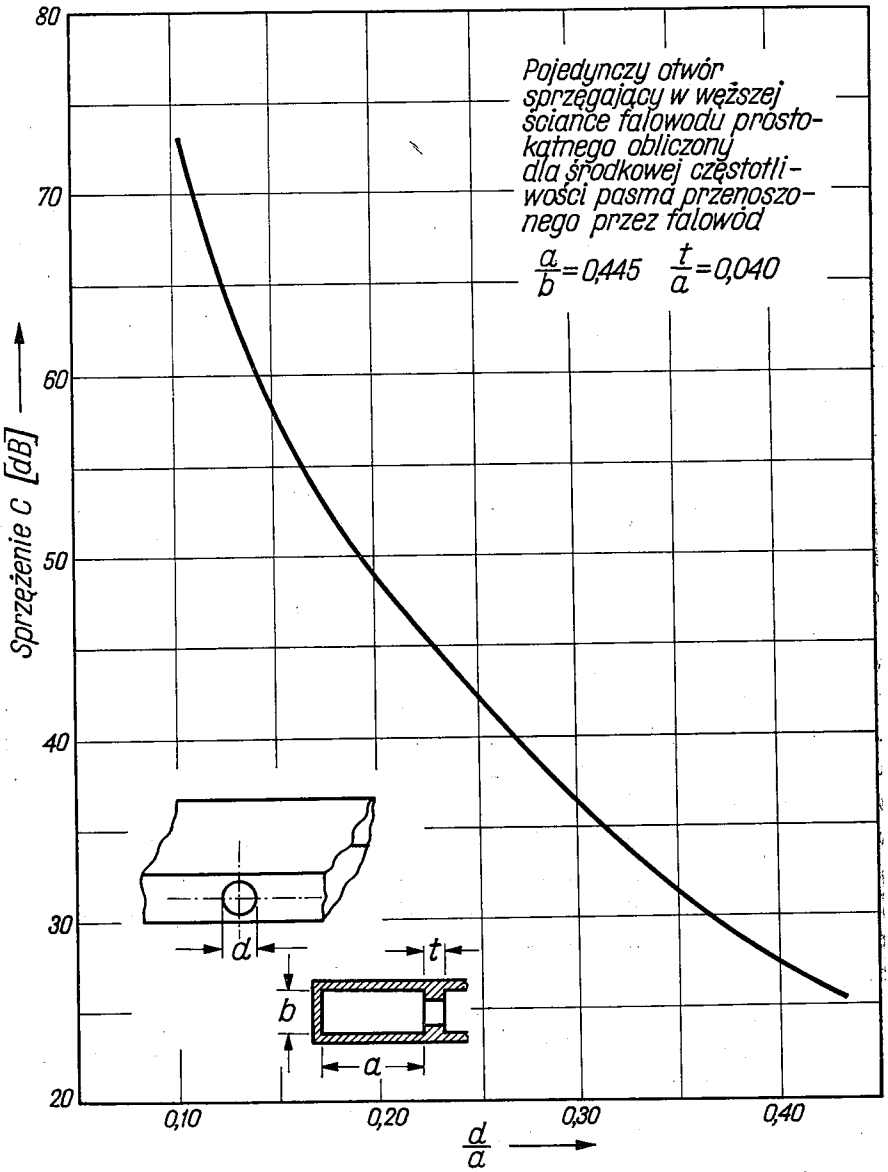
Równanie na wielkość sprzężenia dla jednego otworu umieszczonego w szerszej ściance dwu połączonych ze sobą falowodów prostokątnych (rys. 2b) ma postać

$$C = 20 \log\left[\frac{6a^2 b}{\pi d^3}\,\frac{1}{\sin^2\frac{\pi x}{a}}\left\{1-\left(\frac{1{,}71\,d}{\lambda_0}\right)^2\right\}\right] + 32\,\frac{t}{d}\sqrt{1-\left(\frac{1{,}71\,d}{\lambda_0}\right)^2}\,[\mathrm{dB}], \qquad (5)$$

gdzie wszystkie oznaczenia są identyczne jak w równaniu (4), a x jest odległością pomiędzy osią otworu a osią wzdłużną falowodu. Ponieważ sprzężenie zmienia się w funkcji poprzecznego położenia otworu sprzę-

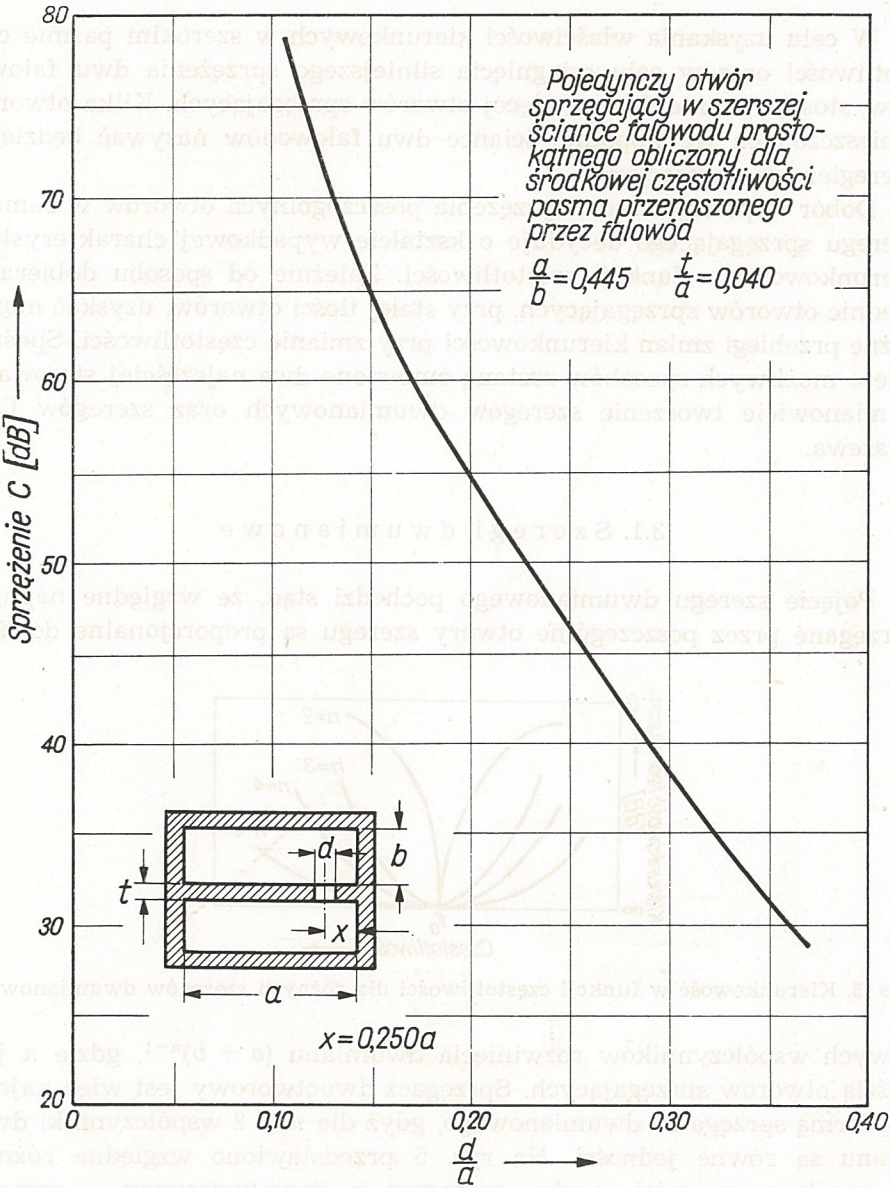
gającego, można znaleźć położenie odpowiadające minimalnej selektywności. Optymalne położenie określono matematycznie jako w przybliżeniu równe połowie odległości pomiędzy osią falowodu a wewnętrzną powierzchnią bocznej ścianki ($x = 0{,}25\,a$) i wtedy równanie (5) przybiera postać

$$C = 20 \log \left\{ \frac{12a^2 b}{\pi d^3} \left[1 - \left(\frac{1{,}71\,d}{\lambda_0} \right)^2 \right] \right\} + 32{,}0 \frac{t}{d} \sqrt{1 - \left(\frac{1{,}71\,d}{\lambda_0} \right)^2} \quad \text{[dB]}. \qquad (6)$$



Rys. 3. Zależność współczynnika sprzężenia od średnicy otworu umieszczonego w węższej ściance falowodu

Wybór któregoś z dwu przedstawionych sposobów sprzężenia wynika z wymagań, jakie musi spełniać sprzęgacz kierunkowy. Otwory sprzęgające umieszczone w węższej ściance nie ograniczają wytrzymałości falowodu na przebicie, zaś otwory w szerszej ściance pozwalają uzyskać dużo mniejszą selektywność. Należy pamiętać, że równania (4), (5) i (6) wyprowadzono przy założeniu $\frac{d}{\lambda_0} \ll 1$, a więc dla małych otworów. Obli-



Rys. 4. Zależność współczynnika sprzężenia od średnicy otworu umieszczonego w szerszej ściance falowodu

czenia będą więc tym bardziej dokładne, im mniejszych otworów będą dotyczyły. Na rys. 3 naniesiono zależność współczynnika sprzężenia C od średnicy otworu umieszczonego w węższej ściance falowodu prostokątnego, zaś na rys. 4 taki sam wykres dla otworu kołowego umieszczonego w szerszej ściance.

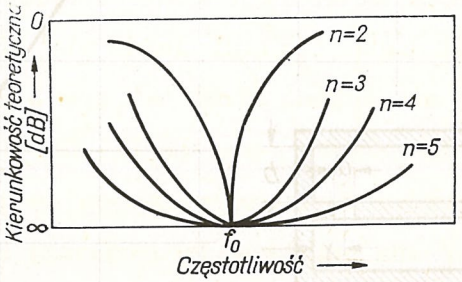
3. SZEREGI SPRZĘGAJĄCE

W celu uzyskania właściwości kierunkowych w szerokim pasmie częstotliwości oraz w celu osiągnięcia silniejszego sprzężenia dwu falowodów, stosuje się nie dwa, a więcej otworów sprzęgających. Kilka otworów umieszczonych we wspólnej ściance dwu falowodów nazywać będziemy szeregiem sprzęgającym.

Dobór współczynników sprzężenia poszczególnych otworów w ramach szeregu sprzęgającego decyduje o kształcie wypadkowej charakterystyki kierunkowości w funkcji częstotliwości. Zależnie od sposobu dobierania średnic otworów sprzęgających, przy stałej ilości otworów, uzyskać można różne przebiegi zmian kierunkowości przy zmianie częstotliwości. Spośród wielu możliwych sposobów zostaną omówione dwa najczęściej stosowane, a mianowicie tworzenie szeregów dwumianowych oraz szeregów Czebyszewa.

3.1. Szeregi dwumianowe

Pojęcie szeregu dwumianowego pochodzi stąd, że względne napięcia sprzęgane przez poszczególne otwory szeregu są proporcjonalne do licz-



Rys. 5. Kierunkowość w funkcji częstotliwości dla różnych szeregów dwumianowych

bowych współczynników rozwinięcia dwumianu $(a + b)^{n-1}$, gdzie n jest ilością otworów sprzęgających. Sprzęgacz dwuotworowy jest więc najniższą formą sprzęgacza dwumianowego, gdyż dla $n = 2$ współczynniki dwumianu są równe jedności. Na rys. 5 przedstawiono względne różnice w szerokopasmowości między sprzęgaczem dwuotworowym a sprzęgaczem zawierającym dwumianowy szereg sprzęgający dla $n = 3, 4, 5$. Wi-

dać, że ze wzrostem ilości otworów zmiany kierunkowości w funkcji częstotliwości stają się coraz to mniejsze.

Tablica 1

Dwumianowe współczynniki sprzężenia

2 otwory	1 (6,00)				1 (6,00)								
3 otwory		1 (12,0)		2 (6,0)		1 (12,0)							
4 otwory			1 (18,1)		3 (8,5)		3 (8,5)		1 (18,1)				
5 otworów		1 (24,1)		4 (12,0)		6 (8,5)		4 (12,0)		1 (24,1)			
6 otworów	1 (30,1)		5 (16,1)		10 (10,1)		10 (10,1)		5 (16,1)		1 (30,1)		
7 otworów	1 (36,1)	6 (20,5)		15 (12,6)		20 (10,1)		15 (12,6)		6 (20,5)		1 (36,1)	

Wybór ilości otworów n podyktowany jest zwykle wymaganą minimalną kierunkowością oraz wielkością sprzężenia. W tablicy 1 zestawiono współczynniki dwumianu oraz względne wartości sprzężenia dla każdego otworu wyrażone w decybelach, dla szeregów składających się z dwu do siedmiu otworów sprzęgających. Współczynniki dwumianu określają względne amplitudy napięć sprzęganych przez każdy z elementów, zaś podane poniżej nich wartości sprzężenia są wartościami, na podstawie których określa się średnicę otworów z równań (4)—(6).

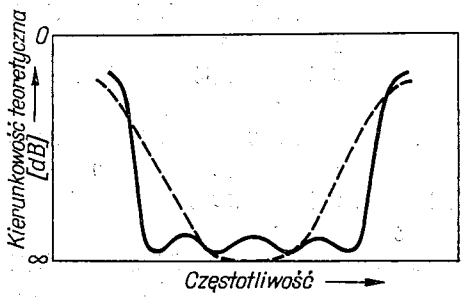
Względne sprzężenie każdego z otworów wyrażone w decybelach określa się biorąc sumę wszystkich współczynników napięciowych w szeregu i biorąc 20 logarytmów ze stosunku sumy do współczynnika poszczególnego otworu. Na przykład dla szeregu $n = 4$, suma wynosi $1 + 3 + 3 + 1 = 8$, a sprzężenie otworów znajduje się biorąc $20 \log \frac{8}{3} = 8{,}5$ dB oraz $20 \log \frac{8}{1} = 18{,}1$ dB. Jeżeli więc sprzężenie sprzęgacza w węższej ściance ma wynosić 30 dB, to z rys. 3 znajduje się średnice otworów dla 38,5 dB $- 0{,}280 \frac{d}{a}$, oraz dla 48,1 dB $- 0{,}207 \frac{d}{a}$.

Odległości pomiędzy otworami w szeregu dwumianowym powinny wynosić zawsze $\frac{\lambda_g}{4}$ dla projektowanej środkowej częstotliwości pasma. W związku z tym mogą wystąpić poważne trudności przy realizowaniu silnego sprzężenia, szczególnie gdy sprzęgacz jest projektowany dla górnego krańca częstotliwości.

3.2. Szeregi Czebyszewa

Nazwa szeregu wywodzi się stąd, że względne amplitudy napięć sprzęganych przez każdy z otworów są proporcjonalne do współczynników

wielomianu Czebyszewa. Zaletą tego szeregu w porównaniu z poprzednim jest możliwość uzyskania szerszego pasma częstotliwości, gdyż kierunkowość oscyluje z równymi garbami wokół danej wartości w maksymalnie szerokim pasmie częstotliwości. Na rys. 6 przedstawiono porównanie kierunkowości szeregu Czebyszewa oraz szeregu dwumianowego o jednakowej ilości elementów sprzęgających.



Rys. 6. Porównanie kierunkowości szeregów Czebyszewa i dwumianowego o równej ilości elementów sprzęgających

Sposób projektowania szeregów Czebyszewa najlepiej przedstawić na przykładzie liczbowym. Przykład opracowano dla pasma 8,4—11,5 GHz w falowodzie FP 22 × 10 dla szeregów zawierających od dwu do sześciu elementów sprzęgających.

Współczynnik pokrycia zakresu

$$p = \frac{\lambda_{g\,max}}{\lambda_{g\,min}} = \frac{57{,}10}{31{,}73} = 1{,}80. \tag{7}$$

Odległości pomiędzy elementami na dolnej i górnej częstotliwości pasma, Φ_1 i Φ_2 określa się jako:

$$\Phi_1 = \frac{180°}{1+p} = \frac{180°}{2{,}80} = 64{,}28°,$$
$$\Phi_2 = 180° - \Phi_1 = 115{,}72°. \tag{8}$$

Współczynnik skali częstotliwości

$$t = \frac{1}{\cos \Phi_1} = \frac{1}{0{,}4336} = 2{,}306. \tag{9}$$

Amplitudy napięć sprzęganych przez poszczególne otwory muszą być proporcjonalne do współczynników wielomianu Czebyszewa pierwszego rodzaju. Wartości tych współczynników dla omawianego przypadku zebrano w tablicy 2. Wartości występujące w rzędach poziomych są proporcjonalne do pożądanych napięć. Wartości w każdym rzędzie dzieli się przez najmniejszą liczbę w rzędzie i w ten sposób uzyskuje się na-

Tablica 2

Wartość współczynników wielomianu Czebyszewa

n \ k	0	1	2	3	4	5
1	2					
2		2,306				
3	8,635		5,317			
4		29,867		12,261		
5	129,111		91,830		28,274	
6		479,623		264,700		65,200

pięcia względne odniesione do jedności. Aby uzyskać wartość sprzężenia wyrażoną w decybelach, należy zsumować współczynniki i sumę podzielić przez poszczególne wartości, tak jak to pokazano w rozdz. 3.1. Wyniki tych działań zebrano w tablicy 3.

Tablica 3

Czebyszewowskie współczynniki sprzężenia

2 otwory				1,000 (6,00)		1,000 (6,00)					
3 otwory			1,000 (11,18)		1,624 (6,97)		1,000 (11,18)				
4 otwory		1,000 (16,74)		2,436 (9,00)		2,436 (9,00)		1,000 (16,74)			
5 otworów	1,000 (22,32)		3,248 (12,08)		4,566 (9,13)		3,248 (12,08)		1,000 (22,32)		
6 otworów	1,000 (27,90)	4,059 (15,87)		7,356 (10,56)		7,356 (10,56)		4,050 (15,87)		1,000 (27,90)	

Odległości pomiędzy osiami otworów w szeregu Czebyszewa powinny wynosić ćwierć długości fali w falowodzie, określonej jako średnica zmian prędkości fazowej

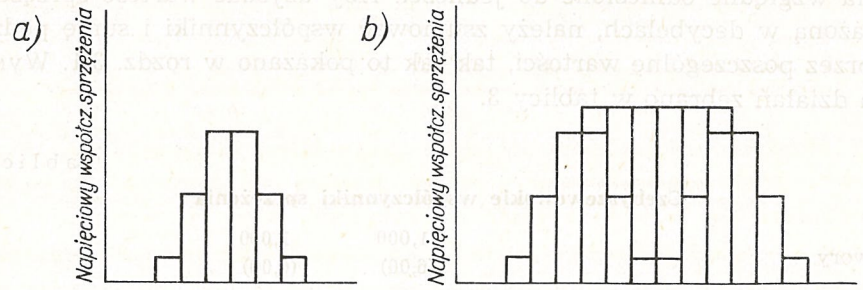
$$\frac{\lambda_{g\acute{s}r}}{4} = \frac{\lambda_{g\,max} \cdot \lambda_{g\,min}}{2\,(\lambda_{g\,max} + \lambda_{g\,min})}. \tag{10}$$

W omawianym przypadku odległość pomiędzy osiami otworów sprzęgających powinna więc wynosić 10,52 mm. Częstotliwość, dla której określa się średnicę otworów, należy wybierać tak, aby wartość sprzężenia, obliczona na brzegach pasma i w kilku jego punktach, przybierała na tej częstotliwości wartość średnią. W przypadku sprzęgaczy z otworami w szerszej ściance selektywność otworu zmienia się zależnie od umieszczenia otworu wzdłuż boku a. Najmniejszą selektywność uzyskuje się dla $x = 0{,}203\,a$, tak że wartość sprzężenia zmienia się o kilka dziesiątych decybela w porównaniu z wartością obliczoną z równania (6).

3.3. Nakładanie szeregów

W wielu przypadkach przy projektowaniu sprzęgaczy kierunkowych z omówionymi powyżej szeregami dwumianowymi czy Czebyszewowskimi, nie można spełnić wszystkich wymagań elektrycznych i mechanicznych. Pojedyncze szeregi sprzęgające nie pozwalają na zaprojektowanie sprzęgaczy o silnym sprzężeniu (C równe kilku dB), gdyż średnice otworów byłyby zbyt duże lub długość szeregu stałaby się niedopuszczalna.

Sprzężenie szeregu jest proporcjonalne do powierzchni pod krzywą napięciowych współczynników sprzężenia poszczególnych otworów pokazaną na rys. 7a. Nakładanie szeregów polega na dodawaniu współczyn-



Rys. 7. Względne wielkości współczynników sprzężenia
a) dla szeregu pojedynczego, b) dla szeregu utworzonego przez nałożenie trzech szeregów pojedynczych

ników sprzężenia, tak aby uzyskać silniejsze sprzężenie przy możliwie najmniejszej ilości otworów. Wynik nałożenia trzech szeregów pokazanych na rys. 7a przedstawiono na rys. 7b.

Krzywa dla szeregu nałożonego obejmuje powierzchnię trzykrotnie większą niż w przypadku szeregu pojedynczego. Napięcie sprzężone przez tak utworzony szereg jest przeto trzykrotnie większe niż napięcie dla szeregu pojedynczego, co odpowiada wzrostowi sprzężenia o 9,5 dB. Najważniejszą zaletą szeregu nałożonego jest to, że sprzężenie zwiększono trzykrotnie przy dwukrotnym powiększeniu ilości otworów. Poza tym połowa otworów ma jednakową średnicę, co również nie jest bez znaczenia w praktyce. Należy podkreślić, że funkcja kierunkowości szeregu nałożonego będzie miała taki sam charakter, jak kierunkowość szeregów składanych. Oczywiście kierunkowość wzrasta ze wzrostem ilości elementów sprzęgających, tak więc kierunkowość szeregu nałożonego będzie większa niż pojedynczego.

Przykład przedstawiony na rys. 7 ilustruje zasadę nakładania szeregów, a w podobny sposób można postępować z dowolną ilością szeregów pojedynczych o dowolnej ilości elementów. Szeregi należy nakładać w taki sposób, aby szereg nałożony był symetryczny względem jak najwięk-

szej ilości otworów o podobnej średnicy. Wtedy otrzymuje się najkrótsze szeregi, gdyż zbytnie ich wydłużanie podważa celowość stosowania szeregów nałożonych.

4. WIELOOTWOROWY SPRZĘGACZ KIERUNKOWY

W celu pokazania praktycznego zastosowania opisanej metody projektowania sprzęgaczy o dużej kierunkowości przytoczono poniżej przebieg projektowania sprzęgacza o sprzężeniu 10 dB i kierunkowości większej od 40 dB w pasmie 8,4—11,5 GHz, wykonanego z prostokątnego falowodu o wymiarach wewnętrznych $a = 22{,}86$ mm i $b = 10{,}16$ mm. Wymagana jest stałość sprzężenia w pasmie oraz powtarzalność egzemplarzy.

4.1. Określenie szeregu sprzęgającego

Ze względu na konieczność zapewnienia dużej kierunkowości w szerokim pasmie częstotliwości wybrano szeregi Czebyszewa. Z szeregu 5-elementowego ($n = 5$) z tabl. 3 utworzono przez nakładanie sześciu szeregów według podanego poniżej schematu następujący szereg sprzęgający:

1,000; 3,248; 4,566; 3,248; 1,000
1,000, 3,248; 4,566; 3,248; 1,000
1,000; 3,248, 4,566; 3,248; 1,000
1,000; 3,248; . . . .

i tak dalej, biorąc sześć szeregów pięcioelementowych. Po zsumowaniu uzyskuje się otwory o następujących zredukowanych wartościach współczynnika sprzężenia:

1,000; 3,248; 4,566, 4,248, 4,248; 4,566; 4,248; 4,248; 4,566; 4,248;
4,248; 4,566; 4,248; 4,248; 4,566; 4,248; 4,248; 4,566; 3;248; 1,000;

Szereg zawiera razem 20 otworów, a suma współczynników sprzężenia wynosi 78,372.

4.2. Określenie otworów sprzęgających

Aby uzyskać współczynniki sprzężenia wyrażone w decybelach dla każdego elementu sprzęgającego, należy sumę współczynników sprzężenia wszystkich elementów szeregu podzielić przez poszczególne współczynniki i wyznaczyć 20 logarytmów z otrzymanych ilozarów.
Wartości sprzężenia poszczególnych otworów są następujące:

$$20 \log \frac{78{,}372}{4{,}566} = 20 \log 17{,}16 = 24{,}69 \text{ dB},$$

$$20 \log \frac{78{,}372}{4{,}248} = 20 \log 18{,}47 = 25{,}23 \text{ dB},$$

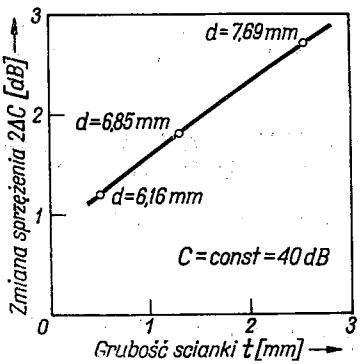
$$20 \log \frac{78{,}372}{3{,}248} = 20 \log 24{,}13 = 27{,}68 \text{ dB},$$

$$20 \log \frac{78{,}372}{1{,}000} = 20 \log 78{,}37 = 37{,}88 \text{ dB}.$$

Ze względu na mniejszą selektywność otworu kołowego umieszczonego w szerszej ściance falowodu wybrano do realizacji ten sposób sprzężenia. Umieszczenie szeregu sprzęgającego w szerszej ściance falowodu ma dodatkową zaletę, że po obu stronach osi falowodu można umieścić dwa równoległe szeregi otworów. Ponieważ podwojona ilość elementów sprzęgających zawsze powiększa całkowite sprzężenie o 6 dB, do uzyskanych powyżej wartości należy dodać po 16 dB i wtedy uzyska się pożądaną wartość sprzężenia dla każdego otworu. Korzystając z równania (6) lub z rys. 4 znajduje się następnie średnice poszczególnych otworów.

4.3. Określenie średnic otworów

Określenie częstotliwości w pasmie pracy, dla której określa się średnice otworów, wymaga szeregu obliczeń, gdyż sprzężenie nie jest liniową funkcją częstotliwości. Na podstawie obliczeń zmian sprzężenia w funkcji częstotliwości otwory określano dla częstotliwości 10,0 GHz, mimo, że środek pasma wypada na częstotliwości 9,68 GHz.



Rys. 8. Zmiana sprzężenia w pasmie 8,4—11,5 GHz w funkcji grubości ścianki

Na wartość sprzężenia wpływ ma również grubość ścianki t pomiędzy dwoma sprzęganymi falowodami. Okazuje się, że selektywność większych otworów jest większa oraz że zwiększenie grubości ścianki powoduje wzrost selektywności. Na rys. 8 przedstawiono zmianę sprzężenia w pasmie pracy w funkcji grubości ścianki dla stałej wartości sprzężenia $C_0 = 40$ dB. Ponieważ należy dążyć do stosowania otworów o możliwie

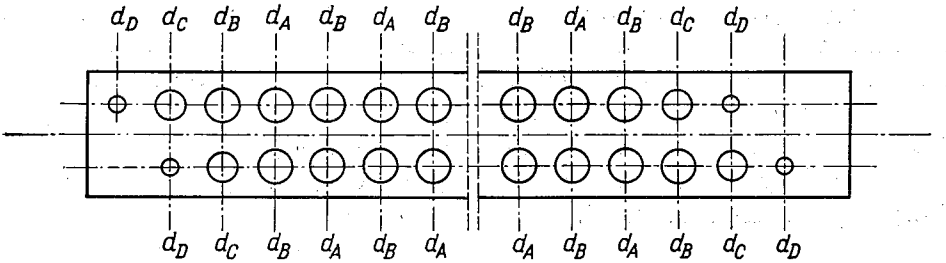
małych średnicach, dla których słuszna jest teoria Bethego, wybrano ściankę o grubości $t = 0{,}50$ mm. W ten sposób sprowadza się selektywność sprzęgacza do minimum.

Korzystając z krzywej podobnej do pokazanej na rys. 4, przeliczonej dla $t = 0{,}50$ mm oraz $f = 10{,}0$ GHz, określono średnice otworów sprzęgających:

A	$C = 40{,}69$ dB	$d_A = 6{,}02$ mm,
B	$C = 41{,}23$ dB	$d_B = 5{,}91$ mm,
C	$C = 43{,}65$ dB	$d_C = 5{,}48$ mm,
D	$C = 53{,}88$ dB	$d_D = 3{,}94$ mm.

Odległości pomiędzy otworami są jednakowe i zostały określone z równania (10) jako $l = 10{,}54$ mm.

Kierunkowość sprzęgacza zawierającego dwa równoległe rzędy otworów sprzęgających umieszczone w szerszej ściance falowodu można dodatkowo poprawić przez przesunięcie obu rzędów o wielkość równą odległości między otworami. Postępowanie takie zapewnia dodatkowy efekt



Rys. 9. Rozmieszczenie otworów w odcinku sprzęgającym sprzęgacza 10 dB

dwumianowy, ponieważ można przyjąć, że każdy szereg ma podobne charakterystyki promieniowania w niepożądanym kierunku, a szczątkowe wartości energii będą się znosić w podobny sposób jak w prostym sprzęgaczu dwuotworowym. Szkic rozmieszczenia otworów wzdłuż ścianki falowodu pokazano na rys. 9.

4.4. Określenie tolerancji wykonania mechanicznego

Przy dopuszczalnym rozrzucie sprzężenia tolerancje wykonania średnic otworów sprzęgających można określić z nachylenia krzywej $C = f(d)$. Ze względu na nieliniowość tej zależności różnym średnicom odpowiadają inne zmiany sprzężenia. Zmiany sprzężenia są tym mniejsze, im większa jest średnica otworu. W zakresie sprzężenia 40,80—43,50 dB średnie nachylenie krzywej wynosi $0{,}054 \frac{\text{dB}}{0{,}01\text{mm}}$, zaś w zakresie

49,50—53,40 dB już 0,073 $\frac{\text{dB}}{0{,}01\ \text{mm}}$. Wykonywanie średnic otworów w sprzęgaczu o sprzężeniu 10 dB z dokładnością ±0,05 mm pozwoli uzyskać nominalną wartość sprzężenia z dokładnością ±0,4 dB.

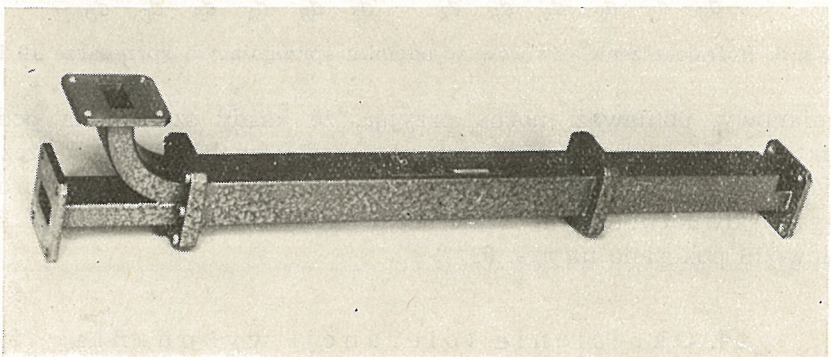
Podobną tolerancję ±0,05 mm nałożono również na odległości pomiędzy otworami oraz na rozstawienie otworów względem osi falowodu.

4.5. Konstrukcja sprzęgacza

Ze względu na wymaganą dużą kierunkowość sprzęgacza, WFS wprowadzany przez obciążenie umieszczone w gałęzi pomocniczej powinien być mniejszy od 1,02. W tym celu zastosowano obciążenie wykonane z płytki mikowej napylonej chromonikieliną [11]. Obciążenie to ma długość 120 mm, co powoduje wydłużenie całej konstrukcji.

Aby umożliwić wygodne wyjście gałęzi pomocniczej, zakończono ją zagięciem 90° wykonanym w płaszczyźnie E promieniem $R = 40$ mm. WFS zagięcia w pasmie pracy nie przekracza wartości 1,08.

Sprzęgacz składa się więc z trzech części skręcanych poprzez specjalne pokrywy. Jedna część zawiera zagięcie falowodowe, część środkowa zawiera elektroformowany odcinek sprzęgający z dwoma szeregami otworów umieszczonymi w płytce o grubości 0,50 mm (zob. rys. 9), zaś trzecia część sprzęgacza zawiera dopasowane obciążenie. Całkowita długość sprzęgacza wynosi 420 mm. Zdjęcie jednego z egzemplarzy modelowych pokazano na rys. 10.

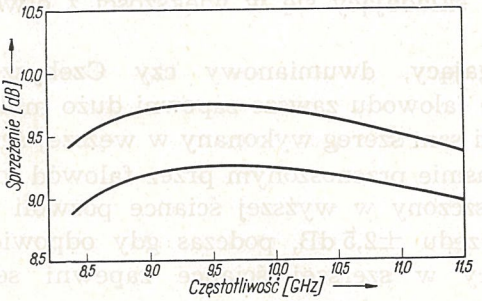


Rys. 10. Widok zewnętrzny sprzęgacza kierunkowego o sprzężeniu 10 dB i kierunkowości ponad 40 dB

4.6. Charakterystyki sprzęgacza

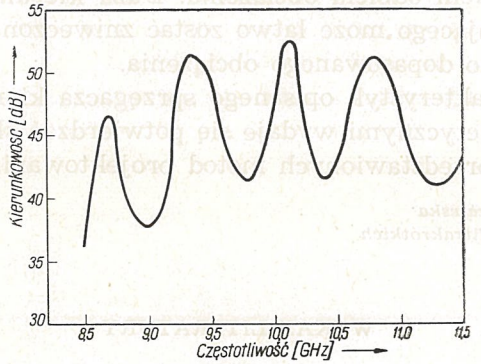
Na rys. 11 przedstawiono przebieg sprzężenia w funkcji częstotliwości dla siedmiu egzemplarzy sprzęgacza. Rozrzut wartości sprzężenia nie przekracza w żadnym przypadku wartości ±0,3 dB. Średnie sprzężenie jest niższe od projektowanego o ok. 0,5 dB, co może wynikać z przyjęcia

zbyt łagodnych tolerancji wykonania mechanicznego, jak również z faktu, że nie spełnione jest podstawowe założenie teorii Bethego, a mianowicie $\frac{d}{\lambda}$ nie jest dużo mniejsze od jedności. Zmiany sprzężenia w funkcji częstotliwości są jednak zadowalające, gdyż selektywność sprzęgacza w pasmie 32% wynosi $\pm 0{,}3$ dB.



Rys. 11. Przebieg sprzężenia w funkcji częstotliwości — dwie skrajne charakterystyki

Na rys. 12 przedstawiono zależność kierunkowości od częstotliwości. Poza nieznacznym zakresem dla niższych częstotliwości, gdzie WFS obciążenia nieznacznie przekracza wartość 1,02, kierunkowość sprzęgacza jest większa od 40 dB. WFS w gałęzi pierwotnej nie przekracza 1,05, zaś WFS w gałęzi pomocniczej jest mniejszy od 1,12, gdy pozostałe końcówki zakończone są dopasowanymi obciążeniami.



Rys. 12. Przebieg kierunkowości w funkcji częstotliwości — średnia siedmiu pomiarów

5. ZAKOŃCZENIE

Opisana metoda nakładania szeregów sprzęgających znajduje zastosowanie do apertur sprzęgających dowolnego kształtu. Zaleta metody polega na zmniejszeniu ilości otworów koniecznych do uzyskania odpowiednio silnego sprzężenia oraz na znacznym uproszczeniu obliczeń. Gdyby

bowiem w opisanym przykładzie nie stosować nałożonego szeregu sprzęgającego, należałoby zastosować 6 szeregów 5-otworowych, czyli razem 30 otworów lub obliczyć szereg zawierający kilkanaście otworów. W długich szeregach pojedynczych stosunek średnicy otworu środkowego do średnicy otworu skrajnego jest bardzo duży, szczególnie duży, gdy ma się do czynienia z szeregiem dwumianowym. Nakładanie szeregów pozwala uzyskać szereg składający się w większości z otworów o podobnych średnicach.

Szereg sprzęgający, dwumianowy czy Czebyszewa, umieszczony w szerszej ściance falowodu zawsze zapewni dużo mniejszą selektywność sprzężenia niż taki sam szereg wykonany w węższej ściance. Jeżeli sprzęgacz pracuje w pasmie przenoszonym przez falowód (ok. 42%), to szereg sprzęgający umieszczony w wyższej ściance pozwoli uzyskać selektywność sprzężenia rzędu $\pm 2{,}5$ dB, podczas gdy odpowiednio umieszczony szereg sprzęgający w szerszej ściance zapewni selektywność rzędu $\pm 0{,}5$ dB.

Zaleca się stosowanie małych otworów sprzęgających, gdyż mimo wydłużenia się szeregu uzyskuje się większą zgodność doświadczeń z obliczeniami teoretycznymi, gdyż wyrażenia na sprzężenie wyprowadzono na podstawie rozważania małych otworów. Poza tym w przypadku sprzężenia w szerszej ściance z równania wynika, że selektywność małych otworów jest mniejsza od selektywności dużych otworów.

Należy również podkreślić konieczność stosowania dobrych obciążeń w falowodzie pomocniczym, gdyż kierunkowość sprzęgacza określona jest współczynnikiem odbicia obciążenia. Duża kierunkowość wewnętrzna układu sprzęgającego może łatwo zostać zniweczona przez zastosowanie nieodpowiednio dopasowanego obciążenia.

Zgodność charakterystyk opisanego sprzęgacza kierunkowego z przewidywaniami teoretycznymi wydaje się potwierdzać słuszność i przydatność praktyczną przedstawionych metod projektowania.

Politechnika Warszawska
Katedra Techniki Fal Ultrakrótkich

WYKAZ LITERATURY

1. Bethe H. A.: Theory of Diffraction by Small Holes.-Phys. Rev., t. 66, Nr 7, str. 163—182, październik 1944.
2. Montgomery C. G.: Technique of Microwave Measurements, rozdz. 14. Mc Graw Hill Book Co., New York 1947.
3. Stevenson A. E.: Theory of Slots in Rectangular Wave — Guides.-Jour. Appl. Phys., t. 19, Nr 1, str. 24—38, styczeń 1948.
4. Surdin M.: Directive Couplers in Waveguides.-Jour. IEE, t. 93, cz. IIIA, Nr 4, str. 725—736, 1946.
5. Mumford W. W.: Directional Couplers.-Proc. IRE, t. 35, Nr 2, str. 160—165, luty 1947.

6. Miller S. E., Mumford W. W.: Multi-Element Directional Couplers.-Proc. IRE, t. 40, Nr 9, str. 1071—1078, wrzesień 1952.
7. Schwartz R. F.: Bibliography on Directional Couplers. — IRE Trans., t. MTT-2, str. 58—63, lipiec 1954.
8. Hensperger E. S.: The Design of Multi-Hole Coupling Arrays. — Microwave Jour., t. 2, Nr 8, str. 38—42, sierpień 1959.
9. Levy R.: A Guide to the Practical Application of Chebyshev Functions.-Proc. IEE, t. 106, cz. C, Nr 10, str. 193—199, wrzesień 1959.
10. Sadd T. S.: Microwave Engineer's Handbook. Horizon House-Microwave Inc., Boston, Mass. 1961.
11. Kraszewski A.: Dopasowane obciążenia falowodowe. — Przegląd Telekomunik., t. 35, nr 3, str. 85—88, marzec 1963.

A. KRASZEWSKI

HIGH DIRECTIVITY WAVEGUIDE DIRECTIONAL COUPLER

Summary

The method of designing broad-band, high directivity directional couplers is described. Fundamental formulas and curves, required to design an arbitrary directional coupler, formed by two coupled rectangular waveguides, are presented. Special attention is given to the multihole couplers, which permits to obtain high directivity in broad frequency band. Comparision is carried out between properties of the coupling arrays with coupling coefficients defined by the coefficients of binomial expanssion and properties of the arrays with coefficients given by Tchebyshev polinomials. Superiority of the latter is shown.

As the single coupling arrays do not allow for the design of couplers with a strong coupling, since the hole diameter would be too large or lenght of the array would be intolerably long, a simple method of overlaying the arrays is presented. In order to give a practical application of this method, the design of a directional coupler with 10 db coupling and directivity greater than 40 db is shown. The coupler was made from standard rectangular waveguide IEC R-100 (WG16 or WR-90), and was designed for operation from 8.4 to 11.5 Gc/s. Validity of the presented designing method is fully confirmed by the characteristics of this coupler.

Several practical notes have been included on the directional coupler construction, admissible mechanical tolerances, mechanical production and so on. The described method of overlaying the coupling arrays may be applied to coupling apertures of arbitrary shape, and similarly as the remaining notes can be referred to directional couplers operating in arbitrary frequency band.

A. KRASZEWSKI

COUPLEUR DIRECTIF À GRANDE DIRECTIVITÉ

Résumé

Après l'introduction donnant les expressions principales et les courbes permettant construire des coupleurs directifs, l'auteur presente l'analyse des coupleurs à grande directivité, possendant une large bande de fréquences.

Le couplage est formé par une série d'ouvertures, qui doit étre doublée, quand on désire un couplage plus fort. On indique la superiorité de la série qui est accordée avec les coefficients de Tschebysheff à celle accordée avec les coefficients de binôme de Newton.

La méthode de calcul du coupleur directif destiné à la bande X (8.4—11.5 Mc, couplage 10 db, directivité plus que 40 db, qui doit étre construit utilisant le guide d'ondes rectangulaire dont les dimensions interieures sont 22.86×10.16 mm) est donnée pour illustrer la théorie. Les résultats des mesures confirment la méthode acceptée par l'auteur.

On donne aussi quelques remarques pratiques concernant la construction mécanique et les erreurs adminibles.

La méthode peut étre utilisée à toutes les bandes de fréquences et à toutes ouvertures.

A. KRASZEWSKI

HOHLLEITER-RICHTKOPPLER VON GROSSER RICHTWIRKUNG

Zusammenfassung

Hier wurde eine Projektierungsmethode für Breitbandfrequenzbereich-Richtkoppler mit grosser Richtwirkung beschrieben. Zu diesem Zweck wurden Grundformeln und Kennlinien dargestellt, welche für die Projektierung beliebiger Richtkoppler von zwei zusammengelegter miteinander gekoppelten Rechteckhohlleiterabschnitten, nötig sind.

Eine besondere Beachtung wurde den Vielöffnungskopplern gewidmet, welche eine grosse Richtwirkung im Breitband-Frequenzbereich zu erreichen erlauben. Es wurde ein Vergleich der Eigenschaften bei Kopplerreihen durchgeführt, deren Kopplungsfaktoren aus den Faktoren des Binomens und den Faktoren des Multinomens von Czebyszew entwickelt wurden. Hierbei wurde auf die entschiedene Überlegenheit der letzteren hingewiesen.

Da die einzelnen Kopplerreihen eine Projektierung von Richtkopplern mit starker Kopplung nicht gestatten, weil ihre Öffnungsdurchmesser zu gross oder die Länge der Reihe unzulässig wäre, wurde eine einfachere Methode für die Auflegung der zu koppelnden Reihen dargestellt.

Zwecks praktischer Veranschaulichung bezüglich der Ausnutzung der beschriebenen Methode für die Projektierung von Kopplern mit grosser Richtwirkung, wurde der Projektierungsverlauf eines Richtkopplers angeführt, dessen Kopplung 10 dB und seine Richtwirkung grösser als 40 dB im Frequenzband von 8,4 bis 11,5 GHz liegt, und der aus standartisiertem Rechteckhohlleiter $22{,}8 \times 10{,}2$ mm (IEC R-100) ausgeführt wurde.

Die dargestellten Kennlinien des Richtkopplers bestätigen vollends die Richtigkeit der dargestellten Projektierungsmethode.

Es wurde ferner eine ganze Reihe praktischer Bemerkungen bezüglich der mechanischen Ausführung, der zulässigen Toleranzen und der konstruktion Lösung dargestellt. Die beschriebene Methode der Auflegung der zu koppelnden Reihen kann für Kopplungsfenster von beliebiger Form Anwendung finden, und ähnlich wie die übrigen Bemerkungen, sich auch auf die Richtkoppler, welche im beliebigen Frequenzband arbeiten, beziehen können.

А. КРАШЕВСКИ

ВОЛНОВОДНЫЙ НАПРАВЛЕННЫЙ ОТВЕТВИТЕЛЬ С БОЛЬШОЙ НАПРАВЛЕННОСТЬЮ

Резюме

В статье приводится метод проектирования широкополосных направленных ответвителей с большой направленностью. Представлены основное формулы и диаграммы, требуемые для проектирования произвольных направленных ответвителей, созданных из двух взаимно связанных отрезков прямоугольного волновода. Особое внимание обращено на многощелевые ответвители, которые позволяют получить большую направленность в широком диапазоне частот. Сравниваются свойства направленных систем, коэффициенты связи которых определены при помощи коэффициентов разложения бинома Ньютона, со свойствами направленных систем, коэффициенты связи которых определены при помощи коэффициентов полиномов Чебышева. Доказано, что эти последние системы отличаются решительным превозходством по сравнению с первыми.

Так как одиночная направленная система не позволяет изготовить ответвитель с сильной связью (ибо диаметры щелей получились бы слишком значительными, или же длина системы получилась бы слишком большой), представлен простой метод наложения направленных систем. Для доказательства практического преимущества описанного метода проектирования направленных ответвителей с большой направленностью, приводится пример проектирования направленного ответвителя со связью 10 дб и направленностью больше 40 дб в диапазоне частот 8,400 — 11,500 мгц, изготовленного из стандартного прямоугольного волновода 22,8 × 10,2 мм (IEC R-100). Представленные характеристики ответвителя полностью подтверждают правильность приведенного метода.

Приводится ряд практических замечаний относительно конструкции ответвителя, его механического изготовления, допускаемых погрешностей и т.д. Изложенный метод наложения направленных систем может быть применен к произвольной формы отверстиям связи, и, так же как и другие замечания, может касаться направленных ответвителей, работающих в произвольном диапазоне частот.

621.396.67

ANDRZEJ KRASZEWSKI

Zastosowanie falowodowych układów hybrydowych w technice mikrofalowej

Rękopis dostarczono 3.12.1962

Omówiono zasadę działania kilku przyrządów mikrofalowych wykorzystujących właściwości układów hybrydowych. Strojnik falowodowy, przesuwnik fazy i dzielnik mocy zawierające hybrydę z krótką szczeliną mają szereg zalet w porównaniu z rozwiązaniami klasycznymi, szczególnie przy pracy na dużym poziomie mocy. Podano przykłady konkretnych rozwiązań konstrukcyjnych oraz omówiono inne możliwości wykorzystania układów hybrydowych w technice mikrofalowej.

1. WPROWADZENIE

Opracowanie mikrofalowych układów hybrydowych rozszerzyło znacznie wachlarz możliwości przy konstrukcji nowych przyrządów mikrofalowych. Szereg konwencjonalnych przyrządów, takich jak przesuwniki fazy, strojniki czy dzielniki mocy można obecnie zastąpić przyrządami wykorzystującymi właściwości układów hybrydowych. Przyrządy takie, poza innymi zaletami elektrycznymi i mechanicznymi, mają tę zaletę, że działają poprawnie na dużym poziomie mocy. Poza tym istnieją możliwości opracowania, w oparciu o właściwości układów hybrydowych innych wygodnych przyrządów, takich jak przełączniki, modulatory, układy rozdzielające sygnały o różnych częstotliwościach, mieszacze zrównoważone, tłumiki, zrównoważone przełączniki antenowe, wzmacniacze i wiele innych.

Pewne podstawowe zasady są wspólne dla wszystkich takich układów i zostaną przytoczone w niniejszej pracy. W zasadzie zostanie rozpatrzony przypadek dwu układów hybrydowych połączonych kaskadowo przy założeniu, że hybryda jest dopasowana i kierunkowa. W pewnych rozwiązaniach występuje jedna tylko hybryda, lecz jest ona wtedy wykorzystywana dwukrotnie, raz do podziału przychodzącego sygnału, a następnie do ponownego jego połączenia. Matematycznie problem jest zupełnie podobny jak w przypadku układu zawierającego dwie hybrydy, tak że oba przypadki można rozpatrywać wspólnie.

Celem niniejszej pracy jest przedstawienie ogólnej zasady działania przyrządów wykorzystujących właściwości układów hybrydowych oraz

wykazanie na konkretnych przykładach zalet tych przyrządów oraz dalszych możliwości ich zastosowania. Przytoczone rozważania są wynikiem prac prowadzonych w tym kierunku w Katedrze Techniki Fal Ultrakrótkich Politechniki Warszawskiej.

2. ANALIZA OGÓLNA

W większości przypadków praktycznych macierz rozproszenia dla układu hybrydowego można zapisać w postaci:

$$\mathbf{S}_H = \begin{bmatrix} 0 & 0 & \hat{S}_{13} & \hat{S}_{14} \\ 0 & 0 & \hat{S}_{14} & \hat{S}_{13} \\ \hat{S}_{13} & \hat{S}_{14} & 0 & 0 \\ \hat{S}_{14} & \hat{S}_{13} & 0 & 0 \end{bmatrix} \tag{1}$$

gdyż zwykle układ taki można traktować jako dopasowany, symetryczny i odwracalny. Wartości współczynników macierzy (1) są wielkościami zespolonymi i można je wyrazić jako

$$\hat{S}_{13} = S_{13}\, e^{-j\theta_{13}} \text{ oraz } \hat{S}_{14} = S_{14}\, e^{-j\theta_{14}}. \tag{2}$$

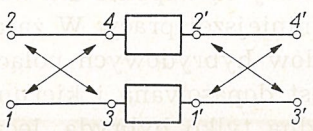
Ponieważ przyjmuje się, że układ jest bezstratny, z warunku zachowania energii wynika, że:

$$S_{14} = \sqrt{1 - S_{13}^2} \text{ oraz } \theta_{14} = \theta_{13} \pm \frac{\pi}{2}. \tag{3}$$

Współczynnik nierównomierności podziału mocy k można wyrazić jako

$$k = \frac{1 - S_{13}^2}{S_{13}^2}. \tag{4}$$

Znak w równaniu (3) zależy od typu użytego układu hybrydowego. W dalszym ciągu analizy znak ten nie posiada zasadniczego znaczenia i korzystając z dowolności stosować będziemy znak plus, który odnosi się do układu hybrydowego z krótką szczeliną [11].



Rys. 1. Ogólny schemat układu

Na rys. 1 przedstawiono ogólny schemat układu zawierającego dwa układy hybrydowe. Hybrydy oraz elementy prowadnicy falowej można scharakteryzować za pomocą napięciowych współczynników rozproszenia zapisanych w postaci $\hat{S}_{ij} = S_{ij}\, e^{-j\theta_{ij}}$.

Można prosto wykazać, że macierz rozproszenia układu pokazanego na rys. 1, zawierającego dwa jednakowe układy hybrydowe, których właściwości określa macierz (1), oraz dwa dowolne czwórniki umieszczone pomiędzy nimi, ma postać:

$$\mathbf{S} = \begin{bmatrix} A & B & C & D \\ B & E & D & F \\ C & D & A & B \\ D & F & B & E \end{bmatrix}, \tag{5}$$

gdzie:

$$A = S_{13}^2 e^{-j2\theta_{13}} S_{33} e^{-j\theta_{33}} + S_{14}^2 e^{-j2\theta_{14}} \cdot S_{44} e^{-j\theta_{44}},$$

$$B = S_{13} S_{14} e^{-j(\theta_{13}+\theta_{14})} [S_{33} e^{-j\theta_{33}} + S_{44} e^{-j\theta_{44}}],$$

$$C = S_{13}^2 S_{31'} e^{-j(2\theta_{13}+\theta_{31'})} + S_{14}^2 S_{42'} e^{-j(2\theta_{14}+\theta_{42'})},$$

$$D = S_{13} \cdot S_{14} e^{-j(\theta_{13}+\theta_{14})} [S_{31'} e^{-j\theta_{31'}} + S_{42'} e^{-j\theta_{42'}}],$$

$$E = S_{13}^2 S_{44} e^{-j(2\theta_{13}+\theta_{44})} + S_{14}^2 S_{33} e^{-j(2\theta_{14}+\theta_{33})},$$

$$F = S_{13}^2 S_{42'} e^{-j(2\theta_{13}+\theta_{42'})} + S_{14}^2 S_{31'} e^{-j(2\theta_{14}+\theta_{31'})}.$$

W ogólnym przypadku, aby znaleźć współczynniki przenoszenia układu, należy rozwiązać równanie macierzowe

$$\mathbf{b} = \mathbf{Sa} \tag{6}$$

gdzie **a** — macierz fal padających, **b** — macierz fal odbitych, zaś **S** — macierz układu określona równ. (5). Ogólne wyrażenia na b_1, b_2, $b_{3'}$ i $b_{4'}$ pozwalają określić przenoszenie energii w układzie z uwzględnieniem niesymetrii czwórników umieszczonych pomiędzy hybrydami oraz przy dowolnym zakończeniu gałęzi wyjściowych ($a_2 = \Gamma_2 b_2$, $a_{3'} = \Gamma_3 b_{3'}$ i $a_{4'} = \Gamma_4 b_{4'}$ [10]. Amplitudę fali padającej z gałęzi 1 (generator) można przyrównać do jedności i tak dobrać płaszczyzny odniesienia, że $a_1 = 1$.

W tym przypadku, gdy we wszystkich gałęziach wyjściowych układu umieszczone są dopasowane obciążenia (przypadek $a_1 = 1$, $a_2 = a_{3'} = a_{4'} = 0$), wyrażenia na współczynniki przenoszenia układu bardzo się upraszczają. Korzystając z równania (4) można napisać:

$$S_{11} e^{-j\theta_{11}} = S_{13}^2 e^{-j2\theta_{13}} S_{33} e^{-j\theta_{33}} - |1 - S_{13}^2| e^{-j2\theta_{13}} S_{44} e^{-j\theta_{44}}, \tag{7}$$

$$S_{12} e^{-j\theta_{12}} = S_{13} \sqrt{1 - S_{13}^2}\, e^{-j\left(2\theta_{13}+\frac{\pi}{2}\right)} [S_{33} e^{-j\theta_{33}} + S_{44} e^{-j\theta_{44}}], \tag{8}$$

$$S_{13'} e^{-j\theta_{13'}} = S_{13}^2 e^{-j2\theta_{13}} S_{31'} e^{-j\theta_{31'}} - |1 - S_{13}^2| e^{-j2\theta_{13}} S_{42'} e^{-j\theta_{42'}}, \tag{9}$$

$$S_{14'} e^{-j\theta_{14'}} = S_{13} \sqrt{1 - S_{13}^2}\, e^{-j\left(2\theta_{13}+\frac{\pi}{2}\right)} [S_{31'} e^{-j\theta_{31'}} + S_{42'} e^{-j\theta_{42'}}]. \tag{10}$$

Równania powyższe można bezpośrednio wykorzystać przy analizowaniu właściwości praktycznych układów zawierających hybrydy, jeżeli słuszne jest założenie uczynione odnośnie dopasowania gałęzi wyjściowych.

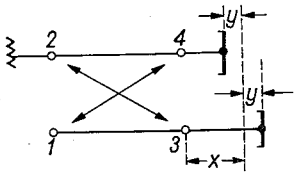
Z równań (8) i (10) wynika, że można rozdzielić wpływ hybryd na współczynniki przenoszenia od wpływu czwórników umieszczonych pomiędzy nimi, gdyż są one od siebie niezależne. Tak więc tłumienie przenoszenia układu będzie sumą tych dwu czynników wyrażonych w decybelach. Dalsze uproszczenia można uzyskać zakładając, że oba czwórniki są identyczne i albo całkowicie bezstratne ($S_{31'} = S_{42'} = 1$), albo odbijają całkowicie energię nań padającą ($S_{33} = S_{44} = 1$). W takich przypadkach można rozdzielić wpływ różnicy przesunięć fazy na współczynniki przenoszenia od wpływu hybryd. Tłumienie wprowadzane przez czwórniki można więc dodawać do tłumienia wynikającego z nierównomiernego podziału mocy w układach hybrydowych oraz tłumienia wywołanego różnicą przesunięć fazy w obu czwórnikach.

Z równań (7) i (9) widać, że czynniki powodujące pogorszenie się izolacji lub wejściowego współczynnika odbicia nie dają się tak wygodnie rozdzielić. Jednakże wprowadzając podobne założenia jak poprzednio można wyznaczyć oddzielnie wpływ układów hybrydowych oraz wpływ różnicy przesunięć fazy na izolację i wejściowy współczynnik odbicia. Gdy przyjmie się, że oba czwórniki są identyczne, jasno wynika fakt, że izolacja gałęzi $3'$ jest powiązana z tłumieniem przenoszenia do gałęzi $4'$, zaś tłumienie przenoszenia do gałęzi 2 jest bezpośrednio związane z odbiciem sygnału w gałęzi 1.

3. ANALIZA SZCZEGÓŁOWA

3.1. Falowodowy strojnik hybrydowy

Jeżeli układ hybrydowy zakończyć dwoma zwieraczami umieszczonymi w sposób pokazany na rys. 2, to przy niezależnym przesuwaniu obu zwieraczy uzyska się ciekawą zależność wejściowego współczynnika od-



Rys. 2. Falowodowy strojnik hybrydowy

bicia (w gałęzi 1). Równanie (6), przy założeniu, że $k = 1$, można dla tego przypadku napisać w postaci

$$S_{11}\, e^{-j\theta_{11}} = \frac{1}{2}\, e^{-j2\theta_{13}} \left[e^{-j2\beta(x+y)} - e^{-j2\beta(x-y)} \right], \tag{11}$$

gdzie $\theta_{33} = \frac{4\pi}{\lambda_g}(x+y)$ oraz $\theta_{44} = \frac{4\pi}{\lambda_g}(x-y)$.

Prosto można znaleźć, że

$$\begin{aligned} S_{11} &= \sin 2\beta y, \\ \theta_{11} &= 2\beta x + 2\theta_{13}. \end{aligned} \tag{12}$$

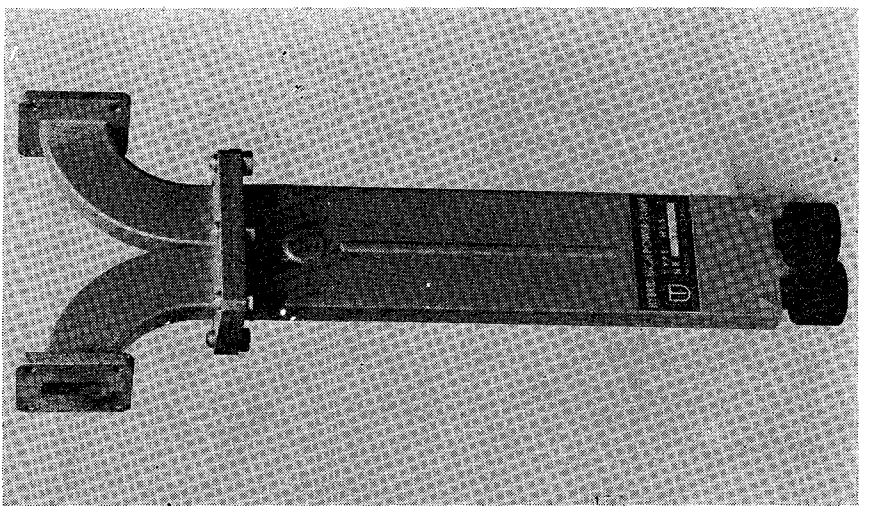
Ponieważ przez odpowiedni dobór płaszczyzny odniesienia można spełnić warunek, że $\Theta_{13} = 0$, z równania (11) widać, że przesuwając zwieracze względem siebie i względem końcówek hybrydy na wejściu można uzyskać dowolną impedancję

$$0 < S_{11} < 1 \qquad \text{dla } 0 \leqslant y \leqslant \frac{\lambda_g}{8}$$

oraz

$$0 < \theta_{11} < 2\pi \qquad \text{dla } 0 \leqslant x \leqslant \frac{\lambda_g}{2}.$$

Układ pokazany na rys. 2 umożliwia zatem dopasowanie praktycznie każdej impedancji do impedancji charakterystycznej prowadnicy falowej.



Rys. 3. Widok strojnika hybrydowego na pasmo 3 cm

Istnieją trzy możliwości dostrajania dowolnej impedancji za pomocą strojnika hybrydowego:

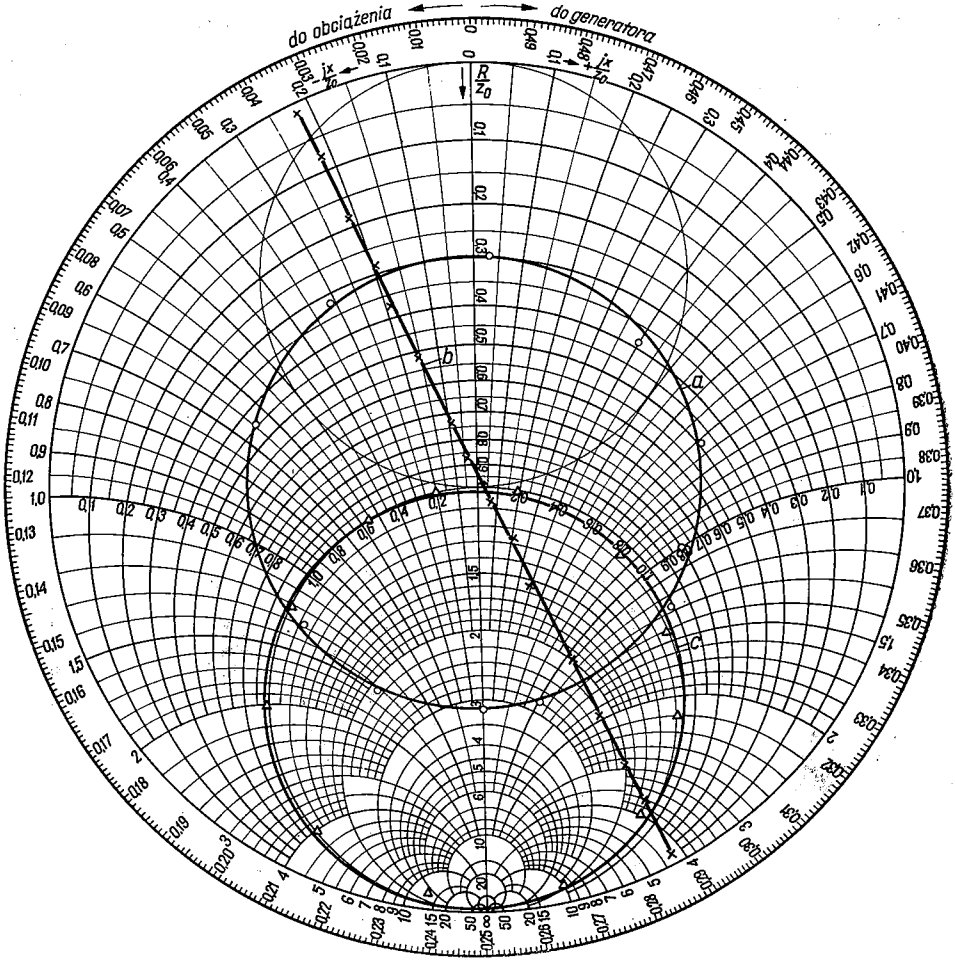
a) y = const, x — zmienne — impedancja zmienia się po kole współosiowym na wykresie Smitha;

b) x = const, y — zmienne — impedancja zmienia się po prostej przechodzącej przez środek wykresu;

c) x — zmienne, y — zmienne — impedancja zmienia się po kole stycznym do środka i brzegu wykresu.

Na rys. 3 przedstawiono zdjęcie strojnika hybrydowego zbudowanego z falowodu prostokątnego FP 22 × 10 i pracującego w paśmie 3 cm, zaś na

rys. 4 pokazano charakterystyki tego układu. Krzywe *a*, *b*, i *c* na rys. 4 odnoszą się do trzech możliwych sposobów dostrajania wymienionych powyżej. Pewne niedoskonałości strojnika wynikają z nieidealnego zachowania się samej hybrydy, która zawsze ma skończoną kierunkowość, oraz zwieraczy, które nigdy nie odbijają całej energii nań padającej. Niecen-



Rys. 4. Możliwe sposoby zmian impedancji wprowadzane przez strojnik pokazany na rys. 3

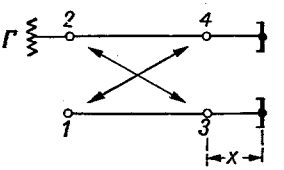
tryczność kół współosiowych wynika z niedopasowania hybrydy, zaś niezamykanie się kół po przesunięciu zwieraczy o $\frac{\lambda_g}{2}$ jest spowodowane stratami energii w odcinku miedzianego falowodu pomiędzy hybrydą a zwieraczem.

Szerokopasmowość strojnika hybrydowego jest określona szerokością pasma przenoszonego przez hybrydę, a więc przekracza 15%, jeżeli zasto-

suje się szerokopasmowe zwieracze dławikowe składające się z dwu lub trzech sekcji. Wytrzymałość na przebicie układu hybrydowego jest tego samego rzędu co wytrzymałość na przebicie samego falowodu, tak więc falowodowy strojnik hybrydowy jest przyrządem mającym szerokie zastosowanie w licznych układach pomiarowych.

3.2. Hybrydowy przesuwnik fazy

Sygnał wchodzący gałęzią *1* do układu pokazanego na rys. 5 dzieli się po połowie pomiędzy gałęzie *3* i *4*. Sygnały wyjściowe w gałęziach *3* i *4* są przesunięte względem siebie o 90°. Rozchodzą się one dalej wzdłuż prowadnicy aż do punktów *5* i *6*, w których umieszczono zwieracz. Oba sygnały odbijają się od płaszczyzny zwarcia i powracają w kierunku hy-



Rys. 5. Hybrydowy przesuwnik fazy

brydy. Jeżeli przesunięcie fazy na drodze *3*—*5* (Θ_{35}) jest równe przesunięciu fazy na drodze *4*—*6* (Θ_{46}), wtedy sygnały przedostające się z gałęzi *3* i *4* do gałęzi *1* będą przesunięte w fazie o 180° i będą się całkowicie znosiły. Natomiast sygnały dochodzące z gałęzi *3* i *4* do *2* będą w fazie i w efekcie cała moc z gałęzi *1* zostanie przeniesiona do gałęzi *2*. Jeżeli w gałęzi *2* umieścić obciążenie wprowadzające współczynnik odbicia Γ, wtedy wyrażenie na wejściowy współczynnik odbicia w gałęzi *1*, przy założeniu, że $k = 1$, przybierze postać:

$$S_{11}\, e^{-j\theta_{11}} = \Gamma e^{-j2\beta x}. \tag{13}$$

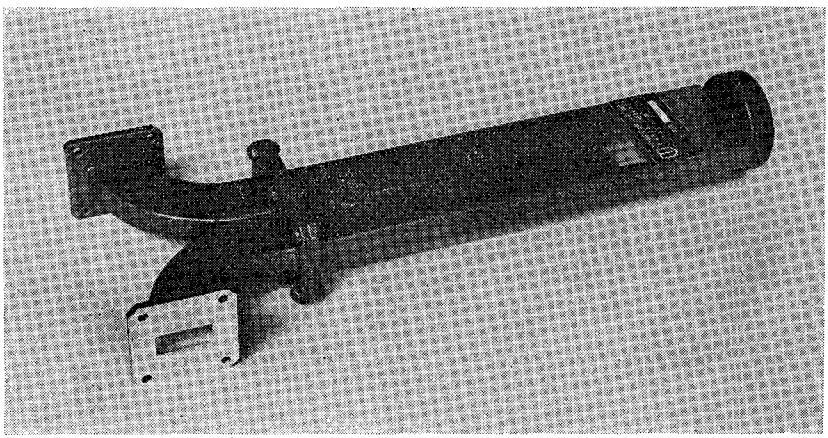
Podobnie jak poprzednio, przez odpowiedni dobór płaszczyzn odniesienia uzyskano $\Theta_{13} = 0$. Widać, że układ spełnia funkcje przesuwnika fazy, w którym przesunięcie fazy $\Delta\Phi$ jest powiązane z przesunięciem Δx zwieracza wyrażeniem

$$\Delta\Phi = \frac{2\pi}{\lambda_g} 2\Delta x. \tag{14}$$

Tak więc przesuwając zwieracz w zakresie $\frac{\lambda_g}{2}$ można uzyskać przesunięcie fazy o 360°.

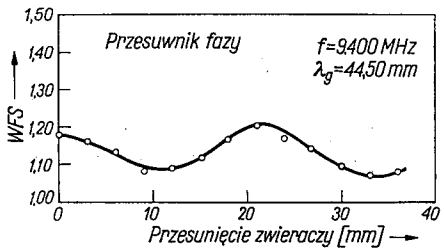
Na rys. 6 przedstawiono zdjęcie hybrydowego przesuwnika fazy zbudowanego z falowodu prostokątnego FP 22 × 10 i pracującego w pasmie 8,4 GHz do 9,8 GHz. Na rys. 7 pokazano przebieg zmian wejściowego

współczynnika fali stojącej w funkcji przesuwania zwieraczy. Mimo że z równania (12) wynika, że układ powinien być dopasowany, w układzie praktycznym dodają się wektorowo odbicia obciążenia, odbicia od hybrydy i odbicia wynikające z niedokładnego podziału mocy. Widać, że okres krzywej pokazanej na rys. 7 wynosi $\frac{\lambda_g}{2}$, co jest zgodne z równaniem (14). Całkowite przesunięcie fazy na najniższej częstotliwości wynosi ponad 400°.



Rys. 6. Widok hybrydowego przesuwnika fazy na pasmo 3 cm

Przesuwnik fazy pokazany na rys. 6 może pracować przy mocy około 200 kW w impulsie. Pod względem wytrzymałości na przebicie przyrząd ten przewyższa znacznie wszystkie rozwiązania klasyczne, w których przesunięcie fazy uzyskuje się dzięki przemieszczeniu płetwy dielektrycznej względem osi falowodu.

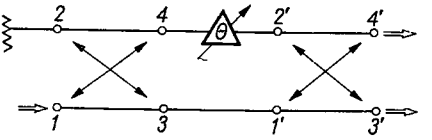


Rys. 7. WFS wejściowy przesuwnika w funkcji położenia zwieraczy

3.3. Hybrydowy dzielnik mocy

Umieszczając dowolny bezstratny przesuwnik fazy pomiędzy dwoma hybrydami połączonymi kaskadowo otrzymuje się inny przyrząd o bardzo interesujących właściwościach. Schemat ideowy układu przedstawiono na rys. 8.

Jeżeli w układzie pokazanym na rys. 8 $\theta_{31'}$ będzie równa $\theta_{42'}$ wtedy sygnał dochodzący z gałęzi *1* przedostanie się w całości do gałęzi *4'*. Jeżeli zaś $\theta_{31'}$ będzie się różnić od $\theta_{42'}$ o 180°, wtedy sygnał z gałęzi *1* w całości przejdzie do gałęzi *3'*. Zmieniając względne przesunięcie fazy $(\theta_{31'} - \theta_{42'})$ w granicach od 0° do 180° można osiągnąć dowolny podział mocy pomiędzy gałęzie *3'* i *4'*.



Rys. 8. Hybrydowy dzielnik mocy

Równanie (10) można teraz napisać w postaci:

$$S_{14'} e^{-j\theta_{14'}} = S_{13}\sqrt{1 - S_{13}^2}\, e^{-j\left(2\theta_{13} + \frac{\pi}{2}\right)} \left[e^{-j\Phi} + 1\right] e^{-j\theta_{42'}}, \tag{15}$$

gdzie $\Phi = \theta_{31'} - \theta_{42'}$ jest przesunięciem fazy wprowadzonym przez przesuwnik. Moduł równania (15) jest określony jako

$$S_{14'} = 2 S_{13}\sqrt{1 - S_{13}^2}\cos\frac{\Phi}{2}, \tag{16}$$

zaś argument

$$\theta_{14'} = 2\theta_{13} + \frac{\pi}{2} + \theta_{42'} + \frac{\Phi}{2}. \tag{17}$$

Optimum $S_{14'}$ względem S_{13} można uzyskać z wyrażenia

$$\frac{dS_{14'}}{dS_{13}} = 0 \tag{18}$$

dla wszystkich możliwych wartości Φ. Będzie więc

$$\frac{dS_{14'}}{dS_{13}} = 2\cos\frac{\Phi}{2}\sqrt{1 - S_{13}^2}\left[1 - \frac{S_{13}^2}{1 - S_{13}^2}\right] = 0. \tag{19}$$

Jedyną możliwą do przyjęcia wartością S_{13} jest więc $\frac{1}{\sqrt{2}}$ i wtedy $S_{14} = \frac{1}{\sqrt{2}}$. Podstawiając te zależności do równania (16) otrzymuje się

$$S_{14'} = \cos\frac{\Phi}{2}, \tag{20}$$

natomiast przenoszenie mocy w układzie jest określane jako

$$S_{14'}^2 = \cos^2\frac{\Phi}{2}$$
$$0 \leqslant S_{14'}^2 \leqslant 1 \quad \text{dla} \quad 0 \leqslant \Phi \leqslant \pi. \tag{21}$$

Dla wszystkich innych wartości $0 \leqslant S_{13} \leqslant 1$ wystąpi zależność

$$0 \leqslant S_{14'}^2 = 4 S_{13}^2 [1 - S_{13}^2] \cos^2 \frac{\Phi}{2} < 1 \tag{22}$$

przy zmianach kąta w zakresie $0 \leqslant \Phi \leqslant \pi$.

Kąty fazowe współczynników rozproszenia można obecnie zapisać w postaci:

$$\theta_{14'} = 2\,\theta_{13} + \frac{\pi}{2} + \frac{\theta_{31'} + \theta_{42'}}{2}, \tag{23}$$

$$\theta_{13'} = \operatorname{arc\,tg} \frac{S_{13}^2 \sin(2\,\theta_{13} - \theta_{31'}) + S_{14}^2 \sin(2\,\theta_{14} - \theta_{42'})}{S_{13}^2 \cos(2\,\theta_{13} - \theta_{31'}) + S_{14}^2 \cos(2\,\theta_{14} - \theta_{42'})}, \tag{24}$$

$$\theta_{24'} = \operatorname{arc\,tg} \frac{S_{13}^2 \sin(2\,\theta_{13} - \theta_{42'}) + S_{14}^2 \sin(2\,\theta_{14} - \theta_{31'})}{S_{13}^2 \cos(2\,\theta_{13} - \theta_{42'}) + S_{14}^2 \cos(2\,\theta_{14} - \theta_{31'})}. \tag{25}$$

Ogólnie rzecz biorąc można więc powiedzieć, że łącząc ze sobą dwa sprzęgacze kierunkowe zawsze uzyska się nowy sprzęgacz kierunkowy. W idealnym przypadku połączenia dwu hybryd jak na rys. 8 uzyskuje się sprzęgacz opisany macierzą rozproszenia

$$\mathbf{S} = \begin{bmatrix} 0 & 0 & \sin\frac{\Phi}{2} & j\cos\frac{\Phi}{2} \\ 0 & 0 & j\cos\frac{\Phi}{2} & \sin\frac{\Phi}{2} \\ \sin\frac{\Phi}{2} & j\cos\frac{\Phi}{2} & 0 & 0 \\ j\cos\frac{\Phi}{2} & \sin\frac{\Phi}{2} & 0 & 0 \end{bmatrix}, \tag{26}$$

a więc sprzęgacz o współczynniku sprzężenia

$$C = \frac{1}{\cos^2 \frac{\Phi}{2}}. \tag{27}$$

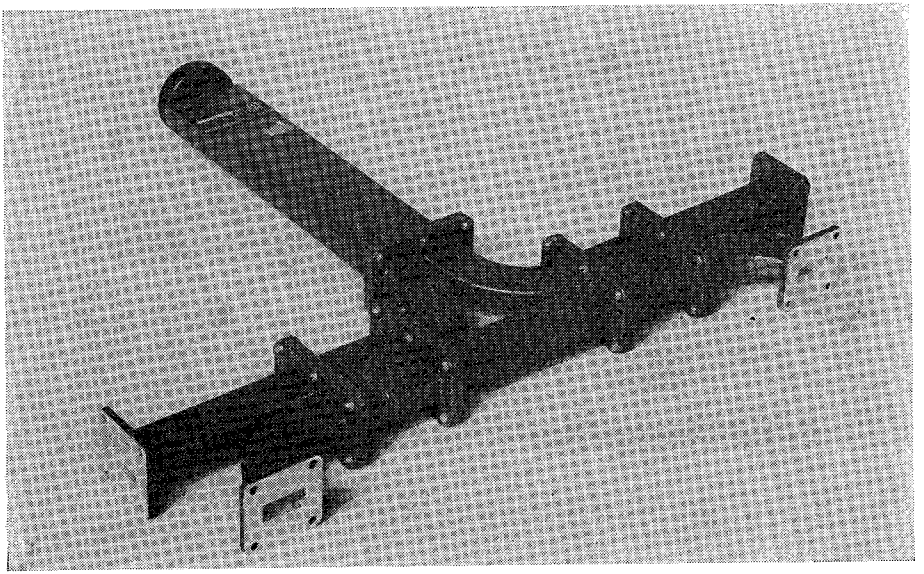
Ponieważ założono, że przesuwnik fazy w układzie na rys. 8 nie wprowadza strat mocy, można napisać wyrażenie na współczynnik nierównomierności podziału mocy

$$K = \frac{S_{13'}^2}{S_{14'}^2} = \frac{1 - 4 S_{13}^2 (1 - S_{13}^2) \cos^2 \frac{\Phi}{2}}{4 S_{13}^2 (1 - S_{13}^2) \cos^2 \frac{\Phi}{2}} = \frac{(1+k)^2}{4k \cos^2 \frac{\Phi}{2}} - 1, \tag{28}$$

gdzie k jest współczynnikiem nierównomierności podziału mocy samej hybrydy, określonym w równaniu (4).

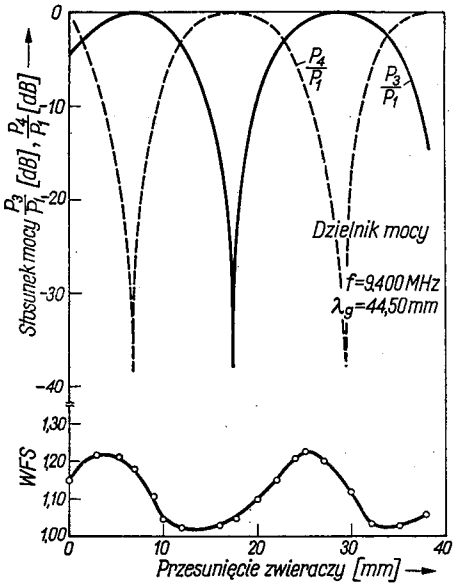
Z przytoczonych powyżej zależności widać, że układ praktyczny może w znacznym stopniu odbiegać od działania układu idealnego. Niedopasowanie przesuwnika fazy, niedopasowanie układu hybrydowego oraz nierównomierność podziału mocy będą powodować wzrost wejściowego współczynnika fali stojącej oraz skończony podział mocy pomiędzy gałęzie wyjściowe.

Na rys. 9 przedstawiono zdjęcie hybrydowego dzielnika mocy zbudowanego z falowodu FP 22 × 10, pracującego w pasmie 8,6 — 9,8 GHz,



Rys. 9. Widok hybrydowego dzielnika mocy na pasmo 3 cm

w którym jako przesuwnik fazy wykorzystano przesuwnik hybrydowy omówiony w rozdz. 3.2. Na rys. 10 przedstawiono charakterystykę podziału mocy oraz dopasowanie tego dzielnika mocy. W pasmie pracy wejściowy *WFS* dzielnika nie przekracza wartości 1,2, zaś podział mocy pomiędzy gałęzie wejściowe zawiera się w granicach od 0 dB do 36 dB. Tak więc współczynnik sprzężenia *C* określony równaniem (27) dla dowolnych zmian kąta przybiera wartości pomiędzy 1 a $2{,}5 \cdot 10^4$. Zerowy poziom odniesienia na rys. 10 odpowiada mocy wejściowej minus 0,3 dB, czyli mniejszej o wartość mocy traconej w miedzianych ściankach falowodu i w zwieraczach. Dzielnik mocy pokazany na rys. 9 pracował zadowalająco przy mocy 200 kW w impulsie bez śladów przebić. W porównaniu z innymi dzielnikami mocy, które zwykle charakteryzują się ograniczoną szerokością pasma, dużym współczynnikiem fali stojącej, ograniczoną wytrzymałością na przebicie, dużymi stratami lub skomplikowaną konstrukcją, hybrydowy dzielnik mocy można uważać za uniwersalny przyrząd laboratoryjny zarówno na małym, jak i na dużym poziomie mocy [9].

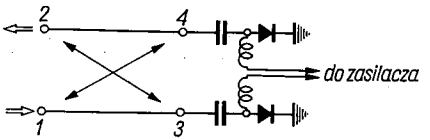


Rys. 10. Podział mocy i wejściowy WFS dzielnika mocy w funkcji położenia zwieraczy

4. INNE ZASTOSOWANIA UKŁADÓW HYBRYDOWYCH

4.1. Przełącznik-modulator

W przełączniku falowodowym pokazanym schematycznie na rys. 11 dwie diody mikrofalowe są polaryzowane napięciem zasilającym. W pewnych warunkach diody mogą być dopasowane do impedancji falowodu — na przykład gdy są spolaryzowane napięciem dodatnim. W innym przypadku, gdy na przykład przyłoży się do diod napięcie ujemne, będą one przedstawiały sobą jednakowe impedancje, dużo większe od impedancji falowodu. Działanie przełącznika falowodowego pokazanego na rys. 11 będzie więc następujące:



Rys. 11. Przełącznik mikrofalowy

Sygnał dochodzący do gałęzi *1* podzieli się w układzie hybrydowym pomiędzy gałęzie *3* i *4*. Jeżeli diody są dopasowane do impedancji falowodu, wtedy cała moc zostanie zaabsorbowana przez diody, a przełącznik znajduje się w pozycji „wyłączony". Gdy zmieni się napięcie zasilające,

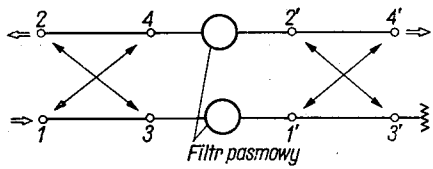
tak że diody będą odbijały sygnały w gałęziach *3* i *4* z powrotem w kierunku hybrydy, wtedy cały sygnał zostanie przeniesiony do gałęzi *2*, co odpowiada przełącznikowi w pozycji „włączony". Tak więc układ taki jest mikrofalowym przełącznikiem sterowanym elektrycznie i dopasowanym w obu położeniach. Jeżeli napięcie zasilające zmienia się w sposób ciągły od wartości ujemnych do dodatnich i z powrotem, na przykład napięcie sinusoidalne lub piłozębne, wtedy układ spełnia rolę modulatora sygnałów wielkiej częstotliwości przesyłanych z gałęzi *1* do gałęzi *2*. Poziom modulowanej mocy jest ograniczony wytrzymałością diod na zniszczenie.

Podobne operacje można przeprowadzać na dużym poziomie mocy (rzędu 5 MW w impulsie i 10 kW mocy średniej w pasmie 10 cm), jeżeli pomiędzy hybrydami umieści się sterowany napięciowo przełącznik wykorzystujący zjawisko rezonansu elektronów wtórnych [12]. Zaletą tego przełącznika jest krótki czas przełączania, rzędu 10^{-9} s, tak, że układ taki można wykorzystać do rozdzielania i dzielenia impulsów dużej mocy, do kształtowania impulsów prostokątnych o czasie trwania rzędu nanosekund oraz do modulacji dużego poziomu mocy.

Diody w układzie pokazanym na rys. 11 można zastąpić waraktorami, które pod wpływem zmian przyłożonego napięcia zmieniłyby swą reaktancję. Faza sygnałów przenoszonych do gałęzi *2* zmieniałaby się wtedy względem fazy sygnałów w gałęzi *1* w takt zmian napięcia zasilającego i układ taki spełniałby funkcje przesuwnika fazy sterowanego elektrycznie.

4.2. Rozdzielanie sygnałów o różnych częstotliwościach

Schemat ideowy układu służącego do rozdzielania sygnałów o różnych częstotliwościach pokazano na rys. 12. Pomiędzy dwoma identycznymi



Rys. 12. Układ rozdzielający

hybrydami umieszczone są dwa identyczne filtry pasmowe. Sygnał wchodzący gałęzią *1* podzieli się równo pomiędzy gałęzie *3* i *4*. Jeżeli częstotliwość sygnału leży w pasmie przenoszenia filtrów, wtedy sygnały przejdą do gałęzi *1'* i *2'* i po połączeniu w drugiej hybrydzie przedostaną się do gałęzi *4'*. Jeżeli zaś częstotliwość przychodzących sygnałów leży w pasmie zaporowym filtrów, wtedy sygnały te zostaną odbite, połączą się

w pierwszej hybrydzie i przedostaną się do gałęzi 2. Tak więc układ umożliwia rozdzielenie sygnałów o różnej częstotliwości dochodzących do gałęzi 1 pomiędzy gałęzie 2 i 4', zależnie od charakterystyki filtrów. Układ jest dopasowany zarówno w pasmie zaporowym, jak i w pasmie przenoszenia filtrów.

Podobne właściwości uzyskać można w układzie, w którym pomiędzy hybrydami umieści się odpowiednio dobrany przesuwnik fazy lub prowadnicę o określonej długości (zob. rys. 8). Jeżeli odcinek prowadnicy posiada taką długość, że różnica dróg równa jest całkowitej ilości fali o częstotliwości f_1 i nieparzystej wielokrotności połówek fali o częstotliwości f_2, wtedy sygnały o częstotliwości f_1 przedostaną się do gałęzi 4', zaś sygnały o częstotliwości f_2 do gałęzi 3'. Różnicę dróg d można prosto wyznaczyć z zależności

$$d = \frac{\lambda_{g1}\,\lambda_{g2}}{2\,(\lambda_{g2} - \lambda_{g1})}, \tag{29}$$

gdzie λg_1 — długość fali w falowodzie dla sygnału o częstotliwości f_1, a λg_2 — długość fali w falowodzie dla częstotliwości f_2.

W podany powyżej sposób można rozdzielać sygnały o wielu częstotliwościach, gdyż podobne układy rozdzielające można łączyć kaskadowo. Możliwe jest również dodawanie sygnałów o różnych częstotliwościach, gdyż zasada działania układów nie zmieni się przez zmianę kierunku strzałek na rys. 8 i 12.

4.3. Przełącznik antenowy

Gdy filtry pokazane na rys. 12 zastąpi się lampami zwierakowymi (lampami NO), układ będzie spełniał wszystkie funkcje przełącznika antenowego [10]. Sygnał o małym poziomie mocy będzie przedostawał się z gałęzi 1 (antena) do gałęzi 4' (odbiornik) poprzez niezjonizowane lampy. Natomiast sygnał dużej mocy z gałęzi 2 (nadajnik) zjonizuje lampy NO i zostanie odbity do gałęzi 1. Moc przeciekająca przez zjonizowane lampy przedostaje się do gałęzi 3', w której umieszczono dopasowane obciążenie.

Szereg dalszych układów, w których zostały zastosowane układy hybrydowe opisano w przytoczonej literaturze.

5. ZAKOŃCZENIE

Przedstawiono ogólną analizę właściwości szerokiej klasy układów mikrofalowych zawierających hybrydy falowodowe, która może być przydatna przy wyznaczaniu parametrów tych układów. Na szeregu przykładach wykazano przewagę takich przyrządów nad przyrządami klasycz-

nymi. Układy zawierające hybrydy są dopasowane w szerokim pasmie częstotliwości bez względu na dopasowanie gałęzi wyjściowych, a ich wytrzymałość na przebicie nie odbiega zasadniczo od wytrzymałości pustego falowodu. Te cechy głównie decydują o rosnącym znaczeniu takich układów w technice mikrofalowej.

Politechnika Warszawska
Katedra Techniki Fal Ultrakrótkich

WYKAZ LITERATURY

1. Sutherland J. W.: Waveguide Hybrid Circuits and their Use in Radar Systems. — Electronic Eng., t. 28, Nr 345, str. 464—469, listopad 1956, również Postępy Radiotechniki Nr 9, str. 13—26, 1957.
2. Teeter W. L., Bushore K. R.: A Variable-Ratio Microwave Power Divider and Multiplexer. — IRE Trans., t. MTT-5, Nr 4, str. 227—229, październik 1957.
3. Miloševic Lj., Vautey R.: Le résonateur à ondes progressives. L'Onde Electrique, t. 37, Nr 360, str. 290—294, marzec 1957.
4. Kraszewski A.: Mikrofalowe układy porównujące. — Przegląd Telekom., t. 33, Nr 1, str. 12—16, styczeń 1960.
5. Vaillancourt R. M.: Analysis of the Variable-Ratio Microwave Power Divider. — IRE Trans., t. MTT-6, Nr 2, str. 238, kwiecień 1958.
6. Mandelsztam M. J.: Riegulirujemyj naprawlennyj otwietwitiel s postojannym sdwigom fazy. — Radiot. i elektronika t. 6, Nr 10, str. 1761—1764, październik 1961.
7. Vařecha K.: Mikrovlnný dělič velkych vykonu. — Slaboupr. Obz., t. 20, Nr 5, str. 292—295, maj 1959.
8. Mohr R. J.: Some Design Aspects of Components Utilizing Symmetric Hybrids. — Microwave Jour. t. 5, Nr 6, str. 90—94, czerwiec 1962.
9. Kraszewski A., Wizner W.: O pewnych właściwościach rezonatorów pierścieniowych z falą bieżącą. — Prace PIT, Nr 38, str. 11—25, 1962.
10. Kraszewski A.: Wpływ niesymetrii kanałów na pracę zrównoważonego przełącznika antenowego. — Zeszyty Naukowe Politechniki Warszawskiej, Elektryka, nr 31, 1963.
11. Kraszewski A.: Falowodowy układ hybrydowy z krótką szczeliną. Prace PIT, nr 43, 1963.
12. Forrer M. P., Milazzo C.: Duplexing and Switching with Multipactor Discharges. — Proc. IRE, t. 50, str. 442—450, kwiecień 1962.
13. Wong J. Y.: A Mechanical Waveguide Hybrid Phase Changer. — IRE Trans., t. MTT-9, str. 364, lipiec 1961.
14. Wanselow R. D.: A Nonreflective Hybrid Stop-Band Filter. — IRE Trans., t. MTT-10, str. 91, styczeń 1962.
15. Wizner W.: Potrójna zwrotnica mikrofalowa. — Prace PIT, Nr 35, str. 25—28, 1961.

A. KRASZEWSKI

APPLICATION OF WAVEGUIDE HYBRID JUNCTIONS IN MICROWAVE TECHNIQUES

Summary

General analysis of properties of the broad class of microwave circuits containing waveguide hybrid junctions is presented, and application of the analysis by means of scattering matrix of the circuit to the design of microwave instruments is indicated. From the scattering matrix of the hybrid junction, the expressions for the scattering coefficients of a circuit composed of two identical hybrids and two arbitrary four-terminal networks between them, have been derived. General expressions can be radically simplified when the outputs of all branches are matched. The derived equations can be directly applied for the practical analysis of circuits having symmetric hybrids.

The analysis of practical circuits using hybrids includes expressions for input reflection coefficient of a hybrid tuner and hybrid phase shifter. It is shown that these instruments have remarkable bandwidth and their power handling capability is about the same as for an empty waveguide. These instruments can thus be used in numerous laboratory sets at high as well as at low power level.

The described hybrid power divider has similiar properties and as follows from theoretical considerations, it is matched for all positions of the phase shifter located between hybrids. Effects of nonideal hybrid on instrument parameters have been discussed. Characteristics of the instruments operating in the 9 Gc/s frequency band and made from the standard waveguide IEC R-100 (WG-16 or WR-90) have been included. The obtained data are in fine agreement with theoretical predictions stated above. Other possible applications of hybrid junctions in microwave techniques are mentioned.

A. KRASZEWSKI

APPLICATION DES ENSEMBLES HYBRIDES À LA TECHNIQUE DES MICROONDES

Résumé

Après une brève analyse montrant les performances d'un large groupe des ensembles hybrides construit en guide d'ondes rectangulaire, on écrit l'utilisation du calcul de matrice dans l'analyse des ensembles microondes. En utilisant la matrice de réflexion, on déduit les expressions générales caractérisant l'ensemble microonde composé de deux coupleurs hybrides, parmi lequels on place deux quadripôles différents. Dans le cas quand les terminaisons adaptées sont placées dans toutes les sorties de l'ensemble, on peut simplifier évidemment les expressions générales. Les équations permettent une analyse directe des ensembles microondes pratiques contenant les hybrides.

On déduit l'expression du T.O.S. présénte par le déphaseur hybride et l'adapteur hybride. L'auteur indique, que les ensembles hybrides possèdent les performances électriques suivantes: largeur de bande et tenue en puissance très remarquable.

On présente aussi les caractéristiques du diviseur hybride de puissance, indiquant leurs dépendance de performance des coupleurs hybrides nonidéals utilisés dans l'ensemble.

Les résultats des mesures des ensembles désignés pour la bande X (9000 Mc), réalisés en guide d'ondes rectangulaire, confirment bien la théorie.

A. KRASZEWSKI

ANWENDUNG VON HOHLLEITERN HYBRIDISCHEN RICHTKOPPLERSYSTEMEN IN DER MIKROWELLENTECHNIK

Zusammenfassung

Hier wurde eine Gesamtanalyse von Eigenschaften der breiten Mikrowellenbereichklasse, welche hybridische Richtkoppler besitzen, sowie die Anwendbarkeit der Streumatrixanalyse bei der Projektierung von Mikrowellenapparaten, dargestellt.

Aus der Streumatrix des hybridischen Systems wurden Bezeichnungen für die Streumatrixkoeffizienten in Systemen, die zwei egale Hybriden sowie beliebige Vierlinge, die zwischen ihnen eingesetzt sind, eingeführt. Die allgemeinen Bezeichnungen werden bedeutend vereinfacht, wenn in allen Ausgangszweigen angepasste Abschlusswiderstände angelegt sind. Die angeführten Gleichungen sind für die praktische Systemanalyse hybridischer Richtkoppler unbedingt notwendig. Im Rahmen der praktischen Systemanalyse wurden sie für den Eingangs-Reflexionskoeffizient des hybridischen Abstimmers sowie für den hybridischen Phasenschieber eingeführt. Es wurde gezeigt, dass Apparate, die auf der Grundlage der Eigenschaften hybridischer Systeme gebaut wurden, einen bedeutenden Frequenzbereich besitzen und ihre Durchschlagsfestigkeit annähernd gleich der Festigkeit des Hohlleiters ist, was ihre Verwendbarkeit in zahlreichen Messystemen, sowohl beim grossen als auch beim kleinen Leistungsniveau, findet.

Ähnliche Eigenschaften besitzt der erwähnte hybridische Leistungsteiler, der wie aus der erwähnten Theorie allen Einstellungen des Phasenschiebers angepasst und zwischen den Hybriden eingesetzt ist. Besprochen wurde auch der Einfluss der nicht idealen Hybriden auf die Parameter der Apparate, in welchen sie arbeiten. Die angegebenen Charakteristiken der Apparate, welche im Frequenzbereich von 9000 MHz arbeiten, beziehen sich auf die, welche aus dem standartisierten Hohlleiter hergestellt sind. Die erzielten Ergebnisse stimmen mit den Werten überein, die bei den theoretischen Betrachtungen vorgesehen waren. Zuletzt wurden noch andere Verwendungsmöglichkeiten für Hybridensysteme in der Mikrowellentechnik angeführt.

А. КРАШЕВСКИ

ANWENDUNG VON HOHLLEITERN HYBRIDISCHEN RICHTKOPPLERSYSTEMEN IN DER MIKROWELLENTECHNIK

Резюме

В статье излагается общий анализ свойств широкого класса микроволноводных контуров, построенных из балансных соединений, и доказывается пригодность анализа матрицы рассеяния контура для проектирования приборов сверх-

высоких частот. Исходя из матрицы рассеяния балансного соединения выведены уравнения для коэффициентов рассеяния цепи, содержащей два балансных соединения и два произвольных четырехполюсника, помещенных между ними. Общие уравнения значительно упрощаются, если в выходных ответвлениях находятся согласованные нагрузки. Приведенные уравнения используются непосредственно для анализа практических схем содержащих балансные соединения.

Выведены уравнения для входного коэффициента отражения волноводного фазовращателя и системы настройки со щелевыми мостами. Указывается, что эти приборы характеризуются весьма значительной широкополосностью и сопротивляемостью на пробой, приближенной к сопротивляемости пустого волновода. Таким образом, они найдут применение в измерительных установках так на большом как и на малом уровне мощности.

Волноводный делитель мощности, состоящий из трех щелевых мостов, как следует из представленной теории, согласован при всех положениях фазовращателя. Рассмотрено влияние неидеального балансного соединения на параметры приборов. Представлены характеристики приборов со щелевыми мостами работающих в диапазоне 9000 мгц изготовленных из стандартного волновода. Эти данные совпадают с величинами результирующими из представленной теории. Указаны другие возможности применения балансных соединений в технике сверхвысоких частот.

621.314.7

ZDZISŁAW KORZEC

Praca tranzystorów warstwowych przy dużych sygnałach w świetle teorii ładunkowej

Rękopis otrzymano 10.9.1962

W artykule omówiono zagadnienia stanów przejściowych w tranzystorach warstwowych pracujących przy dużych sygnałach, w układzie wspólnego emitera, przy obciążeniu czynnym. Wykazano słuszność rozpatrywania tych zagadnień w oparciu o parametry ładunkowe, stanowiące wartości średnie dla dużych sygnałów, w przeciwieństwie do parametrów małych sygnałów określonych dla danego punktu pracy. Podano prostą metodę pomiaru parametrów ładunkowych obszaru czynnego i obszaru nasycenia. Przeprowadzono porównanie zależności charakteryzujących czasy załączenia i wyłączenia, wyprowadzonych w oparciu o parametry ładunkowe i parametry małych sygnałów.

OZNACZENIA

C_{Tc} — pojemność bariery kolektorowej

I_b, I_c, I_e — odpowiednio: prąd stały bazy, kolektora, emitera

Q_b — ładunek w obszarze bazy

Q_V — ładunek pojemności bariery kolektorowej

Q_{bs} — dodatkowy ładunek w obszarze bazy przy nasyceniu

T_b, $_c$, T_e — stałe czasu — odpowiednio: bazy, kolektora emitera (dla obszaru czynnego)

T_{bs} — stała czasu bazy w obszarze nasycenia

U_{ce}, U_{be}, U_{cb} — napięcie między zaciskami — odpowiednio: kolektor-emiter, baza-emiter, kolektor-baza

$U_{cb'}$, $U_{eb'}$ — napięcie na złączach — odpowiednio: kolektor-baza, emiter-baza

U_g — napięcie generatora

U_{cc} — napięcie zasilające obwód kolektora

α_0 — współczynnik wzmocnienia prądowego dla małych sygnałów, w układzie *OB*, przy $U_{CB} = 0$

α_I — jak wyżej, przy wzajemnej zamianie roli emitera i kolektora

β — współczynnik wzmocnienia prądowego w układzie *OE* (dla dużych sygnałów)

β_s — współczynnik wzmocnienia prądowego w obszarze nasycenia
τ_p — czas życia nośników dziurowych
ω_{T0} — pulsacja graniczna (przy $U_{CB} = 0$)
ω_{TI0} — jak wyżej, przy wzajemnej zamianie roli emitera i kolektora

1. WSTĘP

Analiza pracy tranzystorów warstwowych przy dużych sygnałach stanowi z punktu widzenia rozważań teoretycznych zagadnienie trudne. Pierwsze prace z tej dziedziny ukazały się w roku 1954. Są to dziś uznawane za klasyczne dwie prace Molla i Ebersa [5], [6]. W jednej z nich są omawiane własności tranzystora przy dużych sygnałach [1] (zależności między prądami i napięciami w poszczególnych obszarach pracy), w drugiej zaś — zagadnienia stanów przejściowych w tranzystorach przy pracy z dużymi sygnałami. Czasy narastania, magazynowania nośników (opóźnienia) i opadania zostały podane w funkcji parametrów charakteryzujących własności tranzystorów przy małych sygnałach.

W roku 1957 Beaufoy i Sparkes opublikowali swoją pracę „The Junction Transistor as a Charge-Controlled Device". W pracy tej zostało przedstawione całkowicie nowe ujęcie teoretyczne pracy tranzystora zarówno przy małych, jak i dużych sygnałach. Mianowicie tranzystor rozpatrywany jest jako element sterowany ładunkiem gromadzonym w obszarze bazy (w przeciwieństwie do powszechnie stosowanego ujęcia, według którego tranzystor traktowany jest jako element sterowany prądowo). Spotykana w teorii tranzystorów warstwowych analiza rozkładu ładunków w obszarze bazy nie wiąże tego rozkładu z prądami: wejściowym i wyjściowym w taki sposób, jak czyni to Beaufoy i Sparkes. Zdefiniowane przez nich parametry ładunkowe z dużą dokładnością odzwierciedlają własności fizyczne tranzystorów, a jednocześnie mogą być bezpośrednio wykorzystane do projektowania układów impulsowych.

Obecnie szereg firm produkujących tranzystory stosowane w układach impulsowych i przełączających podaje dane techniczne w oparciu o parametry ładunkowe. Ponieważ w literaturze krajowej dotąd bardzo niewiele zostało na ten temat powiedziane, wydaje się celowe nieco szersze omówienie tego zagadnienia.

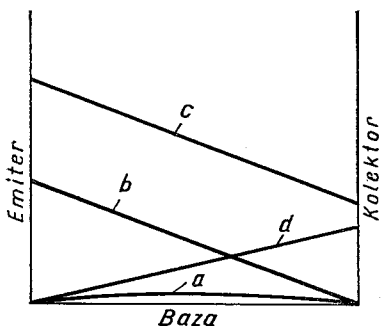
W artykule niniejszym zostaną omówione stany przejściowe w tranzystorach warstwowych w oparciu o parametry ładunkowe. Zależności charakteryzujące stany załączenia i wyłączenia wyprowadzone w oparciu o teorię ładunkową zostaną porównane ze wzorami podanymi przez Molla i Ebersa.

[1] Przez pracę z dużymi sygnałami rozumie się pracę, przy której tranzystor przechodzi z obszaru odcięcia poprzez obszar czynny do obszaru nasycenia — lub odwrotnie.

2. TRANZYSTOR WARSTWOWY JAKO ELEMENT STEROWANY ŁADUNKIEM BAZY

W tranzystorze warstwowym typu p-n-p wstrzykiwane do obszaru bazy dziury poruszają się na skutek mechanizmu dyfuzji [1] w kierunku od emitera do kolektora (lub odwrotnie w przypadku pracy inwersyjnej tranzystora). Gęstość prądu dziurowego jest proporcjonalna do gradientu gęstości dziur w obszarze bazy. Ponieważ wstrzyknięcie dziur do obszaru bazy typu n wymaga — dla zachowania neutralności ładunku — wstrzyknięcia takiej samej ilości dodatkowych elektronów, można mówić o istnieniu w obszarze bazy ładunku Q_b (składającego się z dziur i dodatkowych elektronów [2]). Dodatkowe elektrony są doprowadzane do obszaru bazy poprzez zacisk bazy — przyjmując, że prąd przepływający przez kolektor jest wyłącznie dziurowy.

Gdyby w obszarze bazy nie występowała rekombinacja, a sprawność emitera była równa jedności, wstrzyknięty do obszaru bazy ładunek Q_b powodowałby przepływ prądu od emitera do kolektora aż do chwili, gdy ładunek ten zostałby usunięty. Ponieważ w rzeczywistości występuje rekombinacja powodująca zmniejszenie znajdującego się w obszarze bazy ładunku — musi płynąć określony prąd bazy podtrzymujący wymaganą



Rys. 1. Rozkład ładunku w bazie: a — w obszarze odcięcia, b — w obszarze czynnym, c — w obszarze nasycenia, d — przy pracy inwersyjnej (wzajemnej zamianie roli emitera i kolektora)

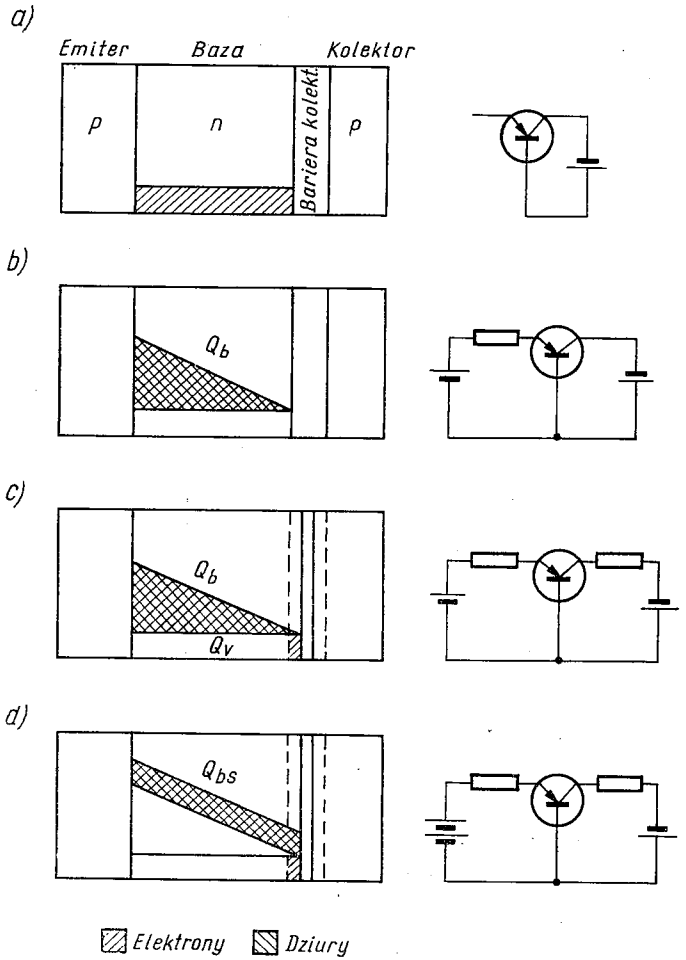
jego wartość. Każdemu przy tym punktowi pracy tranzystora, określonemu w obszarze czynnym i odcięcia przez U_{ce} i I_c, zaś w obszarze nasycenia przez I_c i I_e, odpowiada określona gęstość ładunku rozłożonego w bazie. Dla stanu ustalonego rozkład ten jest jednoznaczny, tzn. każdy rozkład ładunku odpowiada jednej wartości napięcia i prądu — i odwrotnie. Na rys. 1 przedstawiono rozkład ładunku w obszarze bazy dla czterech

[1] Pomijając wpływ działania pola elektrycznego.

[2] W tym ujęciu obszar bazy gromadzący ładunki dodatnie obok ujemnych można uważać za kondensator, co znalazło swoje odzwierciedlenie w schemacie zastępczym tranzystora dla dużych sygnałów (zob. wykaz literatury poz. [1]).

obszarów pracy: odcięcia, nasycenia, obszaru czynnego i pracy inwersyjnej.

W obszarze odcięcia ilość dziur wstrzykiwanych do obszaru bazy, a zatem i ilość dodatkowych elektronów można przyjąć za równą zeru. Jedynym ładunkiem w obszarze bazy jest wtedy ładunek elektronów będący w równowadze z ładunkiem donorów.



Rys. 2. Rozkład ładunku w obszarze bazy z uwzględnieniem ładowania pojemności bariery kolektorowej

a) tranzystor w stanie odcięcia, b) załączenie tranzystora bez zmiany napięcia kolektora, c) załączenie tranzystora przy określonej oporności obciążenia, d) wprowadzenie tranzystora w obszar nasycenia

W obszarze czynnym gęstość dziur przy emiterze utrzymuje się na poziomie odpowiadającym polaryzacji w kierunku przewodzenia, zaś w pobliżu kolektora jest bliska zeru. Gradient gęstości ładunku jest prawie liniowy i proporcjonalny do prądu płynącego od emitera do kolektora. Istnienie pewnej niewielkiej różnicy między prądem kolektora

i emitera znajduje swoje odzwierciedlenie w nieliniowości gęstości gradientu, ale wpływ tego zjawiska na całkowity ładunek bazy może być pominięty, jeżeli współczynnik wzmocnienia prądowego w układzie *OB* jest w przybliżeniu równy jedności.

Podobny rozkład ładunku w obszarze bazy występuje przy pracy inwersyjnej tranzystora (w przypadku tranzystorów asymetrycznych współczynnik wzmocnienia prądowego przy pracy rewersyjnej jest mniejszy niż przy pracy normalnej na skutek zwiększonego udziału rekombinacji powierzchniowej).

W obszarze nasycenia ładunek dziur i elektronów w pobliżu emitera, jak również kolektora, przyjmuje określoną wartość, różną od zera, natomiast gradient ładunku pozostaje nadal liniowy.

2.1. Obszar czynny

2.1.1. Stałe czasu T_c, T_b, T_e

Jeżeli tranzystor zostanie przełączony z obszaru odcięcia do obszaru czynnego przy zwartym obwodzie kolektora (tzn. bez zmiany napięcia na kolektorze), wtedy wstrzyknięty do obszaru bazy ładunek Q_b powoduje przepływ prądu kolektora I_c. Współczynnik proporcjonalności między ładunkiem bazy Q_b i prądem kolektora I_c jest jednym z podstawowych parametrów teorii ładunkowej [1] i nosi nazwę stałej czasowej kolektora

$$T_c = Q_b/I_c . \tag{1}$$

Ładunek magazynowany w obszarze bazy jest zależny od szerokości bazy — im jest węższy obszar bazy, tym mniejszy jest ładunek wymagany do powstania danego prądu. Ponieważ szerokość bazy jest modulowana wartością napięcia kolektor-baza, zatem stała czasowa T_c zależy również od wartości tego napięcia. W większości zastosowań impulsowych stan przewodzenia tranzystora odpowiada warunkowi $U_{cb} = 0$, zatem stałą czasową kolektora T_{c0} można określić jako zależność

$$T_{c0} = Q_b/I_c \text{ przy } U_{cb} = 0. \tag{2}$$

Ponieważ stała czasowa zależy w pewnej mierze od wartości prądu emitera, zatem dokładniej można ją zdefiniować jako stosunek przyrostu ładunku bazy ΔQ_b do przyrostu prądu ΔI_c przy stałej wartości prądu I_e i stałej wartości napięcia U_{cb};

$$T_c = \Delta Q_b/\Delta I_c \tag{3}$$

przy $U_{cb} = \text{const}$ oraz $J_E = \text{const}$.

---

[1] Proporcjonalność ta jest słuszna przy omówionym wyżej założeniu liniowości gradientu ładunku w bazie.

Podobnie można zdefiniować stałą czasową bazy i emitera, a mianowicie

$$T_b = \Delta Q_b / \Delta I_b , \tag{4}$$

$$T_e = \Delta Q_b / \Delta I_e . \tag{5}$$

Stałe te można uzależnić od współczynnika wzmocnienia prądowego α_0 oraz od pulsacji granicznej ω_{T0} [1]

$$T_c = \frac{1}{\alpha_0 \omega_{T0}}, \tag{6}$$

$$T_b = \tau_p = \frac{1}{\omega_{T0}(1-\alpha_0)}, \tag{7}$$

$$T_e = \frac{1}{\omega_{T0}}. \tag{8}$$

Powyższe stałe czasowe określają w pewnym sensie warunki, jakie powinien spełniać obwód wejściowy tranzystora, aby załączenie nastąpiło natychmiastowo. Jeżeli np. tranzystor w układzie *OE* jest sterowany z generatora napięć prostokątnych przez równoległe połączenie oporności R_1 i pojemności C_1, wtedy kształt przebiegu prądu wyjściowego będzie dokładnie odpowiadał kształtowi napięcia wejściowego (przy założeniu, że pojemność kolektora i oporność bazy mogą być pominięte), gdy $T_b = C_1 R_1$. W czasie zmian napięcia generatora pojemność C_1 będzie powodowała wprowadzenie ładunku Q_b do obszaru bazy, zaś oporność R_1 będzie pozwalała na przepływ prądu I_b o wartości potrzebnej do utrzymania wymaganego ładunku. Jeżeli jest pożądane dwukrotne zwiększenie wartości prądu I_c, wtedy C_1 należy dwukrotnie zwiększyć, zaś oporność R_1 dwukrotnie zmniejszyć.

2.1.2. Pojemność złącz emitera i kolektora

Zazwyczaj przełączeniu tranzystora z jednego punktu pracy do drugiego towarzyszy zmiana napięcia kolektor-baza jak i napięcia emiter-baza, co — jak wyżej wspomniano — prowadzi do zmian szerokości bazy i ładunku bazy. Przy tym zwiększeniu prądu płynącego od emitera do kolektora towarzyszy zmniejszenie szerokości bariery kolektor-baza, jak i emiter-baza, a co zatem idzie — zwiększenie szerokości bazy. Ponieważ pojemność bariery emiter-baza jak i zmiany napięcia na tym przejściu są stosunkowo niewielkie, wpływ pojemności emiter-baza może być pominięty, a stałe czasu dużych sygnałów mogą być w przybliżeniu przyjęte jako stałe niezależne od wartości przepływającego prądu.

Zmianom napięcia kolektor-baza towarzyszą dwa zjawiska:

1) zmiana ładunku Q_b na skutek zmiany szerokości bazy,
2) „ „ Q_v potrzebnego do naładowania pojemności bariery kolektor-baza.

Należy zwrócić uwagę na to, że ładunek Q_b magazynowany w obszarze bazy (modulowany napięciem U_{cb}) stanowi ładunek przestrzenny jednakowej ilości dziur i elektronów, natomiast ładunek Q_v potrzebny do naładowania pojemności bariery kolektor-baza od strony bazy składa się wyłącznie z elektronów neutralizujących ładunek donorów.

Ładunek Q_v jest powiązany z pojemnością bariery kolektorowej C_{Tc}, pomierzonej dla małych sygnałów, zależnością:

$$Q_v = M \cdot C_{Tc} \cdot U_{cb}, \tag{9}$$

gdzie $U_{cb} = U_{cb1} - U_{cb2}$ jest zmianą napięcia U_{cb}, C_{Tc} pojemnością bariery kolektorowej pomierzonej przy napięciu początkowym U_{cb1}, M — współczynnikiem o wartości 1 do 2, zależnym od typu tranzystora i wartości napięcia U_{cb} [8].

2.1.3. Załączenie tranzystora w obszarze czynnym (bez nasycenia)

Wprowadzony do obszaru bazy ładunek Q_b (nie podtrzymywany przepływem prądu bazy I_b) na skutek występującej w obszarze bazy rekombinacji dziur i elektronów będzie malał zgodnie z funkcją wykładniczą. Szybkość rekombinacji jest wprost proporcjonalna do ładunku chwilowego, przy czym współczynnikiem proporcjonalności jest odwrotność czasu życia nośników $1/\tau_p = 1/T_b$, zatem

$$Q_b = Q_{b0}\, e^{-t/T_b} \tag{10}$$

oraz

$$\frac{dQ_b}{dt} = -\frac{Q_b}{T_b}, \tag{11}$$

stąd potrzebny do utrzymania danego ładunku Q_b prąd bazy wynosi

$$I_b = -\frac{dQ_b}{dt} = \frac{Q_b}{T_b} = \frac{Q_b}{\beta T_c}. \tag{12}$$

Ponieważ z zależności (1) $I_c = Q_b/T_c$, zatem w powiązaniu z (12) $\beta = I_c/I_b$, gdzie β — współczynnik wzmocnienia prądowego dla dużych sygnałów w układzie OE (pomiaru β dokonuje się zazwyczaj przy $U_{cb} = 0$ i oznacza się wtedy przez β_0).

---

[1]) Współczynnik M wyraża się zależnością:

$$M = \frac{2\left[U_b - U_{cb1} - \sqrt{(U_b + U_{cb1})(U_b - U_{cb2})}\right]}{U_{cb1} - U_{cb2}}$$

Zależność ta jest wynikiem obliczenia całki:

$$Q_v = \int_{U_{cb1}}^{U_{cb2}} C_{Tc} \cdot dU_{cb}$$

dla przypadku złącz stopowych, gdy $C_{Tc} = \text{const}\,(-U_b - U_{cb})^{-\frac{1}{2}}$, co daje w wyniku $Q_v = MC_{Tc1}\Delta U_{cb}$ (U_b — napięcie bariery potencjału).

Z powyższego wynika, że po przyłożeniu w momencie $t = 0$ prądu bazy $I_{b\,zał}$ ładunek w obszarze będzie wzrastał wykładniczo zgodnie z zależnością

$$Q_b = I_{b\,zał}\,\beta T_{c0}\left[1 - e^{-t/\beta T_{c0}}\right]. \tag{13}$$

Jeżeli przejściu tranzystora ze stanu odłączenia do stanu załączenia towarzyszy zmiana napięcia kolektora, wtedy musi być wprowadzony w obszar bazy dodatkowy ładunek Q_v. Całkowity ładunek załączenia, tzn. ładunek, jaki musi być wprowadzony do obszaru bazy, aby tranzystor z obszaru odcięcia przeszedł do granicy obszaru nasycenia wynosi:

$$Q_{zał} = Q_b + Q_v = I_c T_c + Q_v. \tag{14}$$

Ponieważ $\beta Q_v/I_c$ stanowi dodatkowy ładunek przypadający na jednostkę prądu bazy w obszarze czynnym, ładunek ten powoduje zmianę stałej czasowej z βT_c na $\beta\,(T_c + Q_v/I_c)$ [1] (ładunek Q_v związany z pojemnością kolektora nie powoduje zwiększenia rekombinacji). Zatem po przełożeniu skokowo prądu bazy $I_{b\,zał}$ ładunek w obszarze bazy będzie narastał wg zależności

$$Q_{zał} = I_{b\,zał}\,\beta(T_c + Q_v/I_c)\left[1 - e^{-\frac{t}{\beta(T_c + Q_v/I_c)}}\right]. \tag{15}$$

Definiując czas załączenia jako czas potrzebny na to, aby w obszarze bazy ustalił się wymagany ładunek $Q_{bzał}$ po przyłożeniu prądu bazy $I_b = I_{bzał}$ (zakładając, że początkowy prąd bazy był równy zeru) z zależności [15] otrzymujemy:

$$t_{zał} = \beta\,(T_c + Q_v/I_c)\ln I_{b\,zał}/(I_{b\,zał} - I_c/\beta). \tag{16}$$

Jak widać z powyższej zależności, czas załączenia jest tym krótszy, im większą wartość posiada prąd bazy $I_{bzał}$. Jest rzeczą oczywistą, że wstrzyknięcie do obszaru bazy ładunku Q_b w czasie równym zeru wymagałoby nieskończenie dużej wartości prądu sterującego bazy. W praktyce prąd wejściowy jest ograniczony przez impedancję obwodu sterującego, jak również przez oporność bazy. Stosowanie dużego prądu bazy jest korzystne tylko w pierwszym momencie i ma na celu, jak było powiedziane, skrócenie czasu załączania — po załączeniu może wprowadzić tranzystor zbyt głęboko w obszar nasycenia, co zwiększa czas wyłączenia.

---

[1] Tzn. w przypadku nieuwzględnienia ładunku Q_v: $I_b = Q_b/T_b$, gdzie $T_b = \beta T_c$, zaś przy uwzględnieniu ładunku Q_v: $I_b = (Q_b + Q_v)/T_{bv}$, stąd przy tym samym prądzie I_b (ponieważ Q_v nie powoduje zwiększenia rekombinacji)

$$T_{bv} = T_c\beta\,\frac{Q_b + Q_v}{Q_b} = T_c\beta + Q_v T_c\beta/I_c T_c = \beta\,(T_c + Q_v I_c).$$

2.2. Obszar nasycenia

2.2.1. Parametry obszaru nasycenia

Omówiony powyżej ładunek Q_{zal} był wystarczający, aby tranzystor został załączony z obszaru odcięcia do granicy obszaru nasycenia. W układach bardzo często przewiduje się (ze względu na tolerancję elementów, starzenie, różnice parametrów tranzystorów) sterowanie obwodu bazy prądem większym niż I_{bzal}. Powoduje to przejście tranzystora do obszaru nasycenia. W obszarze tym prąd kolektora jest ograniczony wartością oporności obciążenia i wartością napięcia zasilającego. Dalszy wzrost prądu bazy nie może w związku z tym prowadzić do zwiększenia prądu kolektora, a powoduje jedynie zmagazynowanie dodatkowego ładunku w obszarze bazy.

Jeżeli do obszaru bazy zostanie wstrzyknięty ładunek większy od potrzebnego, aby tranzystor znalazł się na granicy nasycenia, o ładunek dodatkowy Q_{bs} i jeżeli prąd bazy potrzebny do utrzymania tego ładunku nasycenia wynosi I_{bs}, wtedy całkowity ładunek bazy wyrazi się zależnością

$$Q_{cal} = I_c T_{c0} + Q_v + Q_{bs}, \tag{17}$$

przy czym całkowity prąd bazy

$$I_b = I_c/\beta + I_{bs} \tag{18}$$

oraz

$$Q_{bs} = T_{bs} I_{bs}, \tag{19}$$

gdzie T_{bs} jest stałą czasową charakteryzującą obszar nasycenia. Stała czasowa T_{bs} definiowana jest jako stosunek ładunku bazy przypadającego na jednostkę prądu bazy w obszarze nasycenia.

W tranzystorowych układach przełączeniowych zdarza się często, że obwód bazy tranzystora jest załączany przy małym prądzie kolektora (np. przy odłączonym obwodzie kolektora), a następnie na skutek zamknięcia obwodu kolektora, wymagany jest znacznie większy prąd kolektora, bez dodatkowej zmiany w obwodzie bazy (rys. 3). Aby przy zamkniętym obwodzie kolektora tranzystor znalazł się na granicy nasycenia, wtedy wymagany do przepływu prądu I_c ładunek bazy wynosi $I_c T_{c0}$. Jeżeli prąd kolektora (na skutek przerwania obwodu kolektora) zostaje zmniejszony do zera, bez zmiany prądu bazy, wtedy cały ładunek obszaru bazy staje się ładunkiem nasycenia $Q_{bs} = I_b T_{bs}$.

Warunkiem, aby po ponownym załączeniu obwodu kolektora popłynął wymagany prąd I_c, jest warunek, aby znajdujący się w obszarze bazy ładunek Q_{bs} był równy (lub większy) ładunkowi $Q_b = I_c T_{c0}$, zatem

$$I_b T_{bs} = I_c T_{c0}, \tag{20}$$

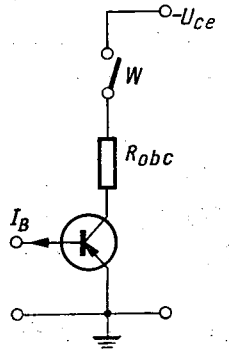
stąd

$$I_c/I_b = T_{bs}/T_{c0} = \beta_s . \tag{21}$$

Współczynnik wzmocnienia prądowego obszaru nasycenia β_s jest definiowany jako stosunek prądu kolektora, jaki może być uzyskany po zamknięciu obwodu kolektora, do stałego prądu sterującego bazy, przy czym

$$\beta_s = T_{bs}/T_{c0} . \tag{22}$$

Współczynnik β_s jest zazwyczaj mniejszy od współczynnika wzmocnienia β dla obszaru czynnego.



Rys. 3. Tranzystor z obwodem kolektora sterowanym dodatkowym wyłącznikiem W

2.2.2. *Wyłączenie*

Proces wyłączenia tranzystora jest związany z przejściem tranzystora z obszaru nasycenia do obszaru odcięcia. Całkowity ładunek, jaki musi być odprowadzony z obszaru bazy, przy założeniu, że tranzystor znajdował się w obszare nasycenia, wynosi

$$Q_{wył} = Q_b + Q_V + Q_{bs} = Q_V + I_c T_{c0} + (I_b - I_c/\beta) \cdot T_{bs} =$$
$$= Q_V + I_c T_{c0}(1 - \beta_s/\beta) + I_b T_{bs} . \tag{23}$$

Czas wyłączenia tranzystora można zatem podzielić na czas zmagazynowania potrzeby do usunięcia ładunku Q_{bs}, tzn. czas potrzebny na przejście punktu pracy z obszaru nasycenia na granicę obszaru czynnego, oraz czas opadania potrzebny do usunięcia pozostałego ładunku i sprowadzenie punktu pracy do obszaru odcięcia.

Zmniejszanie się prądu kolektora następuje dopiero po zakończeniu czasu magazynowania. Wyrażenia na czas magazynowania i czas opadania podano w tablicy. Zazwyczaj przy projektowaniu układu interesuje nas całkowity czas wyłączenia $t_{wył}$, tzn. czas potrzebny do usunięcia z obszaru bazy ładunku $Q_{zał}$ po przyłożeniu do obwodu sterującego bazy prądu $I_{b\,wył}$. Podobnie jak przy procesie załączania zmiana ładunku w obszarze

Czas narastania	Czas załączenia	Prąd załączenia
$t_n^{*)} = \frac{1}{(1-\alpha_N)\,\omega_N} \ln \frac{I_{b1}}{I_{b1} - 0{,}9\,I_c \frac{1-\alpha_N}{\alpha_N}}$	$t_{zał} = \beta\left(T_{c0} + \frac{Q_v}{I_c}\right) \ln \frac{I_{bzał}}{I_{bzał} - I_c/\beta}$	$I_{bzał} = \frac{I_c}{\beta\left[1 - e^{-t_{zał}/\beta(T_{c0} + Q_v/I_c)}\right]}$
$t_n = T_{bv} \ln \frac{I_{b1}}{I_{b1} - 0{,}9\,I_c \frac{1-\alpha_N}{\alpha_N}}$	$t_{zał} \approx \frac{I_c T_{c0} + Q_v}{I_{bzał} - 0{,}58\,I_c/\beta}$	$I_{bzał} \approx \frac{I_c T_{c0} + Q_v}{t_{zał}}\, t\, \frac{0{,}58\,I_c}{\beta}$
Csas magazynowania	**Czas wyłączenia**	**Prąd wyłączenia**
$t_s = T_s \ln \frac{I_{b1} T_b - I_{b2} T_b}{I_c T_{c0} - I_{b2} T_b}$	$t_{wył} \approx \frac{Q_{wył}}{I_{bwył} + 0{,}42\,(Q_{wył} - Q_v)/\beta T_{c0}}$	$I_{bwył} \approx \frac{Q_{wył}}{t_{wył}} - 0{,}42\,\frac{Q_{wył} - Q_v}{\beta T_{c0}}$
$t_s^{*)} = \frac{\omega_N + \omega_s}{\omega_N \omega_I (1 - \alpha_N \alpha_I)} \ln \frac{I_{b1} - I_{b2}}{\frac{I_c(1-\alpha_N)}{\alpha_N} - I_{b2}}$		

*) wg Molla i Ebersa.

bazy przy wyłączaniu będzie miała charakter wykładniczy. Ponieważ stała czasu T_s odnosi się tylko do obszaru nasycenia, zastąpienie jej przez stałą czasową obszaru czynnego βT_c spowoduje tylko nieznaczny błąd w wyrażeniu na czas wyłączenia (jeżeli $T_{bs} = 2$ µsek, $\beta T_c = 2{,}1$ µsek, to błąd z przyjęcia βT_c jako stałej czasowej dla obszaru czynnego i dla obszaru nasycenia wyniesie ok. 5%).

Podobnie jak przy załączaniu ładunek Q_v związany z pojemnością kolektora powoduje zmianę stałej czasowej załączenia βT_c zwiększając ją w stosunku $Q_{wył}/Q_{wył} - Q_v$.

Po przyłożeniu prądu wstecznego bazy $I_{b\,wył}$ ładunek bazy Q_b będzie malał wykładniczo do wartości $I_{b\,wył}\,\beta T_{co}\,Q_{wył}/(Q_{wył} - Q_v)$. Jeżeli założymy, że ładunek $Q_{wył}$ po przyłożeniu prądu $I_{wył}$ ma zostać usunięty po czasie $t_{wył}$, otrzymamy zależność

$$Q_{wył} = \left(Q_{wył} - I_{b\,wył}\,\beta T_{c0}\,\frac{Q_{wył}}{Q_{wył} - Q_v}\right)\left[1 - e^{-\frac{t_{wył}}{\beta T_c Q_{wył}/(Q_{wył} - Q_v)}}\right]. \tag{24}$$

Stosując przybliżenie

$$1 - e^{-t/T} \approx \frac{t/T}{1 + 0{,}58\,t/T} \tag{25}$$

otrzymamy uproszczone wyrażenie na czas wyłączenia

$$t_{wył} = \frac{Q_{wył}}{I_{bwył} + 0{,}42\,(Q_{wył} - Q_v)/\beta T_c}. \tag{26}$$

Podobnie jak przy załączaniu czas wyłączania jest tym krótszy, im większy jest prąd wyłączenia $I_{b\,wył}$.

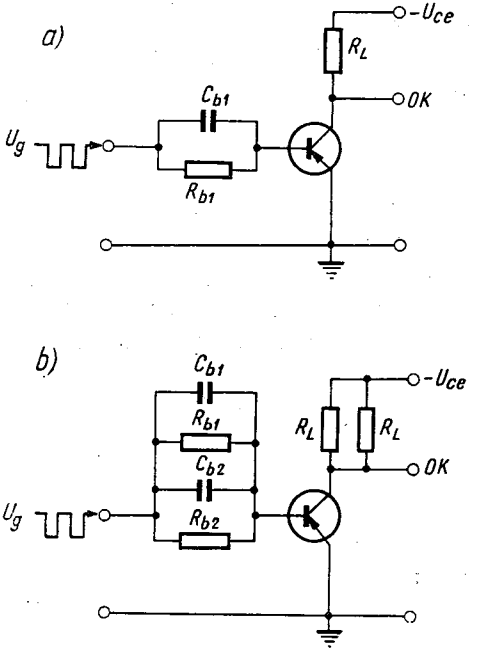
W tablicy 1 zestawiono najważniejsze zależności charakteryzujące pracę impulsową tranzystorów warstwowych. Wyrażenie na czas narastania określone jako czas, po którym prąd kolektora wynosi 0,9 końcowej wartości, zostało podane wg Molla i Ebersa, jak również w teorii ładunkowej. Pomijając wpływ pojemności złącza kolektor-baza (przyjmując $Q_v = 0$) otrzymuje się, korzystając z zależności (7), zgodność obu wyrażeń. Podobnie uzyskuje się zgodność zależności charakteryzujących czas magazynowania, przyjmując [3]

$$T_{bs} = \frac{\omega_{T0} + \omega_{TI0}}{\omega_{T0}\,\omega_{TI0}(1 - \alpha_0\,\alpha_{I0})}. \tag{27}$$

Definicje czasów narastania i opadania objęte są w niektórych krajach normami lub zaleceniami (np. w USA przyjęto definiować czas narastania jako czas potrzebny na to, aby prąd kolektora wzrósł od 0,1 do 0,9 swojej wartości końcowej). Ze względów pomiarowych, jak i ze względu na łatwość wykorzystania do projektowania układów impulsowych, wydaje się słuszne korzystanie z zależności na całkowity czas załączenia i wyłączenia.

3. POMIAR PARAMETRÓW ŁADUNKOWYCH

W omówionych powyżej zależnościach występuje pięć podstawowych parametrów ładunkowych, a mianowicie: T_c — stała czasowa kolektora dla obszaru czynnego, Q_v — ładunek pojemności złącza kolektora-baza, β — współczynnik wzmocnienia prądowego dla dużych sygnałów, β_s — współczynnik wzmocnienia prądowego dla obszaru nasycenia, T_{bs} — stała czasowa bazy w obszarze nasycenia.



Rys. 4. Układ do pomiarów parametrów ładunkowych tranzystorów

Parametry Q_v, T_c, T_{bs} i β_s mogą być w przybliżeniu określone na podstawie pomiarów parametrów małych sygnałów. W praktyce jednak potrzeba jest zbyt dużo pomiarów koniecznych do określenia T_{bs} (i β_s), dlatego też stosowane są inne metody pomiaru.

Stała czasu T_c może być określona z pomiarów częstotliwości granicznej, zakładając, że współczynnik wzmocnienia prądowego maleje przy wzroście częstotliwości o 6 dB/oktawę. Tranzystory, których współczynniki wzmocnienia prądowego dla dużych częstotliwości nie spełniają tego warunku, wykazują małą zgodność między T_c i $1/\alpha_0\omega_T$ i dla takich tranzystorów pomiar metodą małych sygnałów jest niewłaściwy. Wyrażenie (9) jest dostatecznie dokładne, aby określić Q_v na podstawie pojemności złącza kolektora, chociaż współczynnik M jest trudny do oszacowania. Zatem do pomiaru T_{bs} i β_s, jak również do uzyskania dokładniejszych wyników T_c i Q_v, trzeba stosować inne metody pomiarowe. Jedna z nich zostanie opisana poniżej.

Rysunek 4 przedstawia układ do pomiaru wartości T_c i Q_v oraz β_s i T_{bs}. Obwód wejściowy tranzystora, którego parametry mają być po-

mierzone, jest zasilany z generatora napięć prostokątnych poprzez równoległe połączenie oporności R_{b1} i pojemności C_{b1}. Przy wartości $C_{b1} = 0$ oporność R_{b1} dobiera się tak, aby podczas impulsu tranzystor znalazł się na granicy nasycenia (tzn., aby $U_{cb} = 0$). Napięcie zasilające obwód kolektora dobiera się tak, aby było równe napięciu generatora, wtedy dla $U_{cb} = 0$

$$U_{be} = U_g - I_{b1} R_{b1} = U_{ce} = U_{cc} - \beta I_{b1} R_{b1}. \tag{28}$$

Stąd zależność $\beta = R_{b1}/R_L$ pozwala określić wartość β w oparciu o znajomość oporności R_L i R_{b1}. Następnie wartość C_{b1} dobiera się tak, aby na oscylografie uzyskać przebieg fali prostokątnej. Odpowiada to spełnieniu warunku:

$$U_g C_{b1} = Q_{zal} = T_{c0} I_c + Q_v \tag{29}$$

przy założeniu, że $I_{c0} = 0$ oraz że w stanie nasycenia $U_{ce} = 0$. Jeżeli następnie oporność obciążenia zmniejszymy dwukrotnie (rys. 4b) i oporność R_{b2} dopasujemy tak, aby tranzystor znalazł się na granicy nasycenia oraz C_{b2} dobierzemy tak, aby na wyjściu ponownie uzyskać przebieg fal prostokątnych, wtedy będzie spełniona zależność:

$$U_g (C_{b1} + C_{b2}) = Q_{zal} = 2 I_c T_{c0} + Q_v. \tag{30}$$

Ponieważ

$$I_c R_L = U_{cc} = U_g,$$

zatem

$$T_{c0} = R_L C_{b2} \tag{31}$$

oraz

$$Q_v = U_g (C_{b1} - C_{b2}). \tag{32}$$

W układzie tym można również pomierzyć parametry obszaru nasycenia. W tym celu oporność obciążenia zmieniamy z powrotem do wartości R_1. Prąd bazy przechodzący przez R_{b2} będzie teraz prądem nasycenia I_{bs} (ponieważ R_{b1} był dobrany tak, aby tranzystor znajdował się na granicy nasycenia). Jeżeli pozostawiając nie zmienione C_{b2} dobierzemy R_{b2} ponownie tak, aby uzyskać na oscylografie impulsy prostokątne, wtedy

$$Q_{bs} = I_{bs} R_{b2} C_{b2}, \tag{33}$$

stąd

$$T_{bs} = Q_{bs}/I_{bs} = R_{b2} C_{b2}. \tag{34}$$

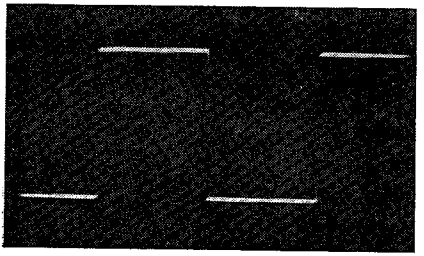
A ponieważ dla poprzednich warunków $C_{b2} = T_{c0}/R_1$,
zatem

$$C_{b2} = T_{c0}/R_1 = T_{bs}/R_{b2}, \tag{35}$$

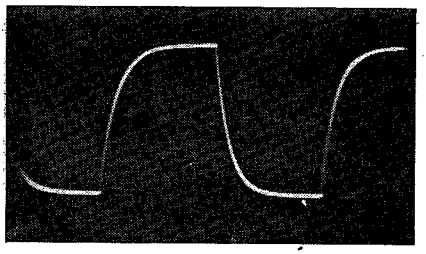
stąd

$$R_{b2} = \beta_s R_L. \tag{36}$$

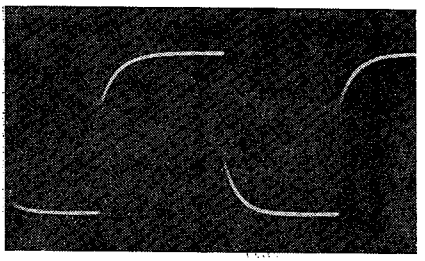
Przyjmując w powyższych pomiarach $R_L = 1$ kom otrzymamy wartość β odczytując wartość R_{b1} w kom, stałą T_c w µsek — odczytując C_{b2} w pF oraz stałą β_s — odczytując R_{b2} w kom. Ładunek Q_v otrzymamy z (32).



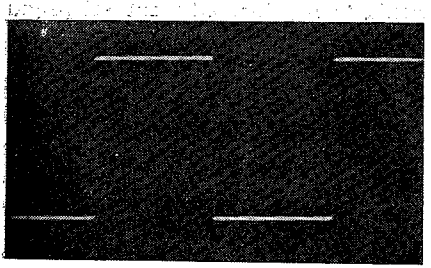
U_g



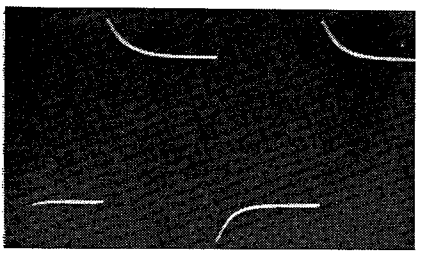
$C_{b1} = 0$



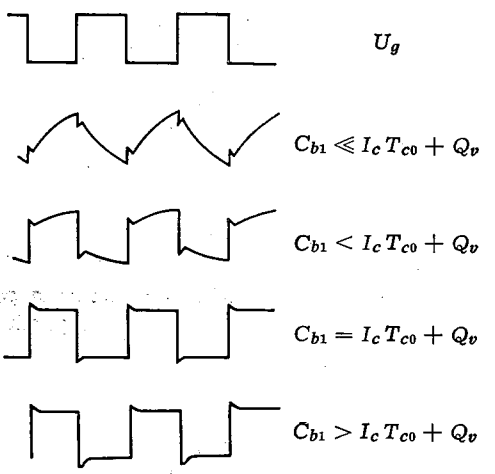
$C_{b1} < I_c T_{c0} + Q_v$



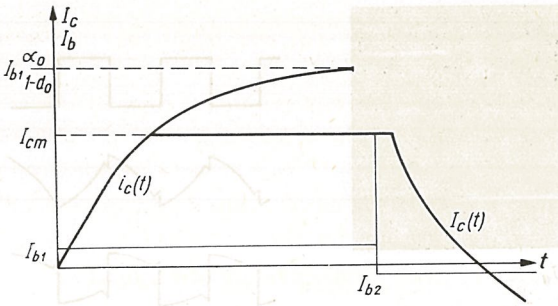
$C_{b1} = I_c T_{c0} + Q_v$



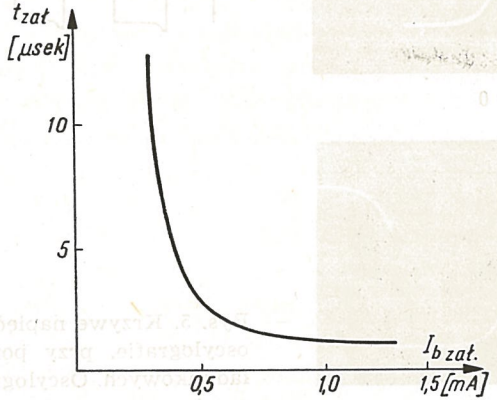
$C_{b1} > I_c T_{c0} + Q_v$



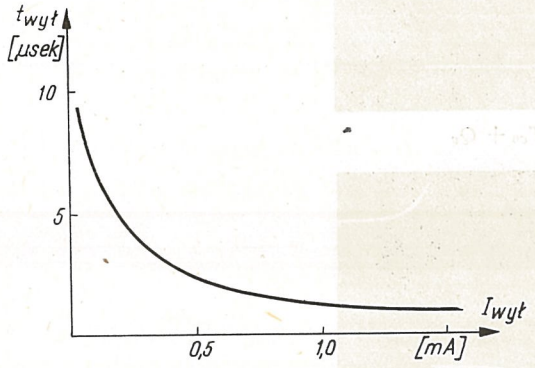
Rys. 5. Krzywe napięć obserwowanych na oscylografie, przy pomiarze parametrów ładunkowych. Oscylogramy krzywych przy pomiarze parametrów ładunkowych tranzystora OC72, $f = 3{,}5$ kHz



Rys. 6. Przebieg narastania i opadania prądu kolektora po przyłożeniu skokowo prądu bazy



Rys. 7. Krzywa obrazująca zależność czasu załączenia tranzystora od wartości prądu bazy



Rys. 8. Krzywa obrazująca zależność czasu wyłączenia tranzystora od wartości prądu bazy

Na rysunku 5 pokazano krzywe przebiegu fal prostokątnych przy pomiarach parametrów ładunkowych (przy czym krzywe zamieszczone po prawej stronie odnoszą się do tranzystorów, dla których spadek wzmocnienia odbiega od słusznego dla większości tranzystorów spadku 6 dB/oktawę).

W celu uzyskania dokładniejszych pomiarów należy uwzględnić to, że w rzeczywistości prąd zerowy kolektora posiada skończoną wartość, jak również U_{ce} na granicy nasycenia nie jest równe zeru.

4. WNIOSKI KOŃCOWE

Z analizy rozkładu ładunku w obszarze bazy wynika zasadniczy wniosek, że prąd wyjściowy tranzystora jest głównie uzależniony od ładunku znajdującego się w obszarze bazy. Prąd sterujący musi spowodować powstanie w bazie odpowiedniego ładunku, a następnie podtrzymywać wymagany prąd wyjściowy. Zatem rozpatrywanie własności tranzystora w oparciu o parametry wiążące ze sobą prądy zaciskowe tranzystora z ładunkiem bazy jest słuszne zarówno dla stanów przejściowych, jak i ustalonych, w przeciwieństwie do określenia własności tranzystora w oparciu o zależności prądowe, słuszne jedynie dla stanów ustalonych (po ustaleniu się odpowiedniego ładunku bazy).

Parametry ładunkowe definiowane jako stosunki prądów zaciskowych do ładunku bazy zmieniają się niewiele z warunkami pracy, w przeciwieństwie do parametrów małych sygnałów, słusznych tylko dla określonego punktu pracy. Z tego względu do pomiaru parametrów ładunkowych mogą być stosowane różne metody, łącznie z wykorzystaniem techniki małych sygnałów, nie powodując znacznych błędów. Parametry ładunkowe stanowią pewne wartości średnie dla obszaru charakterystyk i dlatego określanie przy ich pomocy właściwości impulsowych tranzystora jest bardziej właściwe niż przy pomocy parametrów małych sygnałów. Czasy załączenia i wyłączenia wyrażone przy pomocy parametrów ładunkowych mogą być znacznie łatwiej i szerzej stosowane przy projektowaniu układów impulsowych niż wyrażenia na czas narastania, opadania i magazynowania.

Politechnika Łódzka
Katedra Elektroniki Przemysłowej

WYKAZ LITERATURY

1. Beaufoy R., Sparkes J. J.: The Junction Transistor as a Charge-Controlled Device. — A.T.E. Journal, 1957, 13, nr 4, str. 310—327.
2. Beaufoy R.: Transistor Switching-Circuit Design Using the Charge-Control Parameters. — Proceedings of IEE, 1959, 106, cz. B, str. 1085—1091.
3. Sparkes J. J.: Measurement of Transistor Transient Switching Parameters. — Proc. of IEE, 1959, 106, cz. B, str. 562—567.

4. Kruithof A.: Transient Response of Junction Transistors and its Graphical Representation. — Proc. of IEE, 1959, 106, cz. B, str. 1092—1101.
5. Moll J. L.: Large-Signal Transient Response of Junction Transistors. — Proc. IRE, 1954, XII, str. 1773—1784.
6. Moll J. L., Ebers J. J.: Large-Signal Behaviour of Junction Transistors. — Proc. IRE, str. 1761—1772.
7. Sandler Saratow: Innercjonnyje swojstwa płaskostnych tridow pri rabotie w kluczowom reżimie na aktiwnuju nagruzku. — Wiestnik elektropromyszlennosti, nr 7, 1961.
8. Sparkes J. J.: A Study of the Charge Control Parameters of Transistors. — Proc. of IRE, vol. 48, 1948, str. 1696—1704.
9. Baranowski J., Jankowski J.: Tranzystorowe układy impulsowe. PWT, Warszawa 1961.
10. Stolarski E.: Miernictwo tranzystorowe. PWT, Warszawa 1961.

Z. KORZEC

LARGE-SIGNAL BEHAVIOUR OF JUNCTION TRANSISTOR CONSIDERED AS A CHARGE-CONTROLLED DEVICE

Summary

The present status of the concept of charge control of large signal transistor operation is considered.

Definitions of charge control parameters: collector time constant T_c, saturation time constant T_{bs}, the "on demand current gain" β_s, the collector capitance charge Q_v and dc. current gain β — are discused.

One of the methods of measuring these parameters is described. Expressions are derived for the switching on and switching off times.

Finally the expressions for switching times based on the charge control parameters are compared with these based on small-signal parameters.

Z. KORZEC

ARBEIT DER SCHICHTENTRANSISTOREN BEI GROSSEN SIGNALEN, IN HINSICHT AUF DIE LADUNGS THEORIE

Zusammenfassung

In diesem Aufsatz wurden die Fragen der Übergangszustände bei Schichtentransistoren, die bei grossen Signalen in der gemeinsamen Emitterschaltung bei aktiver Belastung arbeiten, besprochen.

Die Richtigkeit der Behandlung dieser Fragen, gestützt auf die Ladungsparameter, welche die Mittelwerte der grossen im Gegensatz zu den kleinen Signalen, die für den gegebenen Arbeitspunkt bestimmt sind darstellen, wird daselbst nachgewiesen. Es wurde auch eine einfache Messmethode für Ladungsparameter des aktiven Feldes sowie des Sättigungsfeldes entwickelt.

Ausserdem wurde ein Vergleich der Abhängigkeiten, welche die Ein- und Ausschaltzeiten charakterisieren und in Anlehnung an die Ladungsparameter und die Parameter kleiner Signale ausgeführt wurden, angegeben.

З. КОЖЕЦ

РАБОТА ПЛОСКОСТНЫХ ПОЛУПРОВОДНИКОВЫХ ТРИОДОВ ПРИ ЗНАЧИТЕЛЬНЫХ СИГНАЛАХ В СВЕТЕ ЗАРЯДНОЙ ТЕОРИИ

Резюме

Полупроводниковый триод рассматривается как устройство, управляемое зарядом базовой области. Исследуется основное параметры как: постоянная времени коллектора T_c, постоянная времени в области насыщения T_{bs}, коэффициент усиления тока в области насыщения β_s, заряд ёмкости коллекторного перехода Q_V и коэффициент усиления постоянного тока β.

Представлен пример метода измерения этих параметров. Выводятся формулы для определения времени включения и запирания полупроводникового триода. Полученные формулы сравниваются с формулами полученными на основании параметров малого сигнала.