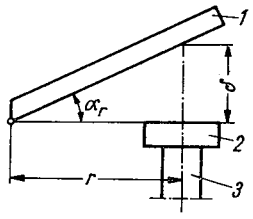
w którym ϕ_δ oznacza strumień magnetyczny przepływający pomiędzy nabiegunnikiem a zworą przez obszar rozpatrywanej umownej szczeliny głównej. Permeancja głównej szczeliny powietrznej jest dla dowolnego położenia otwarcia zwory ($\alpha_r \neq 0$, rys. 2) sumą poszczególnych permeancji cząstkowych dla określonych kątów.



Rys. 2. Uproszczony rysunek fragmentu przekaźnika
1 — zwora, 2 — nabiegunnik, 3 — rdzeń magnetyczny

Pochodną permeancji względem szczeliny δ określono za pomocą wzoru

$$\frac{d\Lambda_\delta}{d\delta} = \frac{\cos^3\alpha_r}{r} \sum_{i=1}^{n} \frac{d\Lambda_{\delta i}}{d\alpha_r}, \tag{18}$$

w którym:

$r = \dfrac{\delta}{\mathrm{tg}\alpha_r}$,

$i = 1, 2, \ldots, n$ liczba permeancji cząstkowych.

Podstawiając we wzorze (17)

$$\phi_\delta = \mu_o s_\delta \frac{(Iz)_r}{\delta}, \tag{19}$$

przy czym:

$(Iz)_r$ — amperozwoje rozruchowe,

s_δ — powierzchnia szczeliny,

$\mu_o = 4\pi \cdot 10^{-7}$ H/m,

oraz czyniąc pewne przekształcenia, można określić warunek rozruchu przekaźnika

$$(Iz)_r \geqslant \frac{\Lambda_\delta \delta}{\mu_o s_\delta} \sqrt{\frac{2Fr}{\left|\dfrac{d\Lambda_\delta}{d\delta}\right|}}. \tag{20}$$

5.2. Warunek odpadania

Warunek odpadania można określić z nierówności (17) odpowiednio ją przekształcając

$$(Iz)_o \leqslant \frac{\Lambda_\delta \Delta}{\mu_o s_\delta} \sqrt{\frac{2F_o}{\left|\dfrac{d\Lambda_\delta}{d\Delta}\right|}}, \tag{21}$$

przy czym:

$(Iz)_o$ — amperozwoje odpadania,

F_o — siła przy której następuje powrót zwory do stanu początkowego,

Δ — szczelina powietrzna przy przyciągniętej zworze.

5.3. Warunek poboru mocy

Moc P_n wydzielana w uzwojeniu przy znamionowym napięciu zasilającym U_n musi spełniać warunek:

$$P_n \leqslant \frac{U_n^2}{R}. \tag{22}$$

Podstawiając w tym wyrażeniu

$$R = \frac{l_{śr} z}{s_{cu} \gamma} = r_{śr} z,$$

uzyskuje się:

$$P_n \leqslant \frac{U_n^2 \cdot \gamma \cdot s_{cu}}{z \cdot l_{śr}}, \tag{23}$$

przy czym:

γ — przewodność miedzi,

s_{cu} — przekrój drutu nawojowego,

$l_{śr}$ — średnia długość zwoju.

Z wyrażenia (23) można wyznaczyć warunek znamionowych amperozwojów

$$(Iz)_n \leqslant \sqrt{\frac{z P_n \gamma s_{cu}}{l_{śr}}}. \tag{24}$$

5.4. Warunek przyrostu temperatury uzwojenia

Przy określaniu warunku przyrostu temperatury uzwojenia przyjęto moc wydzielającą się w uzwojeniu przy pU_n, przy czym p — krotność napięcia znamionowego określona w wymaganiach.

W warunkach ustalonych przyrostu temperatury powinien być spełniony warunek:

$$R \geqslant p^2 \frac{U_n^2}{\Delta\vartheta s \alpha_c}, \tag{25}$$

przy czym:

$\Delta\vartheta$ — dopuszczalny przyrost temperatury,

s — powierzchnia odprowadzająca ciepło,

α_c — współczynnik odprowadzania ciepła

lub

$$P_n \leqslant \frac{\Delta\vartheta}{\alpha_c s} \frac{1}{p^2}. \tag{25a}$$

5.5. Warunek maksymalnych amperozwojów rozruchowych

Przekaźnik powinien działać dostatecznie pewnie i szybko w całym zakresie zmian napięcia zasilającego. W celu spełnienia tego wymagania żąda się aby:

$$(Iz)_z \leqslant p_r(Iz)_n, \tag{26}$$

przy czym:

$(Iz)_n = z\dfrac{U_n}{R}$ — amperozwoje znamionowe

p_r — współczynnik.

Czyniąc pewne przekształcenia wyrażenie (26) można przedstawić w następującej postaci:

$$I_r \leqslant p_r \frac{U_n}{R}. \tag{26a}$$

5.6. Warunek minimalnych amperozwojów odpadania

Przekaźnik po odłączeniu spod napięcia powinien wrócić do stanu początkowego. Jest to spełnione, gdy:

$$(Iz)_o \geqslant p_o(Iz)_n. \tag{27}$$

Podstawiając za $(Iz)_o = I_o z$ i czyniąc pewne przekształcenia, otrzymuje się

$$I_o \geqslant p_o \frac{U_n}{R}, \tag{27a}$$

przy czym p_o — współczynnik.

Warunki (16)...(21) są warunkami określającymi działanie przekaźnika, pozostałe (22)...(27) warunkami ograniczającymi.

5.7. Wyznaczenie obszarów prawidłowego działania i przydatności konstrukcyjnej przekaźnika

Przedstawione warunki nie wyczerpują oczywiście wszystkich możliwości, ponieważ w zależności od wymagań stawianych przekaźnikowi można sformułować szereg dalszych warunków prawidłowego działania.

Mimo tego w podanych warunkach występuje 25 parametrów pracy, które wyznaczają 25-wymiarowy obszar prawidłowego działania. Przed jego określeniem wprowadzono zmienne bezwymiarowe wynikające z podzielenia parametrów pracy przez odpowiadające im parametry jednostkowe:

$$F_r^* = F_r/F, \quad (Iz)_r^* = (Iz)_r/Iz, \quad s_\delta^* = s_\delta/s, \quad \gamma^* = \gamma/\gamma,$$

$$(\mu_o^* = \mu_o)\mu, \quad \delta^* = \delta/\delta, \quad \Lambda^* = \Lambda/\Lambda, \quad \frac{d\Lambda_\delta^*}{d\delta} = \frac{d\Lambda_\delta}{d\delta}\Big/\frac{d\Lambda_\delta}{d\delta},$$

$$(Iz)_o^* = (Iz)_o/Iz, \quad F_o^* = F_o/F, \quad \Lambda_\delta^* = \Lambda_\delta/\Lambda,$$

$$U_n^* = U_n/U, \quad R^* = R/R, \quad z^* = z/z, \quad P_n^* = P_n/P,$$

$$s^* = s/s, \quad \Delta\vartheta^* = \Delta\vartheta/\Delta\vartheta, \quad s_{cu}^* = s_{cu}/s,$$

$$\alpha_c^* = \alpha_c/\alpha_c, \quad l_{śr}^* = l_{śr}/l.$$

Po podstawieniu tak określonych zmiennych do nierówności (20)...(27) i po obustronnym zlogarytmowaniu tych nierówności otrzymano następujące warunki prawidłowego działania:

$$\begin{aligned} C_1 - C_2 - C_3 + C_4 + C_5 - C_6 - C_7 &\geqslant 0, \\ -C_2 + C_4 + C_5 + C_8 - C_9 - C_{10} - C_{11} &\leqslant 0, \\ +C_{12} - C_{13} + C_{14} &\leqslant 0, \\ C_{12} + C_{15} - C_{16} - C_{17} - C_{18} - C_{19} &\leqslant 0, \\ +C_{12} - C_{20} + C_{21} + C_{22} + C_{23} &\leqslant 0, \\ C_1 - C_{15} - C_{24} &\leqslant 0, \\ +C_8 - C_{15} - C_{25} &\geqslant 0, \end{aligned} \tag{28}$$

w których

$$C_1 = \lg(Iz)_r^*; \quad C_2 = \lg \Lambda_\delta^*; \quad C_3 = \lg \delta^*; \quad C_4 = \lg \mu_o^*; \quad C_5 = \lg s_\delta^*;$$

$$C_6 = \frac{1}{2}\lg F_r^*; \quad C_7 = \frac{1}{2}\lg \left| \frac{d\Lambda_\delta}{d\delta} \right|^*; \quad C_8 = \lg(Iz)_o^*; \quad C_9 = \lg \Delta^*;$$

$$C_{10} = \frac{1}{2}\lg F_r^*; \quad C_{11} = \frac{1}{2}\lg \left| \frac{d\Lambda_\delta}{d\Delta} \right|^*; \quad C_{12} = \lg P_n^*; \quad C_{13} = 2\lg U_n^*;$$

$$C_{14} = \lg R^*; \quad C_{15} = 2\lg(Iz)_n^*; \quad C_{16} = \lg z^*; \quad C_{17} = \lg \gamma^*;$$

$$C_{18} = \lg s_{cu}^*; \quad C_{19} = \lg l_{śr}^*; \quad C_{20} = \lg \Delta\vartheta^*; \quad C_{21} = \lg \alpha_c^*;$$

$$C_{22} = \lg s^*; \quad C_{23} = 2\lg p; \quad C_{24} = \lg p_r; \quad C_{25} = \lg p_0.$$

Z przytoczonych zależności widać, że warunki prawidłowego działania (28) są określone zależnościami liniowymi i że do obliczenia optymalnych parametrów przekaźników można wykorzystać zasady programowania liniowego [4]. Trzeba jednak pamiętać, że otrzymane wyniki należy zdelogarytmować.

W celu przedstawienia graficznego analizowanego przykładu, konieczne jest zredukowanie 25 parametrów do 3. Wartości wielu z tych parametrów są bowiem ustalone. Można z góry założyć liczbę zestyków, stąd znane są k, n. Znany jest również materiał drutu nawojowego oraz kształt cewki — stąd znany jest współczynnik α_c.

Z odpowiednich wymagań można określić współczynniki p, p_r, p_o oraz wartości dopuszczalnego przyrostu temperatury $\Delta\vartheta$. Znana jest również przenikalność μ_o oraz wartość napięcia znamionowego U_n. Obliczyć można s i R. Dla uproszczenia założymy, że znane są następujące parametry:

$$\delta, \quad \Delta, \quad \frac{d\Lambda_\delta}{d\delta}, \quad \Lambda_\delta, \quad s_\delta, \quad \frac{d\Lambda_\delta}{d\delta},$$

P_n, U_n, $(Iz)_n$, z, γ, s_{cu}, $l_{śr}$ oraz że $F_r \wedge F_o = \text{const}$.

W ten sposób warunki rozruchu oraz odpadania uproszczą się odpowiednio

$$C_1 - C_6 - C_o \geqslant 0, \tag{29}$$

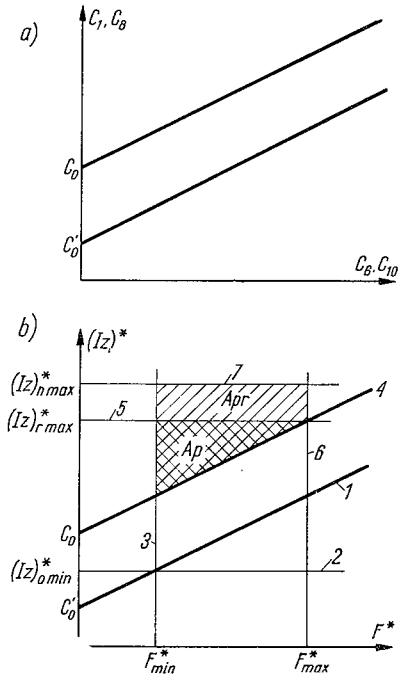
$$C_0 - C_{10} - C_o' \leqslant 0, \tag{30}$$

przy czym C_o i C_o' stałe.

Obszary określone za pomocą nierówności (29) i (30) przedstawiono na jednym wykresie (rys. 3), jako że C_1 i C_8 są zależnościami amperozwojów rozruchowych i odpadania, a C_6 i C_{10} funkcjami siły rozruchowej i odpadania.

Z rysunku 3a widać, że obszar amperozwojów rozruchowych rozciąga się powyżej prostej wyrażającej warunek (29) zamieniony na równość, natomiast obszar amperozwojów odpadania znajduje się poniżej prostej wyrażającej warunek (30) zamieniony również na równość.

Przedstawione obszary są ograniczone (rys. 3b): od góry prostą określoną z warunku (26) oraz od dołu również prostą zgodnie z warunkiem (27). Przecięcie prostych 1 i 2 wy-



Rys. 3. Obszary działania przekaźnika

znacza punkt, którego współrzędna stanowi logarytm minimalnej wartości siły odpadania zwory przekaźnika. Prosta pionowa 3 poprowadzona przez ten punkt ogranicza przedstawione obszary lewostronnie. Przecięcie prostych 4 i 5 wyznacza punkt, którego współrzędna stanowi w tym przypadku logarytm maksymalnej wartości siły docisku zestyków przekaźnika. Prosta pionowa 6 poprowadzona przez ten punkt ogranicza przedstawione obszary prawostronnie.

Zgodnie z definicjami (12) i (13) obszarem prawidłowego działania przekaźnika zabezpieczeniowego jest obszar ograniczony prostymi 3, 4, 7, 6, natomiast obszarem przydatności konstrukcyjnej przekaźnika może być obszar ograniczony prostymi 3, 4, 5, w zależności od przyjętego kryterium przydatności.

Położenie punktu pracy ma wpływ na niezawodność działania przekaźnika, ponieważ zmiana parametrów przekaźnika może przesunąć punkt pracy poza obszar prawidłowego

działania lub przydatności konstrukcyjnej przekaźnika, szczególnie w przypadku gdy punkt ten leży na krawędzi tego obszaru.

W celu zyskania rozwiązania optymalnego, obszar prawidłowego działania należy ograniczyć niezawodnościowo. Analiza tego zagadnienia stanowi jednak odrębne zagadnienie i przekracza ramy tego artykułu.

6. ZAKOŃCZENIE

W wyniku przyjęcia w przedstawionej analizie obszarów prawidłowego działania i przydatności konstrukcyjnej uzyskano możliwość wyboru optymalnego punktu pracy przekaźnika.

Szczególnie korzystna jest możliwość ograniczenia niezawodnościowego tych obszarów i zwiększenie pewności działania przekaźnika. Analiza przedstawionego modelu matematycznego pozwala na wykorzystanie metod elektronicznej techniki obliczeniowej, których istota polega na tym, żeby ze zbioru możliwych warunków pracy przekaźnika wybrać według ustalonego kryterium wariant najlepszy, optymalny.

LITERATURA CYTOWANA W TEKŚCIE

1. W. Bobrowski, Niektóre zagadnienia związane z oceną prawidłowości przekaźników zabezpieczających, Arch. Elektrotechn., 1971, nr 2.
2. S. Budkowski, O pewnych metodach optymalizacji układów logicznych, Arch. Autom. i Telemech., 1969, t. XIV, z. 4.
3. G. H. Goldstick, D. G. Machie, Design of computer cicuits using linear programming techniques, IRE Trans. on Electr. Comp., Vol. EC-9, 1962.
4. I. L. Kalichman, Algebra liniowa i programowanie, PWN, Warszawa 1971.
5. W. Kostecki, Analityczna metoda wyznaczania permeancji głównej szczeliny powietrznej elektromagnesu klapkowego z nabiegunnikiem, Arch. Elektrotechn., 1972, nr 3.
6. A. W. Kurnosow, K woprosu ob optimalnoj geometrii cilindriczeskowo elektromagnita postojannowo toka, Elektriczestwo, 1971, nr 7.
7. N. E. Łysow, A. W. Kurnosow, Ob optimalnych geometriczeskich cootnoszenjach osnobnych razmierow elektromagnitow nastojannowo toka, Elektriczestwo, 1965, nr 8.
8. R. Ramaratham. B. G. Desai, Optimization of polyphase motor design, IEEE Transactions on Power Apparatus aod System, 1971, nr 2.
9. A. Sielicki, Zagadnienia optymalnej syntezy teoretycznej podstawowych układów logicznych, Zeszyty Naukowe Politechniki Wrocławskiej, 1967, nr 178.
10. W. N. Szoffa, Projektnyj mietod rasczeta eliektromagnitow prostojannowo toka kłapannowo tipa, Elektrotiechnika, 1968, nr 5.

J. ŁYSKANOWSKI

SOME PROBLEMS OF THE TECHNICAL ANALYSIS OF THE CONSTRUCTION OF RELAYS

Summary

The paper discusses the designing of the construction of relays using the linear programming method. There has been presented an analytical method of determination of the region of correct performance and the usebility of relays. This method can be used in the estimation of the quality and reliability of relays.

J. ŁYSKANOWSKI

EINIGE PROBLEME TECHNISCHER ANALYSE DER RELAISKONSTRUKTION

Zusammenfassung

Im Artikel wurde eine Analyse der Relaiskonstruktion dargestellt, indem Elemente linearen Programmierens genutzt wurden. Im weiteren wurden Bereiche ordnungsmässigen Funktionierens sowie der Konstruktionsanwendbarkeit von Relais ermittelt. Die in diesem Artikel vorgeführte Methode lässt eine objektive Einschätzung der Güte und der Zuverlässigkeit der Relais zu.

Я. ЛЫСКАНОВСКИ

НЕКОТОРЫЕ ВОПРОСЫ ТЕХНИЧЕСКОГО АНАЛИЗА КОНСТРУКЦИИ РЕЛЕ

Резюме

Статья содержит анализ конструкции реле при использовании элементов линейного программирования. Определены области правильного действия и области конструкционной пригодности реле. Представленный метод может быть использован для оценки качества и надежности реле.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 1, ss. 209-217

621.316.925

Niezawodność układów zabezpieczeń przekaźnikowych zwielokrotnionych

WITOLD BOBROWSKI (BRWINÓW)
Instytut Elektrotechniki, Warszawa

Otrzymano 10.3.1973

W artykule omówiono niektóre zagadnienia związane z obliczaniem niezawodności układów zabezpieczeń przekaźnikowych zwielokrotnionych systemem m/n.

Podano zależności analityczne umożliwiające obliczenie wartości prawdopodobieństw zadziałań brakujących, zbędnych oraz prawidłowych układu zwielokrotnionego. Wyznaczono obszar jednoczesnego zmniejszania się prawdopodobieństw zadziałań brakujących i zbędnych układu zabezpieczeń przekaźnikowych zwielokrotnionych systemem 2/3.

1. WSTĘP

W artykule omówiono niektóre zagadnienia związane z obliczaniem niezawodności układów zabezpieczeń przekaźnikowych zwielokrotnionych. W szczególności rozpatrzono układy zwielokrotnione (rezerwowane) systemem m/n, w którym impuls na wyjściu układu składającego się z n zabezpieczeń pojawia się wówczas, gdy co najmniej m zabezpieczeń poda impuls wyjściowy. Zwielokrotnianie zabezpieczeń przekaźnikowych ma na celu zwiększenie niezawodności działania. Jednak ze względu na dwoisty charakter nieprawidłowych zadziałań przekaźnikowych nie wszystkie układy zwielokrotniane spełniają to zadanie. Wybór odpowiedniego układu zwielokrotniania zabezpieczeń powinien więc być poprzedzony odpowiednią analizą niezawodnościową.

2. ZASTOSOWANIE TEORII NIEZAWODNOŚCI DO OCENY ZABEZPIECZEŃ PRZEKAŹNIKOWYCH

Zastosowanie elementów rachunku prawdopodobieństwa do oceny struktur niezawodnościowych tworzonych przy użyciu zestyków przekaźników omówiono w [1].

Niektóre zagadnienia związane z zastosowaniem teorii niezawodności do oceny zabezpieczeń przekaźnikowych rozpatrzono w [2]. W szczególności zwrócono uwagę na istnienie dwóch rodzajów nieprawidłowych zadziałań zabezpieczeń przekaźnikowych polegających na

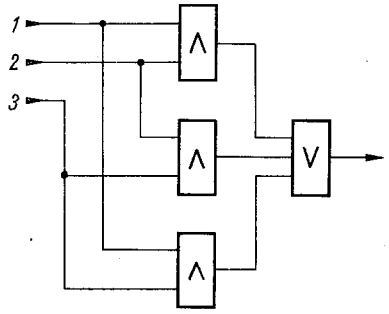
1) pojawieniu się sygnału na wyjściu, w warunkach kiedy takiego sygnału nie powinno być,

2) braku sygnału na wyjściu, w warunkach gdy taki sygnał powinien się pojawić.

Wykazano, że przy łączeniu równoległym zabezpieczeń przekaźnikowych następuje zwiększenie prawdopodobieństwa pojawienia się zbędnego sygnału na wyjściu układu, natomiast przy łączeniu szeregowym następuje zwiększenie prawdopodobieństwa braku sygnału na wyjściu. Omówiono układy zwielokrotniania szeregowo-równoległego zabezpieczeń przekaźnikowych i wykazano, że w układach tych można uzyskać jednocześnie zmniejszenie się wartości prawdopodobieństw obu rodzajów nieprawidłowych zadziałań; wymaga to jednak zastosowania układów złożonych co najmniej z czterech zabezpieczeń.

Najbardziej rozpowszechnionym systemem rezerwowania zabezpieczeń przekaźnikowych [3] są systemy 1/2, 2/2, 2/3, oraz 2/4 nazywane odpowiednio: 1 z 2, 2 z 2, 2 z 3 oraz 2 z 4.

Niektóre problemy rezerwowania zabezpieczeń przekaźnikowych systemem m/n omówiono w [6]. Analizę efektywności rezerwowania równoległego zabezpieczeń przekaźnikowych przeprowadzono w [7]. Problemy zwielokrotniania zabezpieczeń wielkich bloków i związane z nimi zagadnienia układowe rozpatrzono w [4].



Rys. 1. Przykład schematu logicznego układu zabezpieczeń przekaźnikowych zwielokrotnionych systemem m/n (2/3)

Poniżej przeprowadzono analizę niezawodnościową układów zabezpieczeń przekaźnikowych zwielokrotnionych typu 1/2, 2/2, 2/3, ..., m/n zwanych dalej układami zwielokrotnionymi systemem m/n.

Przykład schematu połączeń układu zwielokrotnionego zabezpieczeń systemem 2/3 podano na rys. 1.

W ogólnym przypadku wskaźnik m może przyjmować wartości ze zbioru liczb naturalnych w zakresie od 1 do n

$$1 \leqslant m \leqslant n; \quad n \in N. \tag{1}$$

W przypadkach szczególnych przy wartości wskaźnika

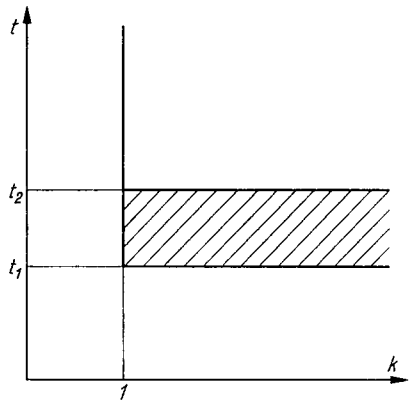
$$m = 1 \tag{2}$$

układ zwielokrotniania systemem $1/n$ jest układem równoległym, natomiast przy wartości wskaźnika

$$m = n \tag{3}$$

układ zwielokrotniania systemem n/n jest układem szeregowym.

Ocenę prawidłowości zadziałania danego zabezpieczenia przekaźnikowego można przeprowadzić przy zastosowaniu kryterium czasowego [5]. Zabezpieczenie przekaźnikowe powinno zadziałać na ogół w określonym czasie, w przypadku gdy przekroczona zostanie wartość rozruchowa wielkości, na którą zabezpieczenie reaguje, natomiast nie powinno działać w pozostałych przypadkach. Przykład wymaganego przebiegu charakterystyki czasowo-prądowej zabezpieczenia nadprądowo-zwłocznego podano na rys. 2.



Rys. 2. Wymagany przebieg charakterystyki zabezpieczenia nadmiarowego zwłocznego
k — krotność wartości rozruchowej, t_1, t_2 — dopuszczalne wartości czasów zadziałania

Przebieg tej charakterystyki można zapisać analitycznie w następującej postaci

$$t_d = \begin{cases} t_d = \infty ; & k < 1, \\ t_1 \leqslant t_d \leqslant t_2 ; & k \geqslant 1, \end{cases} \tag{4}$$

gdzie:

t_d — wymagany czas zadziałania zabezpieczenia,

k — krotność wartości rozruchowej.

Zadziałaniem prawidłowym zabezpieczenia nazywać będziemy zadziałanie, które nastąpiło w wymaganym czasie

$$t_1 \leqslant t_d \leqslant t_2 ; \quad k \geqslant 1 . \tag{5}$$

Należy w tym miejscu zwrócić uwagę, że zgodnie z definicją nie uznajemy za działanie prawidłowe faktu niepobudzenia się np. zabezpieczenia różnicowego przy zwarciu zewnętrznym $k < 0$ (ujemne). Zgodnie z sensem fizycznym „zadziałania" wówczas nie było.

Zadziałaniem zbędnym nazywać będziemy zadziałanie, które nastąpiło w czasie mniejszym od wymaganego

$$\begin{cases} t_d < \infty ; & k < 1, \\ t_d < t_1 ; & k \geqslant 1. \end{cases} \tag{6}$$

Zadziałaniami zbędnymi będą więc wszystkie zadziałania, które nastąpiły przy $k < 1$, a więc również zadziałania przy zwarciach na zewnątrz strefy chronionej przy zabezpieczeniach kierunkowych, różnicowych itp. gdy $k < 0$.

Zadziałaniem brakującym nazywać będziemy zadziałanie, które nastąpiło w czasie większym od wymaganego

$$t_d > t_2; \quad k \geqslant 1. \tag{7}$$

Zadziałaniem brakującym nazywać będziemy więc także „zadziałanie", które w ogóle nie nastąpi, tzn. nastąpi w czasie nieskończenie długim.

Oznaczmy przez p prawdopodobieństwo zdarzenia polegającego na zadziałaniu prawidłowym danego zabezpieczenia w rozpatrywanych warunkach eksploatacyjnych.

Oznaczmy przez z prawdopodobieństwo zdarzenia polegającego na zadziałaniu zbędnym danego zabezpieczenia w rozpatrywanych warunkach eksploatacyjnych.

Oznaczmy przez b prawdopodobieństwo zdarzenia polegającego na zadziałaniu brakującym danego zabezpieczenia w rozpatrywanych warunkach eksploatacyjnych.

Zdarzenia polegające na zadziałaniu prawidłowym, zbędnym lub brakującym danego zabezpieczenia są zdarzeniami wzajemnie wykluczającymi się, w związku z czym zachodzi związek

$$p+b+z = 1. \tag{8}$$

Wartości prawdopodobieństw p, z oraz b w warunkach eksploatacyjnych mogą być wyznaczane na podstawie odpowiednich obserwacji zadziałań zabezpieczeń.

Oprócz analizy niezawodnościowej wybór odpowiedniego układu powinien być poprzedzony odpowiednią analizą ekonomiczną, ze względu na niejednakowe wartości współczynników wagowych odpowiadających ekonomicznym skutkom działań brakujących i zbędnych. Ponieważ wartości tych współczynników zależą od warunków konkretnego układu i wzajemne ich proporcje mogą kształtować się w różny sposób, w niniejszym artykule nie rozważano szczegółowo zagadnień ekonomicznych, a zgodnie z tytułem ograniczono się jedynie do analizy niezawodnościowej.

3. ANALIZA NIEZAWODNOŚCI UKŁADÓW ZABEZPIECZEŃ PRZEKAŹNIKOWYCH ZWIELOKROTNIANYCH, SYSTEMEM m/n

Analizę zabezpieczeń przekaźnikowych zwielokrotnianych systemem m/n przeprowadzono przy następujących założeniach upraszczających.

1) Uszkodzenia występujące w poszczególnych zabezpieczeniach wchodzących w skład układu zwielokrotniania są wzajemnie niezależne. Założenie powyższe jest dopuszczalne, jeżeli zabezpieczenia wchodzące w skład układu zwielokrotniania różnią się budową i mają niezależne układy wejściowe i wyjściowe.

2) Odpowiednie wartości prawdopodobieństw zadziałań zabezpieczeń wchodzących w skład układu zwielokrotniania są sobie równe

$$\begin{aligned} z_1 &= z_2 = \ldots = z_n, \\ b_1 &= b_2 = \ldots = b_n, \\ p_1 &= p_2 = \ldots = p_n. \end{aligned} \tag{9}$$

Założenie powyższe jest dopuszczalne, jeżeli jakość zabezpieczeń rezerwowych odpowiada jakości zabezpieczenia głównego.

3) W obliczeniach nie uwzględniono niezawodności elementów i połączeń dodatkowych samego układu zwielokrotniania.

Założenie powyższe jest dopuszczalne jeżeli niezawodność elementów dodatkowych jest znacznie większa niż niezawodność zabezpieczeń wchodzących w skład układu zwielokrotniania.

Oznaczmy przez $p^{(m/n)}$, $z^{(m/n)}$ oraz $b^{(m/n)}$ prawdopodobieństwa zadziałań prawidłowych, zbędnych i brakujących układu zabezpieczeń przekaźnikowych zwielokrotnionych systemem m/n.

Zadziałanie zbędne układu zabezpieczeń przekaźnikowych zwielokrotnionych systemem m/n może nastąpić wówczas, gdy co najmniej m zabezpieczeń zadziała zbędnie.

Prawdopodobieństwo zdarzenia polegającego na tym, że co najmniej m zabezpieczeń z ogólnej liczby n zadziała zbędnie możemy wyrazić za pomocą następujących zależności

$$z^{(m/n)} = \sum_{r=m}^{n} \binom{n}{r} z^r (1-z)^{n-r}, \tag{10}$$

gdzie r — wskaźnik sumowania.

Zadziałanie brakujące układu zabezpieczeń przekaźnikowych zwielokrotnionych systemem m/n może nastąpić wówczas, gdy co najmniej $n-m+1$ zabezpieczeń zadziała brakująco. Prawdopodobieństwo zdarzenia polegającego na tym, że co najmniej $n-m+1$ zabezpieczeń z ogólnej liczby n zadziała brakująco można wyrazić za pomocą następującej zależności

$$b^{(m/n)} = \sum_{r=n-m+1}^{n} \binom{n}{r} b^r (1-b)^{n-r}. \tag{11}$$

Prawdopodobieństwo zadziałania prawidłowego układu zabezpieczeń przekaźnikowych zwielokrotnionych systemem m/n możemy określić z zależności analogicznej do wzoru (8)

$$p^{(m/n)} = 1 - (z^{(m/n)} + b^{(m/n)}). \tag{12}$$

4. PRZYKŁAD ANALIZY NIEZAWODNOŚCI DZIAŁANIA UKŁADÓW TRZECH ZABEZPIECZEŃ ZWIELOKROTNIONYCH SYSTEMEM m/n

Poniżej rozpatrzono przykładowo niezawodności układów zwielokrotnionych składających się z trzech zabezpieczeń pracujących odpowiednio w systemie 1/3, 2/3 oraz 3/3. Wartości prawdopodobieństw rozpatrywanych zadziałań zbędnych, brakujących oraz prawidłowych układu zwielokrotnienia wyznaczyć można z zależności (10), (11) oraz (12). Prawdopodobieństwa zadziałań zbędnych omawianych układów zwielokrotniania wynoszą odpowiednio:

dla systemu 1/3

$$z^{(1/3)} = \sum_{r=1}^{3} \binom{3}{r} z^r (1-z)^{3-r} = z^3 - 3z^2 + 3z, \tag{13a}$$

dla systemu 2/3

$$z^{(2/3)} = \sum_{r=2}^{3} \binom{3}{r} z^r (1-z)^{3-r} = 3z^2 - 2z^3, \tag{13b}$$

dla systemu 3/3

$$z^{(3/3)} = \sum_{r=3}^{3} \binom{3}{r} z^r (1-z)^{3-r} = z^3. \tag{13c}$$

Prawdopodobieństwa zadziałań brakujących omawianych układów zwielokrotniania wynoszą odpowiednio:
dla systemu 1/3

$$b^{(1/3)} = \sum_{r=3}^{3} \binom{3}{r} b^r (1-b)^{3-r} = b^3, \tag{14a}$$

dla systemu 2/3

$$b^{(2/3)} = \sum_{r=2}^{3} \binom{3}{r} b^r (1-b)^{3-r} = 3b^2 - 2b^3, \tag{14b}$$

dla systemu 3/3

$$b^{(3/3)} = \sum_{r=1}^{3} \binom{3}{r} b^r (1-b)^{3-r} = b^3 - 3b^2 + 3b. \tag{14c}$$

Prawdopodobieństwa zadziałań prawidłowych omawianych układów wynoszą odpowiednio:
dla systemu 1/3

$$p^{(1/3)} = 1 - (z^{(1/3)} + b^{(1/3)}) = 1 - z^3 + 3z^2 - 3z - b^3, \tag{15a}$$

dla systemu 2/3

$$p^{(2/3)} = 1 - (z^{(2/3)} + b^{(2/3)}) = 1 - 3z^2 + 2z^3 - 3b^2 + 2b^3, \tag{15b}$$

dla systemu 3/3

$$p^{(3/3)} = 1 - (z^{(3/3)} + b^{(3/3)}) = 1 - z^3 - b^3 + 3b^2 - 3b. \tag{15c}$$

Z powyższych zależności wynika, że w systemie 1/3 istnieje możliwość zmniejszenia prawdopodobieństwa zadziałań brakujących, natomiast w systemie 3/3 istnieje możliwość zmniejszenia prawdopodobieństwa zadziałań zbędnych, natomiast w systemie 2/3 istnieje możliwość zmniejszenia zarówno prawdopodobieństwa zadziałań zbędnych jak i brakujących.

Wyznaczymy wartość prawdopodobieństw z oraz b, przy których w systemie 2/3 następuje jednoczesne zmniejszenie się prawdopodobieństw zadziałań zbędnych i brakujących.

Z zależności dla prawdopodobieństw zadziałań zbędnych

$$z^{(2/3)} < z \tag{16a}$$

otrzymujemy po uwzględnieniu zależności (13b)

$$2z^3-3z^2+z>0, \tag{16b}$$

stąd

$$o<z<0{,}5. \tag{16c}$$

Z zależności dla prawdopodobieństw zadziałań brakujących

$$b^{(2/3)}<b \tag{17a}$$

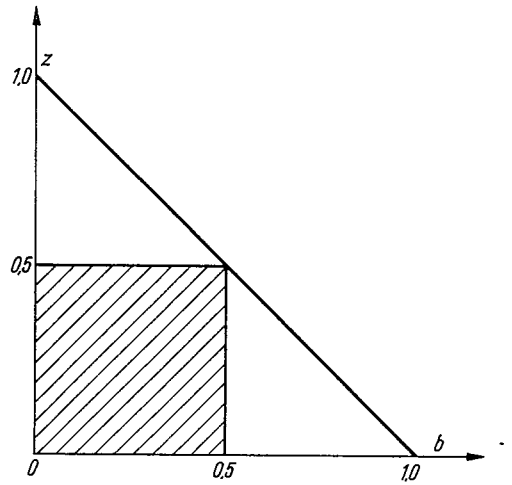
otrzymujemy po uwzględnieniu zależności (14b)

$$2b^3-3b^2+b>0, \tag{17b}$$

stąd

$$0<b<0{,}5. \tag{17c}$$

Na rys. 3 przedstawiono obszar wartości prawdopodobieństw zadziałań brakujących i zbędnych zabezpieczeń, przy których zastosowanie zwielokrotnienia systemem 2/3 powoduje jednoczesne zmniejszenie się wartości prawdopodobieństw obu rodzajów zadziałań nieprawidłowych.



Rys. 3. Obszar jednoczesnego zmniejszania się wartości prawdopodobieństw zadziałań nieprawidłowych układu zwielokrotnionego systemem 2/3

Praktycznie w energetyce nie stosuje się tak zawodnych zabezpieczeń, których prawdopodobieństwa zadziałań nieprawidłowych przekraczałyby wartość 0,5. Można więc sformułować ogólne stwierdzenie, że przy zastosowaniu systemu 2/3 następuje jednoczesne zmniejszenie się prawdopodobieństw zadziałań zbędnych i brakujących.

5. WNIOSKI

1. Zwielokrotnianie zabezpieczeń przekaźnikowych systemem m/n umożliwia zmniejszenie prawdopodobieństw zadziałań zbędnych bądź zadziałań brakujących układu, lub jednoczesne zmniejszenie wartości prawdopodobieństw obu rodzajów nieprawidłowych zadziałań.

2. Zastosowanie układów zwielokrotnionych systemem $1/n$ (połączenie równoległe) powoduje zmniejszenie się wartości prawdopodobieństwa zadziałań brakujących, ale jednoczesny wzrost wartości prawdopodobieństwa zadziałań zbędnych.

3. Zastosowanie układów zwielokrotnionych systemem n/n (połączenie szeregowe) powoduje zmniejszenie się wartości prawdopodobieństwa zadziałań zbędnych ale jednoczesny wzrost wartości prawdopodobieństwa zadziałań brakujących.

4. Dla zabezpieczeń przekaźnikowych spotykanych w praktyce przy zastosowaniu układu zwielokrotnionego systemem 2/3 następuje jednoczesne zmniejszenie się zarówno wartości prawdopodobieństw zadziałań brakujących jak i zbędnych.

5. Wybór odpowiedniego układu zwielokrotnienia zabezpieczeń przekaźnikowych systemem m/n powinien być poprzedzony analizą niezawodnościową rozpatrywanych układów.

LITERATURA CYTOWANA W TEKŚCIE

1. E. F. Moore, C. E. Shannon, Reliable Circuits Using Less Reliable Relays, J. Franklin Inst., 262, No 3 (1956), 191—208; No 4, 281—297.
2. V. L. Fabrikant, O primenenii teorii nadeznosti k ocenke ustrojstv relejnoj zaščity, Električestvo, 1965, nr 9.
3. J. Trojak, Considerations on Protection and Automation Problems for Large Turbogenerators, CIGRE, 1972, nr 34—07.
4. W. Dzierżanowski, Zwielokrotnianie zabezpieczeń wielkich bloków, Zeszyty Naukowe Politechniki Wrocławskiej, Wrocław 1973.
5. W. Bobrowski, Niektóre zagadnienia związane z oceną prawidłowości działania przekaźników zabezpieczających, Archiwum Elektrotechniki, 1971, nr 2.
6. W. Bobrowski, Problemy niezawodności zabezpieczeń przekaźnikowych rezerwowanych systemem m/n, Materiały Konf. Zast. Matem. IMPAN-PTM, Jadwisin, 16—19.X.1972.
7. W. Bobrowski, Optimum efektywności rezerwowania równoległego zabezpieczeń przekaźnikowych, Archiwum Elektrotechniki, 1973, nr 1.

W. BOBROWSKI

RELIABILITY OF RESERVATED PROTECTIVE RELAY SCHEMES

Summary

The paper discusses some problems connected with computation of reliability of m/n system reservated protective relays schemes.

Analytical dependences allowing computation of probability values of missing, needless and correct operations of reservated schemes are given. The surface of simultaneous decreasing of missing and needless operation probabilities of the 2/3 system reservated protective relays scheme is appointed.

W. BOBROWSKI

ZUVERLÄSSIGKEIT VON SICHERHEITSSYSTEMEN VERVIELFACHTER RELAIS

Zusammenfassung

Es wurden manche mit der Zuverlässigkeitberechnungen von Sicherheitssystemen vervielfachter Relais Typ m/n zusammenhängende Probleme erwogen.

Man hat analytische Abhängigkeiten angeführt, die eine Wertberechnung der Wahrscheinlichkeiten des Einsetzens fehlender, unnötiger und ordnungsmässiger vervielfachter Systeme möglich machen. Der Bereich gleichzeitiger Verminderung von Wahrscheinlichkeiten des Einsetzens fehlender und unnötiger Sicherheitssysteme vervielfachter Relais Typ 2/3 wurde ermittelt.

В. БОБРОВСКИ

НАДЕЖНОСТЬ РЕЗЕРВИРОВАННЫХ СХЕМ РЕЛЕЙНОЙ ЗАЩИТЫ

Р е з ю м е

В статье обсуждены некоторые вопросы расчета надежности схем релейной защиты резервированной системой m/n.

Приведены аналитические зависимости дающие возможность расчета значения вероятности отсутствия срабатывания, излищних срабатываний и правильных срабатываний в резервированной схеме. Определена область одновременного уменьшения вероятности отсутствующих и излишних срабатываний резервированных системой 2/3 схем релейной защиты.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 1, ss. 219-250

621.384.3

Fotoluminescencyjne przetworniki podczerwieni

JERZY WOJCIECHOWSKI (WARSZAWA)

Instytut Technologii Elektronowej
przy Naukowo-Produkcyjnym Centrum Półprzewodników, Warszawa

Otrzymano 26.6.1973

Tematem pracy jest nowy rodzaj przetworników energii promieniowania podczerwonego na energię promieniowania w zakresie widzialnym widma elektromagnetycznego, działających na zasadzie wielofotonowego transferu energii z uczulacza do aktywatora i wzbudzania luminescencji antystokesowskiej w jonach pierwiastków z grupy lantanowców (Yb^{3+}-Er^{3+}, Yb^{3+}-Ho^{3+} lub Yb^{3+}-Tm^{3+}). Omówiono zasadę działania przetwornika oraz przedyskutowano warunki decydujące o uzyskaniu jak największej jego sprawności. Przedstawiono problemy związane z zastosowaniem przetwornika w konstrukcji półprzewodnikowych źródeł światła, w których źródłem promieniowania podczerwonego jest dioda elektroluminescencyjna z GaAs:Si.

1. WSTĘP

W ciągu ostatnich lat obserwuje się szybki rozwój półprzewodnikowych źródeł światła — diod elektroluminescencyjnych działających na zasadzie rekombinacji promienistej w złączu *p-n* [1 ÷ 3]. W zależności od rodzaju zastosowanego materiału półprzewodnikowego oraz technologii złącza *p-n* uzyskuje się diody świecące czerwono (dla materiałów typu GaAsP, GaP, GaAlAs), żółtopomarańczowo (dla GaAsP, GaInP) i zielono (dla GaP), a także diody będące źródłem promieniowania w bliskiej podczerwieni, około 900 nm (dla GaAs).

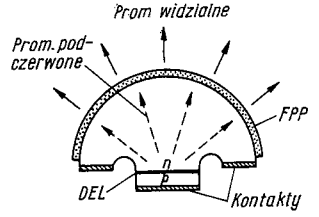
Prowadzone są również prace nad nowymi materiałami (np. GaN) w celu uzyskania diod o niebieskiej barwie świecenia [4].

Ostatnio przedmiotem zainteresowania staje się także nowy rodzaj źródeł światła składających się z diody elektroluminescencyjnej (*DEL*) z arsenku galu oraz warstwy luminoforu spełniającego rolę fotoluminescencyjnego przetwornika podczerwieni (*FPP*) (rys. 1). Część energii promieniowania podczerwonego diody zostaje zaabsorbowana przez luminofor i przetworzona na promieniowanie widzialne. Przez odpowiedni dobór materiału przetwornika jest możliwe uzyskanie świecenia o barwie czerwonej, zielonej lub niebieskiej [5 ÷ 8].

Działanie przetwornika typu *FPP* polega na wzbudzeniu luminescencji antystokesowskiej[1)] w jonach aktywatora-pierwiastka z grupy lantanowców (Er^{3+}, Ho^{3+} lub Tm^{3+}).

[1)] tj. luminescencji nie stosującej się do prawa Stokesa, które mówi, że częstotliwość promieniowania wysyłanego przez fotoluminofor nie może być większa od częstotliwości promieniowania pobudzającego.

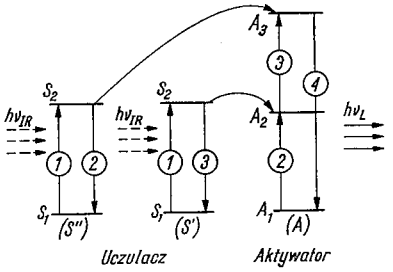
W najprostszym przypadku (rys. 2) ten sam jon aktywatora (A) w wyniku kolejno następującej po sobie absorpcji dwóch fotonów podczerwieni przechodzi ze stanu podstawowego A_1 pośrednio przez stan A_2 do stanu A_3 i przy deekscytacji promienistej daje świecenie w zakresie widzialnym widma elektromagnetycznego. Dla zwiększenia absorpcji podczerwieni przetwornik zawiera znaczną ilość jonów uczulacza (S), którym jest inny z lanta-



Rys. 1. Źródło światła typu *DEL-FPP* (schematycznie)

nowców, a mianowicie iterb Yb^{3+}, mający jedno pasmo absorpcji z maksimum dla około 960 nm. Ze wzbudzonych jonów Yb^{3+} w stanie S_2 następuje przekazywanie energii wzbudzenia do jonu aktywatora i wzbudzenie go do wyższych stanów S_2 i S_3.

Jako źródło podczerwieni stosuje się diodę elektroluminescencyjną z arsenku galu domieszkowanego krzemu po obu stronach złącza *p-n*; przez zmianę koncentracji krzemu można przesuwać maksimum charakterystyki widmowej promieniowania diody GaAs: Si od około 920 nm do 1000 nm [9] i w ten sposób uzyskać dopasowanie charakterystyki promieniowania diody do charakterystyki absorpcji przetwornika.



Rys. 2. Schemat działania przetwornika typu *FPP*

Zainteresowanie tego rodzaju kombinowanym źródłem światła (*DEL-FPP*) wynika m.in. z faktu, że:

— technologia monokrystalizacji GaAs przebiegającej pod normalnym ciśnieniem jest znacznie prostsza w porównaniu do technologii GaP otrzymywanego metodą Czochralskiego pod ciśnieniem około 50 Atm, a przez to cena GaAs jest znacznie niższa,

— sprawność diod z GaAs jest duża; zewnętrzna wydajność kwantowa η_{q_z} diod GaAs: Si z obudową półsferyczną z GaAs sięga 32% [9], natomiast wydajność η_{q_z} świecących zielono diod z GaP jest znacznie niższa i wynosi obecnie maksymalnie 0,6% [10],

— sprawność η_e źródła promieniowania *DEL-FPP* dla zieleni wynosi obecnie $0{,}1 \div 0{,}3\%$, jest więc niewiele mniejsza od uzyskiwanej dla diod z GaP,

— dla świecenia niebieskiego sprawność η_e układu *DEL-FPP* wynosi jak dotychczas nie więcej niż 0,01%, jednak jest o 1 rząd wielkości lepsza od sprawności diod elektroluminescencyjnych z SiC,

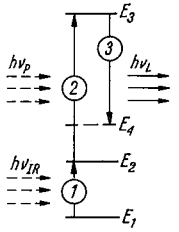
— w związku z pracami prowadzonymi w szeregu ośrodków, głównie w USA, ZSRR, Japonii, Francji i Holandii, oczekuje się, że w najbliższej przyszłości sprawność układu *DEL-FPP* wzrośnie co najmniej do 1% [7].

— w przyszłości realna jest także możliwość opracowania na bazie przetworników typu *FPP* świetlnej mozaiki wielobarwnej na jednym rodzaju podłoża z GaAs.

W dalszej części pracy omówiona będzie szerzej zasada działania fotoluminescencyjnego przetwornika podczerwieni, będzie podany opis stosowanych technologii, a także będą przedstawiane podstawowe parametry przetwornika i układu dioda-przetwornik.

2. ZASADA DZIAŁANIA PRZETWORNIKA

Prace nad przetwornikami typu *FPP* datują się od roku 1959, kiedy to Bloembergen [11] przedstawił koncepcję licznika kwantów energii podczerwieni, pokazaną schematycznie na rys. 3. Przetworzenie promieniowania podczerwonego (zwykle o długości fali $1 \div 2$ μm) na promieniowanie w zakresie widzialnym widma elektromagnetycznego następuje tu na



Rys. 3. Schemat licznika kwantów podczerwieni wg [11]

drodze dwustopniowego wzbudzenia luminescencji jonów aktywatora, przez absorpcję fotonu podczerwieni $h\nu_{IR}$ (przejście 1) oraz fotonu promieniowania pompującego $h\nu_p$ (przejście 2). W nieobecności promieniowania podczerwonego oraz gdy można pominąć obsadzanie poziomu E_2 na drodze cieplnej ($h\nu_{IR} \gg kT$), promieniowanie pompujące nie może być absorbowane i luminescencji nie obserwuje się. Natomiast w obecności promieniowania podczerwonego następuje wzbudzenie jonu aktywatora do stanu metastabilnego E_2 i pompowanie energii na poziom E_3, skąd są możliwe przejścia promieniste do poziomów E_2 i E_1. Świecenie to jest trudne do wykrycia w obecności intensywnego strumienia pompującego, a także jest częściowo reabsorbowane przez licznik. W przypadku gdy między poziomami E_3 i E_2 występuje poziom E_4 luminescencja związana z przejściem 3 może być rejestrowana już bez przeszkód.

Powyższa koncentracja licznika 4-poziomowego została rozwinięta przez Esterowitza i współpracowników [12] do postaci licznika 5-poziomowego i zrealizowana przy użyciu trójdodatnich jonów pierwiastków z grupy lantanowców: Pr^{3+}, Nd^{3+}, Eu^{3+}, Tb^{3+}, Dy^{3+}, Ho^{3+}, Er^{3+}, Tm^{3+}, mających korzystną pod tym względem, wieloprzejściową strukturę

poziomów energetycznych [13]; jako osnowy zastosowano monokryształy, takie jak: CaF_2, CdF_2, CeF_3, $LaCl_3$, $CaWO_4$ i Al_2O_3. Dla wymienionych jonów lantanowców poziomy energetyczne są związane z elektronami znajdującymi się w wewnętrznej, niecałkowicie obsadzonej powłoce elektronowej 4f. Elektrony 4f są skutecznie ekranowane przez zewnętrzne, wypełnione powłoki elektronowe 5s i 5p. W wyniku słabego oddziaływania pola krystalicznego na elektrony powłoki 4f widma absorpcji, a także widma fluorescencji i widma emisyjne mają strukturę wąskopasmową. Jednocześnie położenie poziomów energetycznych tych jonów ulega jedynie niewielkim zmianom przy zmianie rodzaju osnowy.

Licznik zrealizowany według koncepcji Bloembergena jest więc wąskopasmowym przetwornikiem podczerwieni na promieniowanie widzialne. Wydajność jego jest jednak mała, gdyż prawdopodobieństwo absorpcji dwóch kolejnych fotonów w jednym centrum aktywatora jest bardzo małe, m.in. ze względu na ograniczoną (zwykle do około 1%) zawartość jonów aktywatora; przy większych koncentracjach zaczynają przeważać przejścia bezpromieniste związane z tzw. stężeniowym wygaszaniem luminescencji [14].

Koncepcja licznika Bloembergena została zmodyfikowana przez Auzela [15] [16] oraz Owsjankina i Feofiłowa [17], którzy do sieci krystalicznej osnowy wprowadzili dwa rodzaje domieszek: uczulacz (S) i aktywator (A), oba w postaci jonów pierwiastków z grupy lantanowców. Rolę uczulacza wprowadzanego w znacznych ilościach spełniają jony iterbu Yb^{3+}, który charakteryzuje się jednym pasmem absorpcji przypadającym w bliskiej podczerwieni, z maksimum dla około 960 nm, w związku z przejściami między poziomami energetycznymi $^2F_{7/2}$ i $^2F_{5/2}$. Poziom wzbudzony $^2F_{5/2}$ jest poziomem metastabilnym, z którego może być przekazywana energia na poziomy wzbudzone jonów aktywatora.

Jeżeli jony uczulacza S i jon aktywatora A są położone dostatecznie blisko siebie, to może mieć miejsce następujący proces (rys. 2). Jon S' absorbuje foton promieniowania podczerwonego o częstotliwości $\nu_{IR} = \frac{S_2 - S_1}{h}$ i przechodzi ze stanu podstawowego S_1 do stanu wzbudzonego S_2. Przy przejściu do stanu podstawowego S_1 z jonu S' jest przekazywany kwant energii wzbudzenia do jonu A, który w pierwszym stopniu wzbudzenia przechodzi ze stanu podstawowego A_1 do stanu A_2. W czasie trwania wzbudzonego stanu A_2 sąsiedni jon S'' pochłania foton promieniowania podczerwonego (ewentualnie znajduje się już we wzbudzonym stanie S_2) i po przejściu do stanu podstawowego S_1 przekazuje energię wzbudzenia do jonu A, który w drugim stopniu wzbudzenia przechodzi ze stanu A_2 do A_3. Przy przejściu jonu A do stanu podstawowego emitowany jest foton promieniowania widzialnego o częstotliwości $\nu_L = \frac{A_3 - A_1}{h} \cong \frac{2(A_2 - A_1)}{h} \cong \frac{2(S_2 - S_1)}{h}$. Tak więc przez dodanie dwóch fotonów $h\nu_{IR}$ uzyskuje się emisję fotonu o prawie dwukrotnie większej częstotliwości. Dokładna wartość częstotliwości ν_L zależy od tego, czy przenoszenie energii w obu przejściach z jonu S do A jest rezonansowe, czy też odbywa się ze stratą energii (z udziałem przejść fononowych).

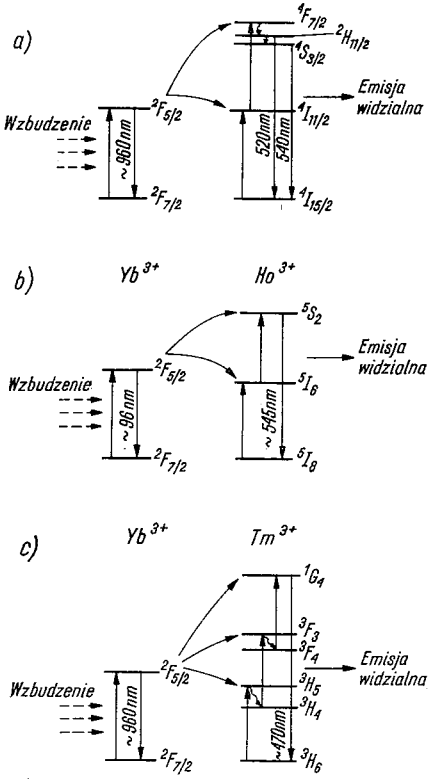
W podobny sposób następuje dodawanie trzech, a nawet czterech fotonów i można uzyskać emisję fotonu o częstotliwości prawie trzy- lub czterokrotnie większej od częstotliwości ν_{IR}.

Z powyższego wynika, że jeśli chodzi o dobór aktywatora, to winien się on charakteryzować układem możliwie równoodległych poziomów energetycznych, przy czym różnica

energii między dwoma kolejnymi poziomami powinna być jak najbliższa energii wzbudzenia jonu Yb^{3+}, aby było możliwe wielokrotne dodawanie fotonów $h\nu_{IR}$.

Jednocześnie powinno być w takim układzie mało poziomów pośrednich, zbędnych z punktu widzenia pompowania energii w aktywatorze, a prowadzących do dodatkowej straty energii.

Analiza struktury poziomów energetycznych jonów lantanowców [13] wykazuje, że największe różnice między poziomami energetycznymi występują dla najcięższych jonów: erbu Er^{3+}, hołmu Ho^{3+} i tulu Tm^{3+}. Te właśnie jony zostały użyte w roli aktywatora przez Auzela oraz dalszych pracach [18÷49].

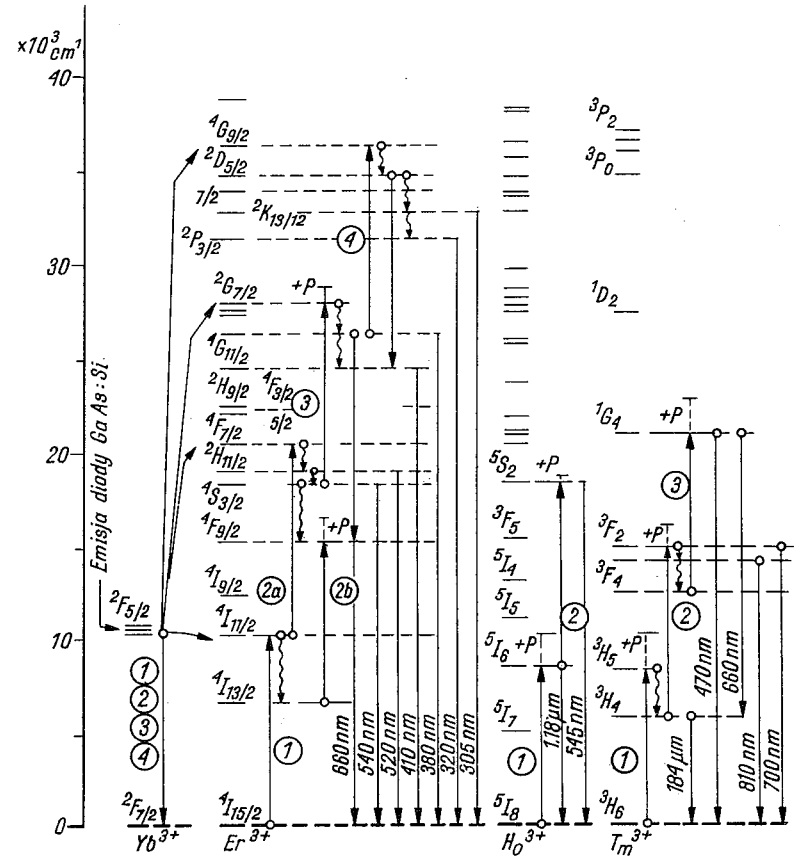


Rys. 4. Schemat transferu energii w układach: a) Yb^{3+}-Er^{3+}, b) Yb^{3+}-Ho^{3+}, c) Yb^{3+}-Tm^{3+}

W wymienionych jonach aktywatorów występują poziomy Er $^4I_{11/2}$, Ho $^5I_6$ i Tm $^3H_5$, na które jest możliwy transfer energii z poziomu Yb $^2F_{5/2}$ uczulacza (rys. 4).

Dla jonu Er^{3+} jest możliwy transfer energii [(Yb $^2F_{5/2} \rightarrow {}^2F_{7/2}$): (Er $^4I_{15/2} \rightarrow {}^4I_{11/2}$)] typu rezonansowego, gdyż względne położenia poziomów Er $^4I_{11/2}$ i Yb $^2F_{5/2}$ pokrywają się. Drugi transfer energii z jonu Yb^{3+} do jonu Er^{3+} [(Yb $^2F_{5/2} \rightarrow {}^2F_{7/2}$): (Er $^4I_{11/2} \rightarrow {}^4F_{7/2}$)] jest również typu rezonansowego. Z poziomu $^4F_{7/2}$ następuje szybkie przejście multifonowe poprzez poziom $^2H_{11/2}$ do poziomu $^4S_{3/2}$, a luminescencja widzialna o długości fali około 540 nm powstaje w wyniku przejścia elektronu z poziomu $^4S_{3/2}$ na poziom podstawowy $^4I_{15/2}$.

Podobny mechanizm przenoszenia energii występuje w układzie Yb^{3+}—Ho^{3+}. Pierwszy transfer energii z jonu Yb^{3+} do jonu Ho^{3+} [(Yb $^2F_{5/2} \rightarrow {}^2F_{7/2}$): (Ho $^5I_8 \rightarrow {}^5I_6$)] jest nierezonansowy, gdyż poziom Ho $^5I_6$ leży poniżej poziomu Yb $^2F_{5/2}$; nadmiar energii jest oddawany do sieci krystalicznej. Drugi transfer [(Yb $^2F_{5/2} \rightarrow {}^2F_{7/2} \rightarrow {}^2F_{7/2}$:(Ho $^5I_6 \rightarrow {}^5S_2$)] można uważać za rezonansowy. Z przejściem $^5S_2 \rightarrow {}^5I_8$ jest związana luminescencja widzialna o długości fali około 545 nm.



Rys. 5. Możliwe przejścia promieniste w układach Yb^{3+}-Er^{3+}, Ho^{3+} i Tm^{3+}

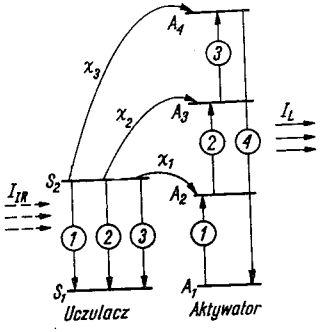
W układzie Yb^{3+}—Tm^{3+} następują kolejno po sobie trzy transfery energii, przy czym wszystkie są typu nierezonansowego. Po pierwszym transferze [(Yb $^2F_{5/2} \rightarrow {}^2F_{7/2}$): (Tm $^3H_6 \rightarrow {}^3H_5$)] jon Tm^{3+} przechodzi ze stanu podstawowego do stanu $^3H_5$, a po deekscytacji multifonowej do stanu $^3H_4$. Po drugim transferze energii [(Yb $^2F_{5/2} \rightarrow {}^2F)_{7/2}$: : (Tm $^3H_4 \rightarrow {}^3F_3$)] i po deekscytacji multifononowej jon Tm^{3+} przechodzi do stanu $^3F_4$, a w wyniku trzeciego transferu [(Yb $^2F_{5/2} \rightarrow {}^2F_{7/2}$): (Tm $^3F_4 \rightarrow {}^1G_4$)] — do stanu $^1G_4$. Świecenie niebieskie o długości fali około 470 nm jest wynikiem przejścia $^1G_4 \rightarrow {}^3H_6$.

We wszystkich trzech opisanych układach przy wzbudzaniu ich podczerwienią otrzymuje się także inne, poza wymienionymi wyżej, pasma promieniowania. Wynika to stąd,

że schematy poziomów wzbudzonych aktywatorów są w rzeczywistości bardziej złożone niż to przedstawiono na rys. 4.

Na rys. 5 pokazano wyznaczone w pracy Dieke'a [13] poziomy energetyczne dla jonów Er^{3+}, Ho^{3+} i Tm^{3+} oraz zaznaczono możliwe przejścia promieniste.

Szczególnie bogate widmo może wystąpić dla Er^{3+} [21] [25]. Świecenie czerwone (dla 660 nm) otrzymuje się wtedy, gdy przeważa proces 2b, w którym po przejściu niepromienistym z poziomu $^4I_{11/2}$ na poziom $^4I_{13/2}$ następuje drugi transfer energii [(Yb $^2F_{5/2} \rightarrow$ $\rightarrow {}^2F_{7/2}$): (Er $^4I_{13/2} \rightarrow {}^4F_{9/2}$)] i wzbudzenie jonu Er^{3+} do stanu $^4F_{9/2}$. Po trzecim transferze energii [(Yb $^2F_{5/2} \rightarrow {}^2F_{7/2}$): (Er $^4F_{9/2} \rightarrow {}^4G_{11/2}$)] możliwe jest przejście promieniste $^4G_{11/2} \rightarrow {}^4I_{15/2}$ ewnetualnie $^2H_{9/2} \rightarrow {}^4I_{15/2}$ dające świecenie we fiolecie.



Rys. 6. Dodawanie trzech fotonów w układzie uczulacz-aktywator (schemat uproszczony)

Dla niektórych typów osnowy może nastąpić także rezonansowy transfer zwrotny energii z Er^{3+} do Yb^{3+}, któremu odpowiada przejście z Er $^4G_{11/2}$ do $^4F_{9/2}$; towarzyszy temu wzmocnienie promieniowania czerwonego. Obserwuje się wtedy zmianę barwy świecenia z zielonej na czerwoną przy wzroście intensywności pobudzania przetwornika [25].

Możliwe jest także dodawania czterech fotonów, w wyniku którego obserwuje się promieniowanie w ultrafiolecie o długości fali 380, 320 i 305 nm.

Z powyższego wynika, że przenoszenie energii w układzie „uczulacz—aktywator" jest procesem złożonym. Prawdopodobieństwo realizacji mechanizmu podstawowego, zgodnie ze schematami podanymi na rys. 4, a także wydajność tego procesu oraz prawdopodobieństwo wystąpienia innych przejść wg rys. 5 zależy od bardzo wielu czynników.

Badanie tych zależności było tematem licznych prac [7] [28] [33] [50 ÷ 53], w których jako podstawę rozważań przyjęto równania fenomenologiczne opisujące zmianę obsadzeń poziomów energetycznych.

Dla przestawienia istoty zagadnienia rozpatrzmy przenoszenie energii do aktywatora jako dodawanie trzech fotonów zachodzące w uproszczonym układzie jak na rys. 6. Układ ten o przejściach typu nierezonansowego odpowiada układowi Yb^{3+}—Tm^{3+}, w którym nie uwzględniono przejść multifonowych w aktywatorze z uwagi na to, że są to przejścia bardzo szybkie, o czasach rzędu 1 μs [32] oraz pominięto dodatkowe przejścia promieniste nie zaznaczone na rys. 4.

Dla tego uproszczonego modelu równania opisujące zmianę obsadzeń w warunkach stacjonarnych mają postać następującą:

$$\frac{dn_{s_2}}{dt} = \frac{\sigma_s I_{IR}}{E_s} n_{s_1} - \frac{n_{s_2}}{\tau_s} - \chi_1 n_{s_2} n_{A_1} - \chi_2 n_{s_2} n_{A_2} - \chi_3 n_{s_2} n_{A_2} = 0, \tag{1}$$

$$\frac{dn_{A_2}}{dt} = \chi_1 n_{s_2} n_{A_1} - \frac{n_{A_2}}{\tau_{A_2}} - \chi_2 n_{s_2} n_{A_2} = 0, \tag{2}$$

$$\frac{dn_{A_3}}{dt} = \chi_2 n_{s_2} n_{A_2} - \frac{n_{A_3}}{\tau_{A_3}} - \chi_3 n_{s_2} n_{A_3} = 0, \tag{3}$$

$$\frac{dn_{A_4}}{dt} = \chi_3 n_{s_2} n_{A_3} - \frac{n_{A_4}}{\tau_{A_4}} = 0, \tag{4}$$

gdzie:

σ_s — przekrój czynny na absorpcję dla przejścia Yb $^2F_{2/7} \rightarrow {}^2F_{5/2}$,

E_s — energia odpowiadająca przejściu Yb $^2F_{7/2} \rightarrow {}^2F_{5/2}$,

I_{IR} — natężenie promieniowania pobudzającego,

n — liczba obsadzeń stanów energetycznych,

τ — średni czas życia wzbudzonego stanu,

χ — współczynnik prawdopodobieństwa transferu energii z uczulacza do aktywatora.

Z rozwiązania tego zespołu równań przy założeniu, że dla małych natężeń promieniowania I_{IR} obsadzenie n poziomów wzbudzonych jest małe w porównaniu z obsadzeniem poziomu podstawowego (tj. $n_{s1} \cong n_s$, $n_{A1} \cong n_A$) i po pominięciu ostatnich członów równań (1) i (3), można otrzymać wyrażenie na moc promieniowania widzialnego (na jednostkę objętości) będącego wynikiem sumowania trzech fotonów

$$P_L^{\mathrm{III}} = \frac{E_{A_{41}}}{\tau'_{A_{41}}} \tau_{A_4} \tau_{A_3} \frac{\varkappa_1 \chi_2 \chi_3 n_A}{(\tau_{A_2})^{-1} + \chi_2 n_{s_2}} n_{s_2}^3, \tag{5}$$

gdzie:

$E_{A_{41}}$ — energia fotonu prom. widzialnego ($= h\nu_L$),

$\tau'_{A_{41}}$ — czas życia jonu Er^{3+} w stanie 4 ze względu na przejście promieniste do stanu 1.

Jednocześnie zależność między intensywnością pobudzania a obsadzeniem stanu $^2F_{5/2}$ jonu Yb^{3+} jest określona jako

$$I_{IR} = \frac{E_s}{\sigma_s \cdot n_s} \left[\frac{1}{\tau_s} + \chi_1 n_A \left(1 + \frac{1}{1 + (\tau_{A_2} \chi_2 n_{s_2})^{-1}} \right) \right] n_{s_2}. \tag{6}$$

Przy słabym pobudzaniu równanie (6) może być uproszczone do postaci

$$I_{IR} \cong \frac{E_s}{\sigma_s \cdot n_s} \left(\frac{1}{\tau_s} + \chi_1 n_A \right) n_{s_2}. \tag{7}$$

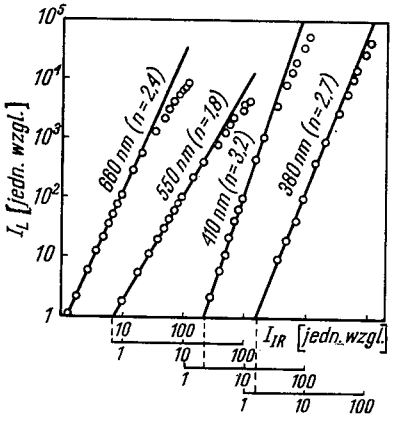
Przy sumowaniu dwóch fotonów, co ma miejsce dla układu Yb^{3+}—Er^{3+} i Yb^{3+}—Ho^{3+}, moc promieniowania widzialnego

$$P_L^{\mathrm{II}} = \frac{E_{A_{31}}}{\tau'_{A_{31}}} \tau_{A_3} \frac{\chi_1 \chi_2 n_A}{(\tau_{A_2})^{-1} + \chi_2 n_{s_2}} n_{s_2}^2. \tag{8}$$

Z powyższych równań (7) i (8) wynika, że zależność mocy promieniowania widzialnego P_L oraz natężenia luminescencji I_L od natężenia pobudzania I_{IR} jest typu wykładniczego, czyli

$$I_L \doteq I_{IR}^{\alpha} \tag{9}$$

Wartość α jest charakterystyczna dla danego pasma emisji i jest związana z krotnością transferu energii z pochłaniacza do aktywatora i w przypadku sumowania dwóch fotonów $\alpha = 2$, a w przypadku sumowania 3 fotonów $\alpha = 3$. Na tej podstawie można z pomierzonych doświadczalnie charakterystyk $I_L(I_{IR})$ określić mechanizm przejść fotonowych. Gdy np. w układzie Yb^{3+}—Er^{3+} wykładnik α dla promieniowania o długości fali 540 nm wynosi $\alpha \leqslant 2$, to następuje dwukrotny transfer energii, gdy dla promieniowania 660 nm $2 < \alpha < 3$, to następuje dwukrotny i trzykrotny transfer energii, a natomiast gdy dla promieniowania 410 nm $\alpha > 3$, to następuje trzykrotny i czterokrotny transfer energii na każdy foton w zakresie widzialnym (rys. 7).



Rys. 7. Zależność $I_L(I_{IR})$ dla *FPP* typu $Y_3OCl_7:Yb^{3+}$, Er^{3+} [21]

Dalszy istotny wniosek to taki, że sprawność przetwornika typu *FPP* zmienia się z $(\alpha - 1)$ potęgą natężenia pobudzenia. Dla barwy czerwonej i zielonej sprawność jest w przybliżeniu wprost proporcjonalna do natężenia pobudzenia, a dla barwy niebieskiej — wzrasta z kwadratem I_{IR}.

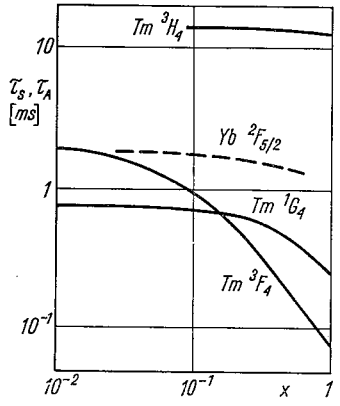
Sprawność przetwornika jest tym większa, im większe są:
— przekrój czynny na absorpcję w uczulaczu,
— koncentracja jonów uczulacza i aktywatora,
— czasy życia stanów wzbudzonych uczulacza i aktywatora,
— prawdopodobieństwo transferu energii z uczulacza do aktywatora.

Ponadto sprawność zależy od nieuwzględnionego w równaniach $(1 \div 4)$ ewentualnego transferu zwrotnego energii z aktywatora do uczulacza w przypadku przejść rezonansowych oraz od występowania przejść promienistych dających dodatkowe pasma emisji. Należy również uwzględnić przejścia bezpromieniste związane ze stężeniowym wygaszaniem luminescencji.

Wpływ wielkości tych czynników zależy w głównej mierze od rodzaju użytej osnowy. Podstawowe wymagania stawiane materiałowi osnowy są takie jak w przypadku innych

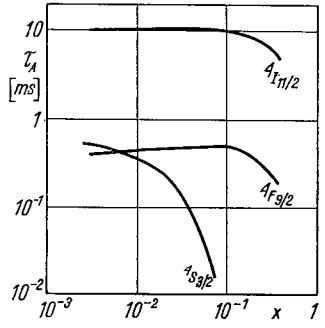
materiałów luminescencyjnych [14]. Dotyczy to niewprowadzania własnych pasm absorpcji w paśmie absorpcji i promieniowania przetwornika, wymiarów jonów zbliżonych do wymiarów jonów uczulacza i aktywatora itp. Stosowane osnowy bazują na związkach itru, lantanu oraz gadolinu, głównie w postaci fluorków, a także tlenków, tlenochlorków i tlenosiarczków.

Jak wynika z [13], widma optyczne jonów lantanowców nie wykazują większych różnic dla różnych rodzajów osnowy. Jednak już np. niewielkie różnice poziomów aktywatora wpływać mogą na mechanizm (rezonansowy lub nierezonansowy) przenoszenia energii



Rys. 8. Zależność $\tau_s(x)$ dla $Y_{1-x}Yb_xF_3$ (— — —) oraz $\tau_A(x)$ dla $Y_{0,9997-x}Yb_xTm_{0,0003}F_3$ (———) [32]

z aktywatora. W praktyce pole krystaliczne osnowy oddziaływuje więc bardzo silnie na takie podstawowe parametry, jak prawdopodobieństwo pompowania energii do aktywatora i czasy życia wzbudzonych stanów, a w związku z tym i na sprawność przetwornika. Dobór odpowiedniego materiału osnowy, a także ustalenie optymalnej koncentracji uczulacza i aktywatora prowadzone są zwykle eksperymentalnie.



Rys. 9. Zależność $\tau_A(x)x$ dla $Y_{1-x}Er_xF_3$ [32]

Najczęściej stosowanym kryterium jest czas życia wzbudzonych stanów, który powinien być możliwie długi, a który łatwo jest wyznaczyć pobudzając przetwornik impulsami promieniowania.

W osnowie domieszkowanej jedynie iterbem czas życia τ_S stanu Yb $^2F_{5/2}$ jest długi, rzędu kilku ms i w niewielkim stopniu zależy od zawartości iterbu. Na rys. 8 [32] pokazana jest przykładowo linią kreskowaną zależność τ_S od zawartości x iterbu dla materiału $Y_{1-x}Yb_xF_3$. Przy większych wartościach x obserwuje się nieco krótsze czasy τ_S ze względu na samowygaszanie stężeniowe wzbudzenia.

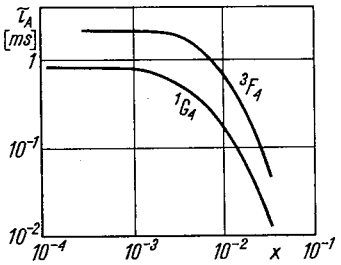
Jeśli chodzi o czasy życia stanów wzbudzonych aktywatora, to wyniki pomiarów są następujące [32] [33]. Dla materiału $Y_{1-x}Er_xF_3$ czas życia τ_A stanu $^4I_{11/2}$ jest rzędu 10 ms i przy wzroście wartości x obserwuje się jedynie niewielki wpływ wygaszania stężeniowego (rys. 9). Natomiast czas życia stanu $^4S_{3/2}$ maleje gwałtownie, gdy zawartość erbu wzrasta powyżej paru procent. Tłumaczy się to następującymi możliwymi rezonansowymi procesami relaksacyjnymi:

$$(^2H_{11/2} \rightarrow {}^4I_{9/2}): (^4I_{15/2} \rightarrow {}^4I_{13/2})$$

lub

$$(^2H_{11/2} \rightarrow {}^4I_{13/2}): (^4I_{15/2} \rightarrow {}^4I_{9/2}),$$

przy czym poziom $^2H_{11/2}$ jest obsadzany w temperaturze pokojowej z blisko położonego poziomu $^4S_{3/2}$.



Rys. 10. Zależność $\tau_A(x)$ dla $Y_{1-x}Tm_xF_3$ [32]

W praktyce zawartość erbu jest w związku z tym ograniczona do $1-2\%$.

Dla aktywatora Ho^{3+} występuje podobny proces relaksacji

$$(^5S_2 \rightarrow {}^5I_4): (^5I_8 \rightarrow {}^5I_7),$$

ograniczający zawartość holmu do $0{,}1 \div 0{,}3\%$. Jednocześnie czasy życia wzbudzonych stanów holmu są krótsze niż dla erbu, np. czas życia wzbudzonego stanu $^5S_2$ wynosi około 0,13 ms [26], co preferuje erb jako aktywator w układzie z dwufotonowym transferem energii.

Dla Tm^{3+} czasy życia wzbudzonych stanów $^3F_4$ i $^1G_4$ są rzędu $1 \div 3$ ms, jednak maleją szybko ze wzrostem zawartości aktywatora (rys. 10), zawartość tulu jest więc ograniczona do około 0,3%. Z poziomu $^3F_4$ następuje relaksacja

$$(^3F_4 \rightarrow {}^3H_4): (^3H_6 \rightarrow {}^3H_4),$$

a z poziomu $^1G_4$ są możliwe relaksacje:

$$(^1G_4 \rightarrow {}^3F_2): (^3H_6 \rightarrow {}^3H_4),$$
$$(^1G_4 \rightarrow {}^3H_4): (^3H_6 \rightarrow {}^3F_2),$$
$$(^1G_4 \rightarrow {}^3F_4): (^3H_6 \rightarrow {}^3H_5),$$

lub

$$(^1G_4 \rightarrow {}^3H_5): (^3H_6 \rightarrow {}^3F_4).$$

W osnowie zawierającej zarówno aktywator jak i uczulacz występuje poza opisanymi procesami relaksacyjnymi w aktywatorze także relaksacja uczulacz-aktywator. Obserwuje się mianowicie wpływ zawartości uczulacza na czas życia na poziomach wzbudzonych aktywatora. W układzie Yb^{3+}—Er^{3+} może to być proces relaksacji

$$(Er\ {}^4S_{3/2} \rightarrow {}^4I_{11/2}):(Yb\ {}^2F_{7/2} \rightarrow {}^2F_{5/2}),$$

a w układzie Yb^{3+}—Tm^{3+} procesy

$$(Tm\ {}^3F_4 \rightarrow {}^3H_6):\ (Yb\ {}^2F_{7/2} \rightarrow {}^2F_{5/2}),$$

$$(Tm\ {}^1G_4 \rightarrow {}^3H_5):\ (Yb\ {}^2F_{7/2} \rightarrow {}^2F_{5/2})\ \text{i inne.}$$

Na rys. 8 pokazano (linią ciągłą) zależność czasu życia stanów wzbudzonych aktywatora od zawartości iterbu w przetworniku $Y_{0,9997-x}Yb_xTm_{0,0003}F_3$.

W związku z powyższymi procesami relaksacji optymalna zawartość iterbu w osnowie YF_3 wynosi przy aktywacji erbem około 20%, a przy aktywacji tulem — 35%.

Dalsze czynniki definiujące pracę przetwornika typu *FPP* to przekrój czynny σ_s na absorpcję w uczulaczu oraz współczynnik prawdopodobieństwa transferu energii z uczulacza do aktywatora.

Przekrój czynny σ_s jest to stosunek liczby przejść na wyższy poziom wzbudzenia, przypadających w sekundzie na jeden jon uczulacza do liczby fotonów padających w sekundzie na 1 cm^2 powierzchni przetwornika. Przekrój czynny można wyznaczyć przez całkowanie graficzne iloczynu z charakterystyki widmowej pochłaniania uczulacza i charakterystyki emisji diody GaAs: Si. Dla uzyskania dużej sprawności przetwornika przekrój czynny σ_s powinien być możliwie duży: oznacza to, że charakterystyki pochłaniania uczulacza i promieniowania diody powinny jak najbardziej zachodzić na siebie.

Dla jonu Yb^{3+} mającego konfigurację elektronową $4f^{13}$ pasmo absorpcji występujące w bliskiej podczerwieni jest znacznie szersze od pasm absorpcji dla innych lantanowców, takich jak Er^{3+}, Ho^{3+} i Tm^3, gdyż jest prawdopodobnie związane z przejściami elektronów do powłoki zewnętrznej ($4f \rightarrow 5d$). Niemniej jednak maksima charakterystyk absorpcji Yb^{3+} i emisji diody nie pokrywają się, gdyż typowe diody GaAs: Si mają maksimum promieniowania dla 930÷940 nm, a dla stosowanych najczęściej osnów maksimum absorpcji Yb^{3+} przypada dla 950÷980 nm. Zagadnienie dopasowania źródła podczerwieni i przetwornika będzie rozpatrzone szerzej w rozdz. 4.

Współczynnik prawdopodobieństwa transferu energii na poziomy aktywatora może być wyznaczony w oparciu o równania (1÷4). W ten sposób w pracy[33] wyznaczono wartości χ_1, χ_2, χ_3 dla przejść w układzie Yb^{3+}—Tm^{3+}. Okazało się, że prawdopodobieństwo transferu energii w bardzo małym stopniu zależy od zawartości jonów uczulacza i aktywatora, natomiast o wartości decyduje głównie to, o ile transfer energii odbiega od transferu rezonansowego. Dla przejścia

$$(Yb\ {}^2F_{5/2} \rightarrow {}^2F_{7/2}):\ (Tm\ {}^3H_6 \rightarrow {}^3H_5),$$

dla którego niedopasowanie poziomów Yb $^2F_{5/2}$ i Tm $^3H_5$ wyrażone w liczbach falowych było rzędu 1600 cm^{-1} $\chi_1 = 10^{-17}$ cm^3 s^{-1}, dla przejścia związanego z dodawaniem dru-

giego fotonu (przy niedopasowaniu ok. 1000 cm^{-1}) $\chi_2 = 9.10^{-15}$ cm^3 s^{-1}, a dla przejścia związanego z dodawaniem trzeciego fotonu (przy niedopasowaniu ok. 1400 cm^{-1}) $\chi_3 = 2{,}7 \cdot 10^{-16}$ cm^3 s^{-1}.

Wzajemne położenie poziomów wzbudzonych uczulacza i aktywatora jest szczególnie istotne dla układu Yb^{3+}—Er^{3+}, w którym jest możliwy rezonansowy transfer energii. Zagadnienie to badali Kuroda i inni [47] stosując jako osnowy YF_3, YOCl i Y_3OCl_7, dla których w zależności od rodzaju osnowy obserwuje się dominujące świecenie zielone lub czerwone.

Dla wymienionych osnow wyznaczono m.in. położenie poziomów wzbudzonych Yb $^2F_{5/2}$ i Er $^4I_{11/2}$ w temperaturze 4°K, a także prawdopodobieństwo wystąpienia procesów multifonowych, takich jak wewnętrzna relaksacja multifonowa (ang. *MPR*) i transfer energii z udziałem fononów (*PAT*) [50].

W YF_3 poziom Er $^4I_{11/2}$ jest położony w skali energii wyrażonej w liczbach falowych 27 cm^{-1} powyżej poziomu Yb $^2F_{5/2}$. Przy wzroście temperatury (absorpcja fononu 27 cm^{-1}) jest więc możliwy transfer energii [(Yb $^2F_{5/2} \rightarrow {}^2F_{7/2}$): (Er $^4I_{15/2} \rightarrow {}^4I_{11/2}$)], a po drugim transferze energii [(Yb $^2F_{5/2} \rightarrow {}^2F_{7/2}$): (Er $^4I_{11/2} \rightarrow {}^4F_{7/2}$)] występuje świecenie zielone. Jednocześnie prawdopodobieństwo χ_{MPR} dla relaksacji (Er $^4I_{11/2} \rightarrow {}^4I_{13/2}$) jest małe i emisja czerwona jest bardzo słaba.

W YOCl poziom Er $^4I_{11/2}$ leży 148 cm^{-1} powyżej poziomu Yb $^2F_{5/2}$, co uniemożliwia transfer energii na ten poziom aktywatora. Natomiast jest możliwy w tej osnowie transfer energii [(Yb $^2F_{5/2} \rightarrow {}^2F_{7/2}$): (Er $^4I_{15/2} \rightarrow {}^4I_{13/2}$)] z udziałem fononów, gdyż znaczna jest wartość χ_{PAT}. Po drugim transferze energii [(Yb $^2F_{5/2} \rightarrow {}^2F_{7/2}$): (Er $^4I_{13/2} \rightarrow {}^4F_{9/2}$)] następuje emisja Er^{3+} w czerwieni.

W Y_3OCl_7 poziom Er $^4I_{11/2}$ leży 41 cm^{-1} poniżej poziomu Yb $^2F_{5/2}$ i po pierwszym transferze energii następuje obsadzenie tego poziomu aktywatora. Jednocześnie w Er^{3+} w osnowie Y_3OCl_7 występuje duże prawdopodobieństwo χ_{MPR} przejścia relaksacyjnego $^4I_{11/2} \rightarrow {}^4I_{13/2}$. Po drugim transferze energii (przejście $^4I_{13/2} \rightarrow {}^4F_{9/2}$) występuje również głównie emisja w czerwieni przy przejściu promienistym $^4S_{3/2} \rightarrow {}^4I_{15/2}$.

Kończąc rozważania nad czynnikami wypływającymi na sprawność *FPP* należy wspomnieć o efekcie nasycenia występującym zazwyczaj przy dużych natężeniach pobudzenia; przebiegi $I_L(I_{IR})$ są mianowicie tego rodzaju, że wykładnik α maleje, co ilustrują charakterystyki pokazane na rys. 7. Przyjmuje się, że przyczyną tego zjawiska jest tendencja do nasycania obsadzeń poziomów pośrednich aktywatora, np. poziomu A_2 (rys. 6) występująca wtedy, gdy zmniejszanie obsadzenia tego poziomu wywołane drugim transferem energii (na poziom A_3) jest większe od zmiany obsadzenia wywołanej przejściem spontanicznym z poziomu A_2 do A_1. Na przykład dla *FPP* typu Y_2O_2S: Yb^{3+}, Er^{3+} o barwie świecenia zielonej wykładnik $\alpha \cong 2$ tylko wtedy, gdy gęstość mocy pobudzania wynosi poniżej 0,3 W/cm^2; dla większych gęstości mocy wartość α wynosi około 1,5, a dla gęstości mocy powyżej 5 W/cm^2 nawet $\alpha \cong 1$, czyli otrzymuje się liniowy przebieg charakterystyki $I_L(I_{IR})$ [37].

Dla przetworników wykonanych na osnowie fluorków efekt nasycenia występuje zwykle przy większych gęstościach mocy pobudzenia. Dla *FPP* typu $BaYF_5$: Yb^{3+}, Er^{3+} Johnson i inn. [7] nie stwierdzili odchylenia od $\alpha \cong 2$ dla gęstości mocy do 10 W/cm^2.

3. TECHNOLOGIA PRZETWORNIKÓW

Ważniejsze rodzaje opracowanych dotychczas przetworników *FPP* podano w tabl. 1. Pod względem rodzaju zastosowanej osnowy oraz technologii przetworniki można podzielić na wykonane na bazie fluorków oraz tlenków, przy czym są to przeważnie związki itru, lantanu, gadolinu ewentualnie lutetu. Najczęściej stosuje się związki itru, rzadziej lantanu i gadolinu, co można uzasadnić małą różnicą w wartościach promieni jonów itru oraz jonów uczulacza i aktywatora, które mogą być wtedy wprowadzone w znacznych ilościach

Ważniejsze rodzaje przetworników typu FPP

Tablica 1

Lp.	*FPP* (wzór sumaryczny)	Główne pasmo emisji (nm)	Technologia	Poz. lit.
1	$Y_{0,89}Yb_{0,1}Er_{0,01}F_3$	540	Krystalizacja fluorków w atmosferze HF(900÷1000C)	20
2	$Y_{0,5985}Yb_{0,4}Tm_{0,0015}F_3$	480		
3	$La_{0,86}Yb_{0,12}Er_{0,02}F_3$	540		
4	$La_{0,79}Yb_{0,2}Ho_{0,01}F_3$	540		
5	$La_{0,7985}Yb_{0,2}Tm_{0,0015}F_3$	480		
6	$Gd_{0,89}Yb_{0,1}Er_{0,01}F_3$	540		
7	$NaY_{0,80}Yb_{0,18}Er_{0,02}F_4$	540	Krystalizacja fluorków z dod. 20% NaF w atmosferze HF (750°C) i w argonie (1000°C)	46
8	$BaLu_{0,92}Yb_{0,05}Ho_{0,03}F_5$	540	Krystalizacja $BaF_2 + MF_3$*) w atmosferze HF (1000°C)	22
9	$BaY_{0,3}Yb_{0,6}Er_{0,1}F_5$	550		
10	$LiY_{0,799}Yb_{0,2}Ho_{0,001}F_4$	540	Krystalizacja fluorków w atmosferze HF, wyciąganie kryształów $LiMF_4$	26
11	$Y_{0,8}Yb_{0,19}Er_{0,01}F_3$	550	Krystalizacja fluorków w atmosferze N_2 (1000°C) pod osłoną BeF_2	24
12	$Y_{0,8}Yb_{0,198}Ho_{0,002}F_3$	540	Krystalizacja fluorków w atmosferze N_2 (1000°C) pod osłoną BeF_2	24
13	$Y_{0,649}Yb_{0,35}Tm_{0,001}F_3$	470		
14	$Y_{0,8}Yb_{0,17}Er_{0,03}F_3$	550	Krystalizacja fluorków w wysokiej próżni (1175°C)	41
15	$Y_{0,74}Yb_{0,25}Er_{0,01}OCl$	660	Krystalizacja chlorków w atmosferze Cl_2 (850°C)	23
16	$(Y_{0,71}Yb_{0,28}Er_{0,01})_3OCl_7$	550 660		
17	$(Y_{0,715}Yb_{0,28}Ho_{0,005})_3OCl_7$	550		
18	$(Y_{0,85}Yb_{0,11}Er_{0,04})_2O_3$	650	Krystalizacja w atmosferze powietrza (1300°C)	42
19	$(Y_{0,86}Yb_{0,08}Er_{0,06})_2O_2S$	550	Krystalizaja z dod. Na_2CO_3, K_3PO_4 i S w atmosferze powietrza (1100°C)	37
20	$(Gd_{0,86}Yb_{0,08}Er_{0,06})_2O_2S$	550		
21	$(La_{0,86}Yb_{0,08}Er_{0,06})_2O_2S$	550		

*) gdzie M = Y, La, Gd lub Lu (z dodatkiem uczulacza i aktywatora)

do sieci krystalicznej osnowy. Np. promienie wymienionych jonów wyznaczone dla tlenosiarczków mają następujące wartości [37]:

jon Yb^{3+} — promień jonu 0,90 *A*,
Tm^{3+} — promień jonu 0,91 *A*,
Er^{3+} — promień jonu 0,92 *A*,
Ho^{3+} — promień jonu 0,95 *A*,
Y^{3+} — promień jonu 0,93 A,
Gd^{3+} — promień jonu 0,98 A,
La^{3+} — promień jonu 1,10 A,
Lu^{3+} — promień jonu 0,91 A.

Materiałami wyjściowymi w technologii przetworników są wysokiej czystości tlenki: Y_2O_3 i La_2O_3 o czystości 99,9999% ewentualnie 99,999%, Yb_2O_3 jako wprowadzany w dużych ilościach procentowych zwykle 99,999%, a pozostałe tlenki lantanowców — o czystości 99,99% ewentualnie 99,9%.

Rodzaj i ilość zanieczyszczeń zawartych w materiałach wyjściowych a także wprowadzanych w trakcie procesu technologicznego ma istotny wpływ na sprawność przetwornika i barwę świecenia. Szczególnie szkodliwe są domieszki innych pierwiastków z grupy lantanowców charakteryzujących się licznymi, blisko położonymi poziomami energetycznymi, przez co wzrasta liczba przejść niepromienistych z poziomów wzbudzonych uczulacza i aktywatora i w ten sposób obniżona zostaje sprawność przetwornika. Autorzy pracy [43] podają, że np. wzrost zawartości dysprozu Dy^{3+} z 1 ppm do 10 ppm powoduje spadek sprawności przetwornika co najmniej o jeden rząd wielkości. Przyczyną tego zjawiska jest transfer energii z poziomów Er $^4I_{11/2}$ i Yb $^2F_{5/2}$ na niżej położone poziomy Dy^{3+} [49].

Najczęściej stosowane przetworniki są wykonywane na osnowie fluorków. W pierwszej fazie procesu technologicznego otrzymuje się mieszaninę fluorków typu MF_3 według składów podanych w tabl. 1, zwykle przez rozpuszczenie odpowiednich tlenków w kwasie azotowym, dodanie nadmiaru kwasu fluorowodorowego i wytrącenie fluorków, a następnie wyizolowanie ich przez odfiltrowanie i dehydratację. Przeprowadzenie wszystkich składników do roztworu daje gwarancję całkowitego ich wymieszania. Także w przypadku fluoryzacji tlenków na drodze suchej (w atmosferze gazowego fluorowodoru) mieszaninę odpowiednich tlenków rozpuszcza się uprzednio w kwasie azotowym, wytrąca w postaci szczawianów i przez prażenie w atmosferze powietrza otrzymuje się ponownie tlenki [20].

Strącanie fluorków i dehydratacja prowadzone są w taki sposób, aby zapobiec okludacji zanieczyszczeń obecnych w roztworach oraz usunąć resztę wody, która powoduje hydrolizę fluorków wpływając ujemnie na własności luminescencyjne. M. in. Parker i Johnson [48] zalecają wielokrotne przemywanie i strącanie osadu oraz długotrwałe wygrzewanie w temperaturze 70°C, co powoduje przejście z postaci galaretowatej do bardziej zbitej, proszkowej, natomiast w pracy [41] stosuje się wielogodzinne suszenie osadu w próżni w temperaturze 60°C.

W następnej operacji otrzymaną mieszaninę fluorków poddaje się prażeniu (krystalizacji) w wysokiej temperaturze (zwykle 900÷1100°C). Jedynie Heves i Sarver [20] łączyli fluoryzację tlenków i krystalizację w jeden proces działając na mieszaninę tlenków suchym gazowym fluorowodorem w temperaturze 900÷1000°C i w ten sposób unikali strącenia fluorków na drodze mokrej.

Prażenie w wysokiej temperaturze ma na celu umożliwienie jonom uczulacza i aktywatora wejścia do sieci krystalicznej osnowy.

Przez powolne oziębianie dąży się do otrzymania materiału mającego strukturę krystaliczną o możliwie małej liczbie defektów. Proces prowadzi się w atmosferze beztlenowej i zwykle z dodatkiem gazowego fluorowodoru. W ten sposób zapewnia się fluoryzację tlenków, które mogą powstać w związku z obecnością tlenu zaabsorbowanego przez osady fluorków lub z hydrolizą fluorków podczas grzania ich w piecu.

Prażenie przeprowadza się z dodatkiem topników, znacznie obniżających wymaganą temperaturę wygrzewania oraz przeciwdziałających procesowi utleniania. Jako topnik stosuje się zazwyczaj fluorek amonu, w połączeniu ze znacznie wyżej topliwymi fluorkami sodu, litu lub baru [22] [26] [46], które tworzą z osnową złożone fluorki $NAYF_4$, $LiYF_4$ lub $BaYF_5$. Najlepsze wyniki uzyskuje się dodając NaF, którego temperatura topnienia wynosi 980°C, a promień jonowy Na^+ jest równy 0,99 A. Na przykład jednym z najbardziej wydajnych przetworników jest przetwornik o składzie $NaY_{0,80}Yb_{0,18}Er_{0,02}F_4$, który otrzymano w pracy [46] stosując najpierw dehydratację mieszaniny fluorków w strumieniu gazowego fluorowodoru w temperaturze 750°C, a następnie krystalizację z dodatkiem 20% NaF w atmosferze argonu w temperaturze 1000°C.

Stosowanie atmosfery gazowego fluorowodoru podczas krystalizacji wymaga kosztownej aparatury platynowej. Dąży się więc do wyeliminowania fluorowodoru z tego procesu.

Oryginalną, często stosowaną metodę opracowali Van Uitert i współpracownicy [24], którzy zastosowali ochronę fluorków przed utlenianiem za pomocą niskotopliwego szkliwa fluorkowego na bazie BeF_2. Do mieszaniny fluorków MF_3 umieszczonej w tyglu platynowym dodaje się fluorek berylu BeF_2 i fluorek amonu NH_4F i całość wygrzewa w 1000°C w przepływie azotu. BeF_2 jest materiałem o temperaturze topnienia 800°C i o małej gęstości, przez co osłania warstwą ochronną krystalizujący luminofor, a jednocześnie na zasadzie reakcji wymiany z tlenkami itru i lantanowców oddaje im fluor, a sam przechodzi w tlenek, tworząc warstwę szkliwa typu $BeO\text{-}BeF_2$.

Ze względu na toksyczne właściwości związków berylu BeF_2 zastępuje się innymi fluorkami, a mianowicie AlF_3 i MgF_3 [41], o temperaturze topnienia odpowiednio 1040°C i 1260°C. Jednak fluorki te dają w reakcji z tlenem związki o strukturze krystalicznej i sprawność otrzymanych przetworników jest na ogół nieco gorsza.

Inny sposób krystalizacji bez udziału gazowego fluorowodoru polega na wyprażaniu fluorków w wysokiej próżni. Low i Major [41] prowadzili prażenie w próżni rzędu 10^{-6} Tr w temperaturze 1175°C, a następnie stosowali długotrwałe ochładzanie do temperatury pokojowej; po tym procesie sprawność przetworników była większa niż otrzymanych przy zastosowaniu dodatku BeF_2 i krystalizacji w atmosferze azotowej.

Technologia przetworników wytwarzanych na osnowie tlenku jest prostsza w porównaniu z technologią fluorków. W przetwornikach tlenkowych dominuje jednak na ogół czerwona barwa świecenia, a ich sprawność jest niższa niż przetworników na osnowie fluorków.

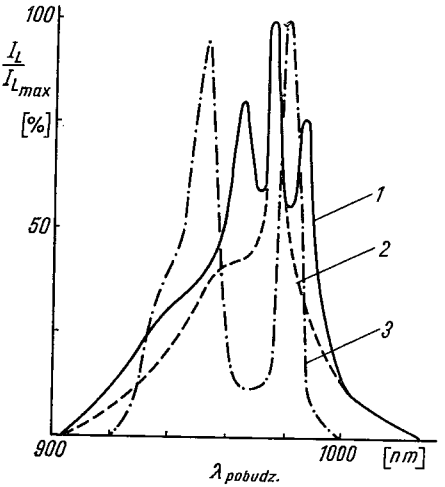
Wittke i współpracownicy [43] otrzymali przetwornik świecący czerwono stosując rozpuszczanie mieszaniny tlenków w kwasie azotowym, strącenie szczawianów i ponowne przejście do tlenków przez prażenie szczawianów w temperaturze 1300°C w atmosferze powietrza.

Zespół Van Uiterta [23] [25] opracował technologię przetworników na bazie tleno-chlorków. Odpowiednie tlenki rozpuszczano w kwasie solnym, a otrzymane chlorki odwadniano i wygrzewano w atmosferze chloru w 850°C. W zależności od warunków dehydratacji i zawartości resztkowego tlenu podczas krystalizacji otrzymywano przetworniki o osnowie YOCl lub Y_3OCl_7 dające w układzie Yb^{3+}—Er^{3+} emisję dla 410 nm, 550 nm i 660 nm (najsilniejsze pasmo emisji).

Przetwornik świecący zielono (z aktywatorem Er^{3+}) oraz niebiesko (z Tm^{3+}) otrzymał Yocom i in. [37] na osnowie tleno-siarczków Y_2O_2S, Gd_2O_2S i La_2O_2S. Mieszanina tlenków z dodatkiem siarki oraz Na_2CO_3 i K_3PO_4 była wyprażana w temperaturze 1100°C w atmosferze powietrza.

4. DOBÓR ŹRÓDŁA PODCZERWIENI

Zagadnienie to obejmuje problemy dopasowania charakterystyki widmowej promieniowania diody elektroluminescencyjnej do charakterystyki absorpcji *FPP* przy zapewnieniu jednoczesnej dużej zewnętrznej wydajności kwantowej η_{q_z} diody, oraz problemy optymalizacji konstrukcji diody i geometrii układu dioda—*FPP* dla uzyskania maksymalnej skuteczności świetlnej tego źródła światła.

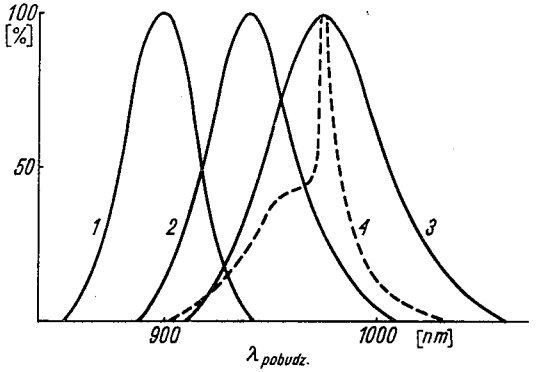


Rys. 11. Charakterystyki wzbudzenia *FPP* dla 1) YF_3:Yb^{3+}, Er^{3+}, 2) LaF_3:Yb^{3+}, Er^{3+} [18] oraz 3)Y_2O_2S:Yb^{3+}, Er^{3+} [37]

Jak wspomniano w rozdz. 2, maksimum charakterystyki absorpcji uczulacza Yb^{3+} przypada w przedziale 950÷980 nm, w zależności od rodzaju zastosowanej osnowy. Pasmo absorpcji można wyznaczyć na podstawie badania widma odbicia promieniowania od warstwy *FPP*, jednak zazwyczaj korzysta się z charakterystyki widmowej wzbudzenia przetwornika [20] [47]. Charakterystyka wzbudzenia *FPP* jest to zależność natężenia promieniowania widzialnego od długości fali promieniowania pobudzającego, przy stałym poziomie energetyki wzbudzania. Na rys. 11 pokazano przykładowo typowe charakterystyki

tego rodzaju, wyznaczone dla trzech rodzajów przetworników; na osi rzędnych wyznaczone są wartości natężenia luminescencji dla głównego pasma emisji przetwornika. Charakterystyki wzbudzenia różnią się nieco od wyznaczonych na podstawie pomiarów odbicia, a mianowicie zakres widmowy jest węższy a przebieg bardziej selektywny. Różnica wynika stąd, że charakterystyka wzbudzenia odzwierciedla tylko te procesy absorpcji i wzbudzenia, które kończą się deekscytacją promienistą.

Maksimum promieniowania typowej diody elektroluminescencyjnej z GaAs, w której rekombinacja promienista zachodzi za pośrednictwem płytkich poziomów akceptorowych (np. Zn), przypada dla około 900 nm (rys. 12, krzywa 1), dioda taka jest więc mało przydatna



Rys. 12. Typowe charakterystyki widmowe promieniowania *DEL*: 1 — z GaAs, 2 — z GaAs:Si, 3 — — z GaAs:Si o maksimum emisji pokrywającym się z maksimum charakterystyki wzbudzenia *FPP* (krzywa 4)

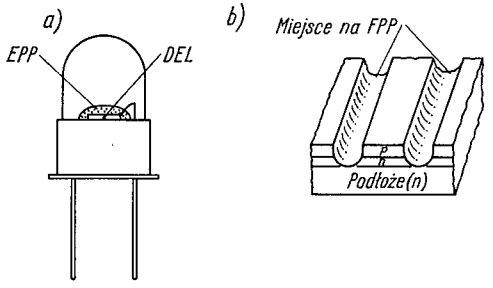
jako źródło pobudzania *FPP*. Okazuje się jednak, że przez domieszkowanie GaAs krzemem można uzyskać diody o maksimum promieniowania przesuniętym w podczerwień i jednocześnie o większej wydajności [9] [54]. Diody GaAs: Si są wykonywane w procesie epitaksji z fazy ciekłej, z zastosowaniem krzemu jako domieszki po obu stronach złącza *p-n*. Rekombinacja promienista w złączu zachodzi za pośrednictwem głębiej położonych poziomów akceptorowych. W związku z tym maksimum charakterystyki widmowej promieniowania przesuwa się w kierunku fal dłuższych, a jednocześnie maleje absorpcja promieniowania w materiale złącza i wzrasta zewnętrzna wydajność kwantowa. Według Ladany'ego [9] zmianie zawartości krzemu w GaAs w granicach od 0,2% do 1% odpowiada przesunięcie maksimum emisji od 920 nm do 1000 nm. W pracy tej otrzymano diody GaAs: Si o zewnętrznej wydajności kwantowej η_{q_z} wynoszącej średnio 6% (maksymalnie do 10%) oraz diody z osłoną wykonaną ze szkliwa o dużym współczynniku załamania, których wydajność η_{q_z} sięgała 32%.

Zazwyczaj diody z GaAs: Si mają maksimum emisji dla około 940 nm, nie zapewniają więc optymalnych warunków wzbudzenia przetwornika. Na rys. 12 (krzywa 2) pokazano charakterystykę emisji takiej diody obok charakterystyki wzbudzenia typowego przetwornika LaF_3: Yb^{3+}, Er^{3+} (krzywa 4). W tym przypadku w procesie wzbudzenia *FPP* może być użyteczne jedynie około 26% energii promieniowania diody, a w przypadku zastosowania osnowy YF_3 — około 33%.

Gdyby maksima charakterystyk emisji i wzbudzenia pokrywały się, to ze względu na różne kształty charakterystyk odpowiednie wartości wynosiłyby około 38% i 48% dla obu wymienionych osnow [6].

Ze względu na nieliniową charakterystykę $I_L(I_{IR})$ przetwornika istotne jest, aby dla źródła o danej charakterystyce widmowej natężenie promieniowania podczerwonego padającego na powierzchnię przetwornika było jak największe.

Decyduje o tym geometria układu dioda—*FPP*. W szeregu prac stosuje się konstrukcję układu taką jak pokazana na rys. 1, z warstwą *FPP* nałożoną na powierzchnię półkulistej czaszy. Zastosowanie czaszy z GaAs lub z niskotopliwego szkliwa arseno-tlenowcowo-chlorowcowego [55] o współczynniku załamania $n = 2{,}4 \div 2{,}9$ zbliżonym do współczynnika załamania dla GaAs ($n = 3{,}6$) pozwala na zwiększenie zewnętrznej wydajności kwantowej diody przez zmniejszenie strat promieniowania wywołanych przez odbicia wewnętrzne na granicy półprzewodnik-powietrze.



Rys. 13. a) Konstrukcja źródła światła typu *DEL-FPP* b) Fragment konstrukcji *DEL* dla wskaźnika cyfrowego typu *DEL-FPP* (schematycznie)

Natężenie promieniowania pobudzającego warstwę *FPP* nałożoną na powierzchnię czaszy jest jednak znacznie mniejsze od uzyskiwanego wtedy, gdy przetwornik jest umieszczony w bezpośredniej bliskości złącza i w tym ostatnim przypadku można otrzymać większą sprawność całkowitą źródła promieniowania, mimo iż wydajność samej diody będzie niższa. W realizacji takiej konstrukcji może być przydatna dioda heterozłączowa planarna, dla której uzyskano wydajność η_{q_z} około 10% [31].

Firma General Electric, która pierwsza uruchomiła produkcję półprzewodnikowych źródeł światła opartych na wykorzystaniu przetwarzania promieniowania podczerwonego na widzialne, stosuje w diodzie typu *SSL*-3 nakładanie warstwy *FPP* bezpośrednio na powierzchnię złącza (rys. 13a).

Podobną konstrukcję układu dioda-*FPP* zastosowano przy budowie siedmiosegmentowego wskaźnika cyfrowego [6]. Poszczególne segmenty wskaźnika są wykonane w sposób pokazany na rys. 13b; dla wzbudzenia świecenia segmentu wykorzystuje się emisję z krawędzi wydłużonego złącza *p-n* otoczonego z obu stron rowkami wypełnionymi materiałem typu *FPP*.

Ze względu na wykładniczy charakter zależności $I_L(I_{IR})$ przetwornika wskazane jest zasilanie diody GaAs: Si w sposób impulsowy, gdyż uzyskuje się wtedy znaczny wzrost natężenia promieniowania widzialnego przy tej samej średniej wartości natężenia prądu co przy zasilaniu prądem stałym. Na przykład dla *FPP* o barwie świecenia zielonej, zasilanie

diody impulsami prądu 200 mA przy współczynniku wypełnienia 0,1 daje około 10-krotny wzrost średniej wartości natężenia świecenia w porównaniu do świecenia uzyskiwanego przy zasilaniu prądem stałym 20 mA. Stosuje się zwykle impulsy o szerokości około 3 ms [6].

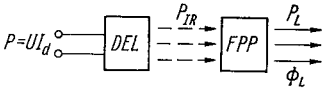
Osobnym zagadnieniem jest dobór optymalnej grubości warstwy przetwornika oraz jego uziarnienia. Warstwa *FPP* jest pobudzana od strony diody, a promieniowanie widzialne jest obserwowane od strony przeciwnej. Gdy warstwa jest zbyt cienka, jest w niej absorbowana tylko mała część energii wzbudzenia, natomiast przy warstwie zbyt grubej następuje znaczne tłumienie promieniowania widzialnego. Jeśli chodzi o wpływ uziarnienia, to większą sprawność uzyskuje się dla materiałów o większych rozmiarach krystalitów [7].

5. PODSTAWOWE PARAMETRY PRZETWORNIKÓW

Właściwości przetworników typu *FPP* oraz źródeł światła typu *DEL-FPP* są określone przez zespół parametrów energetycznych, optycznych i elektrycznych.

S p r a w n o ś ć i s k u t e c z n o ś ć ś w i e t l n a

Najważniejszym parametrem przetwornika typu *FPP* oraz układu *DEL-FPP* jako przetworników energii i źródeł promieniowania jest ich sprawność. W przetwornikach tych za-



Rys. 14. Uproszczony schemat *DEL-FPP*

chodzi przemiana energii: w *DEL* energii elektrycznej dostarczonej do złącza na energię promieniowania podczerwieni wychodzącego na zewnątrz diody, w *FPP* — energii promieniowania podczerwonego na energię promieniowania widzialnego. Sprawność może być określona za pomocą następujących wielkości, których definicje wynikają z rozpatrzenia uproszczonego schematu *DEL-FPP* (rys. 14):

— sprawność źródła promieniowania *FPP*

$$\eta_e = \frac{P_L}{P_{IR}}, \tag{10}$$

— sprawność źródła promieniowania *DEL*

$$\eta_{ed} = \frac{P_{IR}}{P}, \tag{11}$$

— sprawność całkowita źródła promieniowania *DEL-FPP*

$$\eta_{ec} = \frac{P_L}{P}. \tag{12}$$

Oczywiście

$$\eta_{ec} = \eta_e \cdot \eta_{ed}. \tag{13}$$

Przetworniki *FPP* oraz *DEL-FPP* traktowane jako źródła światła mogą być scharakteryzowane za pomocą następujących wielkości:

— skuteczność świetlna źródła światła *FPP*

$$\eta_z = \frac{\phi_L}{P_{IR}}, \tag{14}$$

— skuteczność świetlna źródła światła *DEL-FPP*

$$\eta_{zc} = \frac{\phi_L}{P}, \tag{15}$$

— skuteczność świetlna promieniowania *FPP*

$$K_L = \frac{\phi_L}{P_L}, \tag{16}$$

przy czym

$$\eta_z = \eta_e \cdot K_L \tag{17}$$

oraz

$$\eta_{zc} = \eta_z \cdot \eta_{ed}. \tag{18}$$

Skuteczność świetlna K_L jest zdefiniowana jako

$$K_L = K_m \frac{\int_0^\infty P_L(\lambda) \cdot V(\lambda) d\lambda}{\int_0^\infty P_L(\lambda) d\lambda} \left[\frac{lm}{W}\right], \tag{19}$$

gdzie:

$P_L(\lambda)$ — rozkład widmowy mocy P_L,

$V(\lambda)$ — względna światłoczułość widmowa oka normalnego obserwatora,

$K_m = 680 \frac{lm}{W}$ — fotometryczny równoważnik promieniowania.

Ponieważ przy ocenie sprawności *DEL* zazwyczaj operuje się pojęciem zewnętrznej wydajności kwantowej η_{qz}, określonej jako stosunek ilości fotonów generowanych w obszarze czynnym do ilości elektronów przekraczających złącze *p-n*, należy dodać, że sprawność η_{ed}, uwzględniająca straty mocy związane z opornością szeregową diody

$$\eta_{ed} = \eta_{qz} \frac{h\nu/e}{U}, \tag{20}$$

gdzie U — spadek napięcia na diodzie.

W przypadku przetworników *FPP* i *DEL-FPP* najczęściej operuje się wielkością sprawności źródła promieniowania η_e i η_{ec}.

Dla pierwszych tego rodzaju źródeł promieniowania opracowanych przez Galgainitisa i Fennera [19] przy użyciu *FPP* świecącego zielono (typu LaF_3: Yb^{3+}, Er^{3+}, opisanego w [20]) i diody GaAs: Si przy prądzie $I_d = 50$ mA otrzymano całkowitą sprawność $\eta_{ec} = 0{,}003\%$. Sprawność η_{ec} wzrastała liniowo w funkcji natężenia prądu diody i przy zasilaniu impulsami prądu 1 A wynosiła około 0,05%.

Opracowanie innych typów *FPP* i nowych technologii pozwoliło na uzyskanie większych sprawności. Dla *FPP* zielonego o składzie $BaYF_5: Yb^{3+}$, Er^{3+} pobudzanego diodą przy prądzie $I_d = 300$ mA η_{ec} jest rzędu 0,1%, a o składzie $YF_3: YB^{3+}$, Er^{3+} i $NaYF_4: Yb^{3+}$ Er^{3+} — rzędu 0,2÷0,3% [7] [31] [46] [48].

Geusic i współprac. opublikowali w [31] wyniki pomiarów sprawności źródeł promieniowania, które podano w tablicy 2.

Sprawność η_{ec} źródeł światła *DEL-FPP* według [31]

Tablica 2

FPP (wzór sumaryczny)	Pasmo emisji [nm]	K_L [lm/W]	η_{ec} (%)	
			dla diody o $\eta_{qz} = 10\%$	dla diody o $\eta_{qz} = 28\%$
$Y_{0,74}Yb_{0,25}Er_{0,01}OCl$	660	30	1	7,8
$Y_{0,84}Yb_{0,15}Er_{0,01}F_3$	550	660	0,11	0,86
$Y_{0,65}Yb_{0,35}Tm_{0,001}F_3$	470	60	0,011	0,086

W pomiarach zastosowano *DEL* o wydajności $\eta_{qz} = 10\%$, o maksimum promieniowania przy 930 nm, zasilaną prądem 300 mA. Na powierzchni półkulistej obudowy diody była nałożona warstwa *FPP* o grubości 0,5 mm. *FPP* wykonano techniką prażenia pod warstwą ochronną BeF_2 [24]. Największą sprawność (1%) otrzymano dla źródła o barwie świecenia czerwonej, a dla źródła o barwie świecenia zielonej i niebieskiej odpowiednio 0,11% i 0,0011%.

Jeśli chodzi o skuteczność *DEL-FPP* jako źródła światła, to ze względu na dopasowanie do maksimum czułości oka ludzkiego ($K_L = 660$ lm/W) skuteczność świetlna η_{zc} przy zastosowaniu *FPP* świecącego zielono jest wyższa niż przy zastosowaniu *FPP* świecącego czerwono i wynosi odpowiednio 0,73 lm/W oraz 0,3 lm/W. Jak dotychczas η_{zc} dla źródła o barwie świecenia niebieskiej jest znacznie niższa i wynosi 0,006 lm/W.

W tablicy 2 podano również szacunkowe wartości η_{ec}, które można by uzyskać dla *DEL-FPP* stosując *DEL* o wydajności 28%, czyli największej uzyskiwanej dotychczas w firmie Texas Instruments, nie uwzględniając przy tym efektów cieplnych mogących osłabić świecenie.

Przy zastosowaniu przetwornika typu *FPP* świecącego zielono odpowiednie wartości sprawności η_{ec} i η_{zc} wynosiłyby wtedy 0,86% i 5,6 lm/W, a przy zastosowaniu przetwornika świecącego czerwono — 7,8% i 2,3 lm/W.

Szybki wzrost wartości η_{ec} i η_{zc} przy wzroście sprawności źródła pobudzania *DEL* tłumaczy się tym, że zgodnie z mechanizmem transferu energii w *FPP*, opisanym w rozdz. 2

$$\eta_e \doteq I_{IR}^{\alpha-1}, \tag{21}$$

lub

$$\eta_e \doteq \eta_{ed}^{\alpha-1}, \tag{22}$$

gdzie α — wykładnik potęgi występujący w równaniu (9).

Dla *FPP* o barwie świecenia zielonej lub czerwonej, przy 2-fotonowym transferze energii

$$\eta_e \doteq \eta_{ed}. \tag{23}$$

Przy 3-fotonowym transferze energii (świecenie niebieskie) powinno być:

$$\eta_e \doteq \eta_{ed}^2. \tag{24}$$

Jednak dla przetwornika typu YF_3: Yb^{3+}, Tm^{3+} występuje tendencja nasycenia obsadzenia poziomu wzbudzonego Tm $^3H_4$ wzrastająca w funkcji natężenia pobudzenia i w praktyce $\alpha \cong 2$ [31].

Parametrami fotometrycznymi charakteryzującymi promieniowanie widzialne *FPP* i *DEL-FPP* są: strumień świetlny ϕ_L, światłość czyli natężenie luminescencji I_L, luminancja L oraz emitancja świetlna M.

Jak wiadomo między wielkościami tymi istnieją następujące zależności.

Dla dowolnego źródła światła światłość w danym kierunku

$$I = \frac{d\phi}{d\omega}\left[\frac{\text{lm}}{\text{sr}} = \text{cd}\right], \tag{25}$$

luminancja elementarnego pola powierzchni dA w danym kierunku

$$L = \frac{dI}{dA\cos\varepsilon}\left[\frac{\text{cd}}{\text{m}^2} = \text{nt}\right] \tag{26}$$

gdzie ε — kąt jaki tworzy kierunek obserwacji z normalną do elementu powierzchni dA, natomiast emitancja świetlna elementarnego pola powierzchni świecącej

$$M = \frac{d\phi}{dA}\left[\frac{\text{lm}}{\text{m}^2}\right]. \tag{27}$$

Dla elementu powierzchni świecącej w sposób idealnie rozproszony (zgodnie z prawem kosinusów)

$$\phi = \pi I_m, \tag{28}$$

gdzie I_m — światłość w kierunku normalnym do powierzchni oraz

$$M = \pi \cdot L. \tag{29}$$

Z równania (29) wynika, że luminancja powierzchni świecącej w sposób idealnie rozproszony może być także wyrażona w jednostkach emitancji świetlnej, na przykład 1 lambert jest to luminancja powierzchni o emitancji świetlnej $1\ \frac{\text{lm}}{\text{cm}^2}$ $\left(1\ \text{lambert} = \frac{10^4}{\pi}\ \text{nt}\right)$.

W pracach amerykańskich i angielskich jest używana jako jednostka luminancji 1 stopolambert (footlambert). Jest to luminancja powierzchni o emitancji świetlnej $1\ \frac{\text{lm}}{\text{stopa}^2}$ (1 ft-L = 3,426 nt).

Dla źródła światła *DEL-FPP* wielkość emitancji świetlnej wiąże się ze sprawnością energetyczną *DEL* oraz *FPP*, a także ze skutecznością świetlną promieniowania *FPP*

Zgodnie z przyjętymi oznaczeniami

$$M = \frac{P \cdot \eta_{ed} \cdot \eta_e \cdot K_L}{A}\left[\frac{\text{lm}}{\text{m}^2}\right]. \tag{30}$$

Dobrze świecące źródła światła *DEL-FPP* przy prądach zasilania diody rzędu 100 mA charakteryzują się zwykle luminancją rzędu 100÷300 ft-L (340÷1000 nt); są to źródła światła świecące zielono, przy $\lambda_{max} = 540$ nm.

Charakterystyka prądowa luminescencji źródła światła *DEL-FPP* jest określona na podstawie wzoru (9) przez

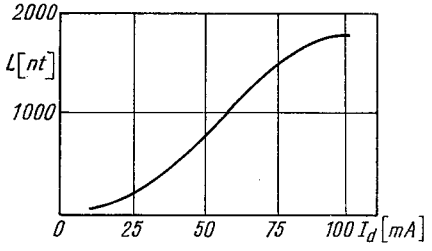
$$L = C \cdot I_d^{\alpha}, \tag{31}$$

gdzie C — stała proporcjonalności.

Wykładnik potęgi α, który w przypadku zastosowania *FPP* świecącego zielono lub czerwono powinien być równy 2, przybiera często inne wartości w zależności od rodzaju użytej osnowy. I tak na przykład:

dla YF_3: Yb^{3+}, Er^{3+}	przy $\lambda = 550$ nm	$\alpha = 2{,}0$,
	przy $\lambda = 660$ nm	$\alpha = 2{,}8$,
dla YOCl = Yb^{3+}, Er^{3+}	przy $\lambda = 550$ nm	$\alpha = 2{,}5$,
	przy $\lambda = 660$ nm	$\alpha = 2{,}0$.

Typowy przebieg charakterystyki $L(I_d)$ dla *FPP* typu LaF_3: Yb^{3+}, Er^{3+} pokazano na rys. 15. W początkowym zakresie charakterystyki luminancja L wzrasta z kwadratem natężenia prądu. W tym zakresie dla prądu 50 mA otrzymano luminancję około 200 ft-L (680 nt).



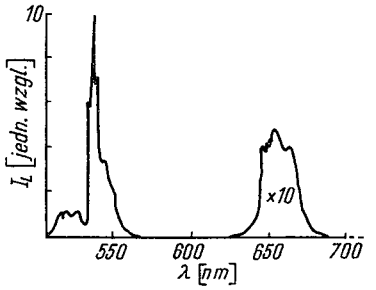
Rys. 15. Charakterystyka $L(I_d)$ układu *DEL-FPP* wg [19]

Przy większych prądach przyrosty luminancji maleją na skutek wzrostu temperatury złącza oraz warstwy *FPP*. W warunkach pracy impulsowej obszar zależności kwadratowej $L(I_d)$ rozszerza się i np. przy zasilaniu impulsami o szerokości 8 ms sięgał w omawianym przypadku do $I_d = 300$ mA.

Widmo promieniowania FPP składa się z jednego ewentualnie kilku wąskich pasm emisji. Na rys. 16 pokazano przykładowo charakterystykę widmową promieniowania przetwornika typu LaF_3: Yb^{3+}, Er^{3+}. Główne pasmo emisji jest w tym przypadku dla zieleni, w przedziale 520÷560 nm, co tłumaczy się schematem przejść Yb^{3+}—Er^{3+} wg rys. 5. Promieniowanie w czerwieni, związane z deekscytacją promienistą z poziomu Er $^2F_{9/2}$ reprezentuje około 5% całkowitej energii emitowanej przez przetwornik. Pasma emisji wykazują złożoną strukturę na skutek rozszczepienia w polu krystalicznym poziomów energetycznych aktywatora.

Ze względu na wąskopasmowy charakter emisji i dominację jednego z pasm *FPP* można traktować jako w przybliżeniu monochromatyczne źródło światła.

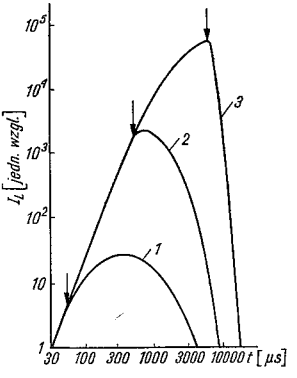
W widmie promieniowania układu *DEL-FPP* występuje poza świeceniem *FPP* także intensywne promieniowanie podczerwone *DEL*, gdyż promieniowanie to jest jedynie częściowo absorbowane przez przetwornik.



Rys. 16. Charakterystyka widmowa promieniowania dla $LaF_3:Yb^{3+}$, Er^{3+} [19]

Przebiegi czasowe świecenia FPP

Narastanie i zanik świecenia *FPP* przy włączaniu i wyłączaniu źródła podczerwieni ma charakter podobny do obserwowanego dla fotoluminoforów czy elektronoluminoforów, tzn. świecenie wzrasta w czasie dążąc do wartości ustalonej oraz zanika z opóźnieniem.



Rys. 17. Przebiegi chwilowe $I_L(t)$ dla $LaF_3:Yb^{3+}$, Er^{3+} pobudzonego impulsem o czasie trwania 50 μs (krzywa 1), 500 μs (krzywa 2) i 7 ms (krzywa 3) [52]

Przebieg narastania i zaniku świecenia uwidacznia się wyraźnie przy pobudzaniu *FPP* w krótkich przedziałach czasowych. Na rys. 17 pokazano przebiegi czasowe świecenia *FPP* typu $LaF_3: Yb^{3+}$, Er^{3+} pobudzanego przez *DEL* zasilaną impulsem prądu o szerokości 50 μs (krzywa 1), 500 μs (krzywa 2) i 7 ms (krzywa 3). Strzałkami zaznaczono koniec trwania każdego z impulsów.

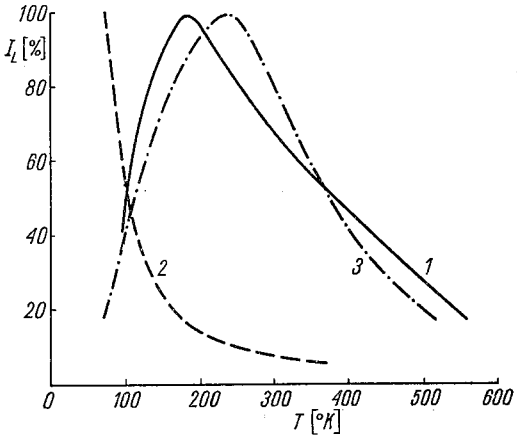
Szybkość działania *DEL* jest rzędu kilkuset ns i praktycznie nie ma wpływu na przebieg świecenia *FPP*.

Z rys. 17 wynika, że promieniowanie *FPP* narasta podczas pobudzania i dopiero po ok. 10 ms osiąga wartość ustaloną.

Jest charakterystyczne, że przy krótkotrwałym impulsie (50 μs) emisja widzialna po ustaniu pobudzenia początkowo dalej wzrasta i osiąga maksimum po około 250 μs. Przy dłuższych impulsach narastanie świecenia jest nieznaczne lub nie występuje wcale. To magazynujące działanie *FPP* przy krótkotrwałym pobudzeniu tłumaczy się długimi (rzędu ms) czasami życia stanów wzbudzonych uczulacza i aktywatora, w związku z czym transfer energii do aktywatora zachodzi także po ustaniu pobudzenia.

Zanik świecenia dla $LaF_3: Yb^{3+}$, Er^{3+} ma charakter eksponencjalny o stałej czasu ok. 1,1 ms.

Narastanie i zanik świecenia innych typów *FPP* przebiega podobnie, a stałe czasowe przebiegów są rzędu pojedynczych ms.



Rys. 18. Zależność $I_L(T)$ dla $YF_3:Yb^{3+}$, Er^{3+} (krzywa 1), $YF_3:Yb^{3+}$, Ho^{3+} (krzywa 2) i $YF_3:Yb^{3+}$, Tm^{3+} (krzywa 3) [20]

Trwałość świecenia FPP określona przez przebieg charakterystyki $L(t)$, czyli spadek luminancji L w funkcji czasu pracy, nie była dotychczas szczegółowo badana. Pewne informacje zawarte są w pracy Kano i współpr. [56], którzy podają, że dla diody GaAs: Si pokrytej warstwą *FPP* w lepiszczu poliuretanowym, zasilanej prądem 50 mA luminancja po czasie pracy 4414 h spadła do 86% wartości początkowej, natomiast dla diody pokrytej warstwą *FPP* w żywicy epoksydowej spadek po okresie czasu 3477 h wynosił 91%. W obu przypadkach spadek luminancji był spowodowany nie tylko starzeniem się przetwornika, lecz także starzeniem diody i lepiszcza.

Kano i współpr. [56] badali również *odporność FPP na działanie czynników zewnętrznych*. Mianowicie przechowywanie luminoforu typu *FPP* w atmosferze o 100% wilgotności przez przeciąg 118 dób nie spowodowało spadku luminancji przetwornika, a powtarzane na przemian gotowanie luminoforu w wodzie i suszenie w atmosferze powietrza w temperaturze 100°C (w ciągu 4 dób) spowodowało spadek luminancji do 88% wartości początkowej.

Zależność parametrów FPP od temperatury była badana w pracach [20] i [47]; typowe przebiegi $I_L(T)$ przedstawia rys. 18. Powyżej 300°K przy wzroście temperatury następuje spadek natężenia emisji w związku ze wzrostem liczby przejść niepromienistych wywołanych obsadzeniem poziomów pośrednich w wyniku oddziaływania aktywatora z polem fononów. Przy spadku temperatury obserwuje się wzrost natężenia emisji, przy czym dla *FPP* z aktywatorami Er^{3+} i Tm^3 występuje w zakresie temperatur $200 \div 250$°K maksimum charakterystyki $I_L(T)$, co tłumaczyć można udziałem fononów w transferze energii. Na przykład w $YF_3: Yb^{3+}$, Er^{3+} poziom Er $^4I_{11/2}$ w temperaturze 4°K leży nieco wyżej poziomu Yb $^2F_{5/2}$. Przy wzroście temperatury w wyniku oddziaływania uczulacza z fononami jest możliwy coraz wydajniejszy transport energii wzbudzenia do aktywatora i wzrasta natężenie promieniowania zielonego.

6. PODSUMOWANIE DOTYCHCZASOWYCH WYNIKÓW

Wnioski

Powyżej opisano technologię i podstawowe właściwości przetworników transformujących promieniowanie podczerwone diody GaAs: Si na światło widzialne. Według aktualnych danych opartych na informacji firm zaawansowanych w tej technice (General Electric, Bell Telephone Laboratories, Texas Instruments) najbardziej wydajne przetworniki to:

dla barwy zielonej $BaYF_5: Yb^{3+}, Er^{3+}$ (Bell),
$YF_3: Yb^{3+}, Er^{3+}$ (T.I.),
$LaF_3: Yb^{3+}, Er^{3+}$ (G.E.),
dla barwy czerwonej $YOCl: Yb^{3+}, Er^{3+}$ (Bell),
$Y_3OCl_7: Yb^{3+}, Er^{3+}$.

Sprawność η_e przetworników wynosi maksymalnie dla promieniowania czerwonego do 10%, a dla promieniowania zielonego $1 \div 3$%, a sprawność η_{ec} źródła promieniowania *DEL-FPP* wykonanego na diodzie o wydajności 10% wynosi odpowiednio 1% i $0{,}1 \div 0{,}3$%.

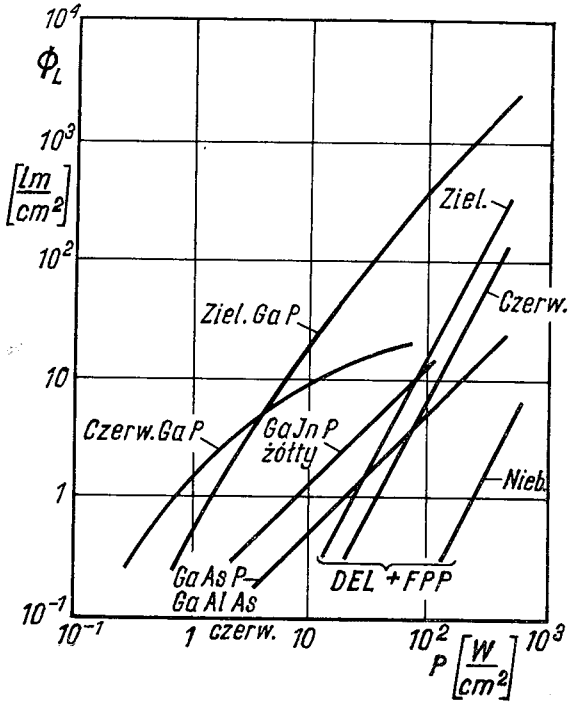
Szczególnie interesujące są źródła światła o barwie zielonej, gdyż skuteczność świetlna K_L promieniowania jest wtedy maksymalna i wynosi około 660 lm/W. Najlepsze „zielone" źródła światła *DEL-FPP* mają skuteczność świetlną $\eta_{zc} = 0{,}6 \div 2$ lm/W. Trudność polega na tym, że ze względu na wykładniczy przebieg charakterystyki $I_L(I_{IR})$ te znaczne sprawności uzyskuje się przy dużych gęstościach prądu diody, rzędu 300 A/cm^2.

Realne są szanse zwiększenia sprawności *FPP*, a przez to zmniejszenie prądu I_d. Prowadzone są prace nad optymalizacją zawartości uczulacza i aktywatora [57], a także poszukiwania nowych rodzajów osnowy, dla których transfer energii z Yb^{3+} do Er^{3+} zachodziłby w korzystniejszych niż dotychczas warunkach. Na przykład dużą sprawność η_{ec} (do 2,8%) uzyskano dla osnowy $NaYF_4$ [46] [56], dla której występuje także znaczna odporność na działanie czynników zewnętrznych.

Sprawność całkowitą η_{ec} układu *DEL-FPP* można także poprawić przez
— zastosowanie *DEL* o większej wydajności,
— lepsze dopasowanie charakterystyk widmowych absorpcji *FPP* i promieniowania *DEL*,
— optymalizację konstrukcji układu *DEL-FPP*.

Przypuszcza się, że na tej drodze można uzyskać wzrost sprawności nawet o 1 rząd wielkości [7]. Ponadto sprawność zależy od sposobu zasilania diody; preferowane jest zasilanie impulsowe, przy którym uzyskuje się większe chwilowe gęstości mocy pobudzenia.

Na tle tych badań interesująca jest ocena możliwości jakie stwarza opracowanie *FPP* dla praktycznego zastosowania źródeł światła typu *DEL-FPP*.



Rys. 19. Zależność $\phi_L(P)$ (na jednostkę powierzchni złącza) dla różnych typów półprzewodnikowych źródeł światła wg [1]

Na rys. 19 przedstawiono porównanie aktualnych właściwości optycznych *DEL-FPP* z innymi typami półprzewodnikowych źródeł światła, a mianowicie podano zależność maksymalnego strumienia świetlnego uzyskiwanego z jednostki powierzchni złącza od gęstości mocy zasilania. Z porównania tego wynika oczywista przewaga „zielonych" *DEL* z GaP nad wszystkimi pozostałymi typami źródeł. Jednak trudności technologiczne i koszt urządzeń do GaP powodują, że jedynie niektóre koncerny elektroniczne angażują się w prace badawcze i rozwojowe nad diodami i wskaźnikami wykonanymi z GaP.

Natomiast z rys. 19 wynika, że przy wzroście gęstości prądu powyżej 100 A/cm^2 skuteczność świetlna wskaźników *DEL-FPP* świecących zielono oraz czerwono jest tego samego rzędu co skuteczność świetlna *DEL* z GaAsP czy GaAlAs.

W ostatnich latach niektóre firmy podjęły produkcję wskaźników świetlnych typu *DEL-FPP*. W r. 1971 firma General Electric uruchomiła produkcję wskaźników świecących zielono, oznaczonych jako SSL-3 i SSL-3F. Podstawowe parametry tych wskaźników podano w tabl. 3.

Podstawowe parametry wskaźników świetlnych typu SSL-3 i SSL-3F firmy General Electric

Tablica 3

	Zakres widzialny	Podczerwień
Promieniowanie przy I_d = 100 mA	Luminancja 80 ft.L. (= 274 nt)	moc 1 mW
Czas narastania do 2/3 wart. max.	4 ms	150 μs
Czas zaniku do 1/3 wart. max.	2 ms	150 μs
Długość fali promieniowania	539 nm	940 nm
Szerokość charakterystyki widmowej (do 0,707 wart. max.)	538 ÷ 541 nm	930 ÷ 974 nm

W firmie tej opracowano również wskaźniki cyfrowe świecące zielono, o konstrukcji 7-segmentowej i wysokości cyfr 4 mm, złożone z segmentów *DEL-FPP* wykonanych według rys. 13b; z jednej płytki GaAs uzyskuje się 60 takich wskaźników.

Produkcję „zielonych" wskaźników świetlnych podjęła w r. 1972 firma japońska Matsushita Electronics.

Wraz ze wzrostem sprawności *DEL-FPP* będzie z pewnością możliwe obniżenie prądu diody ze 100 mA do 20 ÷ 50 mA przy zachowaniu lub niewielkim zwiększeniu poziomu luminancji powierzchni świecącej.

Poziom luminancji wymagany w przypadku każdego rodzaju świetlnego elementu sygnalizacyjnego czy elementu ekranu zależy od warunków oświetlenia otoczenia.

Jak wiadomo natężenie oświetlenia może wahać się od 20 ÷ 30 lx w pomieszczeniach słabo oświetlonych do 200 ÷ 250 lx w pomieszczeniach dobrze oświetlonych, a na stanowiskach pracy może wynosić około 500 lx. Jest zasadą, że luminancja wskaźnika winna być zbliżona do luminancji powierzchni otaczającej (w praktyce nieco większa), aby nie była konieczna ciągła adaptacja oka przy przejściu od obserwacji otoczenia do obserwacji wskaźnika i odwrotnie.

Na tej podstawie szacuje się, że luminancja wskaźnika świetlnego w pomieszczeniu dobrze oświetlonym winna wynosić 80 ÷ 160 nt, a na stanowisku pracy 160 ÷ 320 nt; firma Bell określa wymagania na 340 nt [58]. Taki poziom luminancji zapewnia np. wskaźnik jarzeniowy typu Nixie. Według [19] dla *FPP* świecącego zielono poziom 340 nt można uzyskać przy $I_d \cong 35$ mA, a według danych z tabl. 2 przy $I_d \cong 20$ mA.

Jeśli chodzi o perspektywy rozwoju układów *DEL-FPP*, to następnym krokiem powinno być opracowanie wskaźników świetlnych o barwie niebieskiej. W przyszłości natomiast wydaje się realna możliwość zastosowania *FPP* o trzech barwach świecenia: czerwonej, zielonej i niebieskiej do budowy na jednym rodzaju podłoża ekranów wielobarwnych.

Na zakończenie należy stwierdzić, że w dziedzinie nowoczesnych wskaźników świetlnych dla celów sygnalizacji i zobrazowania istnieje mnogość nie wykluczających się typów pracujących na różnych zasadach i mających określone zalety i wady [59] [60]. Wydaje się bezsporne, że do techniki półprzewodnikowej i układów scalonych najbardziej odpowiednie są źródła promieniowania oparte na elektroluminescencji w złączach *p-n*. Jednym z rozwiązań tego typu może być źródło światła *DEL-FPP*, które wykazuje szereg zalet w porów-

naniu do *DEL* na zakres widzialny, m.in. operuje dobrze poznanym materiałem (GaAs), zapewniając jednocześnie możliwość uzyskania wszystkich trzech podstawowych barw świecenia.

LITERATURA CYTOWANA W TEKŚCIE

1. A. A. Bergh, P. J. Dean, Light-emitting diodes, Proc. IEEE, **60**, 156—223, 1972.
2. D. Witt, Lumineszenzdioden, Elektronik, 369—374, 388, 1972.
3. M. A. Herman, Fosforek galu jako materiał elektroluminescencyjny, Prace ITE, 1970.
4. J. I. Pankove, E. A. Miller, J. E. Berkeyheiser, GaN blue light-emitting diodes, J. Luminescence, **5**, 84—86, 1972.
5. F. Auzel, Diodes electroluminescentes de toutes les couleurs, Electronique, **161**, 31—33, 1972.
6. S. V. Galginaitis, The use of up-converter phosphors in semiconductor light-emitting devices and displays, Metalurgical Trans., **2**, 754—761, 1971.
7. L. F. Johnson, H. J. Guggenheim, T. C. Rich, F. W. Ostermayer, Infrared-to-visible conversion by rare-earth ions in crystals, J. Appl. Phys, **43**, 1125—1137, 1972.
8. E. Ja. Arapowa i in., Sozdanie swietodiodow na osnowie antistoksowych ljuminoforow i połuprowodnikowych istocznikow IK izłuczenia, Optika i Spektroskopia, **32**, 435—437, 1972.
9. I. Ladany, Electroluminescence characteristics and efficiency of GaAs: Si diodes, J. Appl. Phys., **42**, 654—657, 1971.
10. R. A. Logan, H. G. White, W. Wiegmann, Efficient green electroluminescent junctions in GaP, Solid-State Electronics, **14**, 55—70, 1971.
11. N. Bloembergen, Solid-state infrared quantum counter, Phys. Rev. Lett., **2**, 84—85, 1959.
12. L. Esterowitz i in., Rare earth infrared quantum counter, Applied Optics, **7**, 2053—2070, 1968.
13. G. H. Dieke, Spectra and energy levels of rare earth ions in crystals, J. Wiley, N. York 1968.
14. D. Curie, Luminescencja fosforów krystalicznych, PWN, Warszawa 1965.
15. F. Auzel, Compteur quantique par transfert d'energie entre deux ions de terres rares dans un tungstate mixte et dans un verre, C.R. Acad. Sc. Paris, **262B**, 1016—19, 1966.
16. F. Auzel, Compteur quantique par transfert d'energie de Yb^{3+} á Tm^{3+} dans un tungstate mixte et dans verre germanate, C.R. Acad. Sc. Paris, **263B**, 819—821, 1966.
17. W. W. Owsjankin, P. P. Feofiłow, Kooperatiwnaja sensibilizacja luminescencji w kristałłach aktiwirowanych redkozemelnymi ionami, Ż.E.T.F. Pisma w Redakcju, **4**, 471—474, 1966.
18. R. A. Hewes, J. F. Sarver, Energy transfer between Yb^{3+} and Er^{3+} or Tm^{3+} in LaF_3 with application to infrared excited visible luminescence, Bull. Am. Phys. Soc., **13**, 687—8, 1968.
19. S. V. Galginaitis, G. E. Fenner, A visible light source utilizing a GaAs electroluminescent diode and a stepwise excitable phosphor, 1968, Symposium on GaAs, Dallas, USA, 131—135.
20. R. A. Hewes, J. F. Sarver, Infrared excitation processes for the visible luminescence of Er^{3+}, Ho^{3+} and Tm^{3+} in Yb^{3+} — sensitized rare earth fluorides, Phys. Rev., **182**, 427—436, 1969.
21. L. F. Johnson i in., Comments on materials for efficient infrared conversion, Appl. Phys. Letters, **15**, 48—50, 1969.
22. H. J. Guggenheim, L. F. Johnson, New fluoride compounds for efficient infrared-to-visible conversion, Appl. Phys. Letters, **15**, 51—52, 1969.
23. L. G. Van Uitert i in., Efficient infrared-to-visible conversion by rare earth oxychlorides, Appl. Phys. Letters, **15**, 53—54, 1969.
24. L. G. Van Uitert i in., Preparation of efficient infrared stimulable rare earth fluoride phosphors, Materials Res. Bull., **4**, 777—780, 1969.
25. L. G. Van Uitert i in., Infrared stimulable rare-earth oxy-halide phospors, their synthesis, properties and applications, Materials Res. Bul., **4**, 381—9, 1969.
26. R. K. Watts, Infrared to visible conversion in $LiYE_4$: Yb, Ho, J. Chem. Phys., **53**, 3552—7, 1970.

27. J. P. Wittke i in., Properties of a blue-emitting IR pumped YF_3: Yb, Tm diode, Proc. IEEE, **58**, 1283—5, 1970.
28. R. A. Heves, Multiphoton excitation and efficiency in the Yb^{3+} — R.E. (Ho^{3+}, Er^{3+}, Tm^{3+}) systems. J. Luminescence, **1—2**, 778—796, 1970.
29. J. P. van der Ziel i in., Factors controlling infrared-pumped visible emission of Yb^{3+}-Er^{3+} in the scheelites, J. Appl. Phys., **41**, 3308—15, 1970.
30. M. Tamatani i in., Infrared to visible conversion in Y_2O_2S:Yb, Er, J. Phys. Soc. Jap., **29**, 1099, 1970.
31. J. E. Geusic i in., Efficiency of red, green and blue infrared-to-visible conversion sources. J. Appl. Phys., **42**, 1958—60, 1971.
32. F. W. Ostermayer, Preparation and properties of infrared-to-visible conversion phosphors Metall. Trans., **2**, 747—755, 1971.
33. F. W. Ostermayer i in., Frequency upconversion in YF_3:Yb^{3+}, Tm^{3+}, Phys. Rev. B., **3**, 2698—2705, 1971.
34. W. W. Owsjankin, P. P. Feofiłow, Kooperatiwnaja ljuminescencja floridow baria i ittraja, Optika i Spektr., **31**, 944—8, 1971.
35. W. W. Azarow i in., Koncentracjonnoje tuszenie ljuminescencji i wchożdienie jonów RE w monokristały LaF_3, Optika i Spektr., **30**, 918—20, 1971.
36. T. Matsubara, Infrared stimulable phosphors YOF:Yb, Er, Jap. J. Appl. Phys., **10**, 1947, 1971.
37. P. N. Yocom i in., Rare-earth-doped oxysulfides for GaAs-pumped luminescent devices, Metall. Trans., **2**, 763—7, 1971.
38. J. L. Sommerdijk i in., Infrared-excited luminescence in oxidic lattices doped with Yb^{3+} and Er^{3+}, J. Luminescence, **4**, 404—416, 1971.
39. J. L. Sommerdijk, On the excitation mechanisms of the infrared-excited visible luminescence in Yb^{3+}, Er^{3+} — doped fluorides, J. Luminescence, **4**, 441—9, 1971.
40. E. Ja. Aprapowa in., Sozdanie swietodiodow na osnowie antistoksowych ljuminoforów i połuprowodnikowych istoczników IK izłuczenia, Optika i Sepktr., **32**, 435—7, 1972.
41. N. M. P. Low, A. L. Major, Methods of preparation of rare earth trifluoride up-conversion phosphors, Materials Res. Bull., **7**, 203—212, 1972.
42. R. K. Watts, H. J. Richter, Diffusion and transfer of optical excitation in YF_3:Yb, Ho, Phys. Rev. B., **6**, 1584—89, 1971.
43. J. P. Wittke i in., Y_2O_3:Yb, Er — new red-emitting infrared-excited phosphor, J. Appl. Phys., **43**, 595—600, 1972.
44. J. L. Sommerdijk i in., Influence of host lattice on the infrared-excited visible luminescence in Yb^{3+}, Er^{3+} — doped oxides, J. Luminescence, **5**, 297—307, 1972.
45. J. L. Sommerdijk, A. Bril, Visible luminescence of Yb^{3+}, Er^{3+} under IR excitation, Intern. Conf. on Lum, Leningrad 1972.
46. N. Menyuk i in., $NaYF_4$:Yb, Er — an efficient upconversion phosphor, Appl. Phys. Letters, **21**, 159—161, 1972.
47. H. Kuroda i in., Mechanism and controlling factors of infrared-to-visible conversion process, in Er^{3+} and Yb^{3+} — doped phosphors, J. Phys. Soc. Jap., **33**, 125—141, 1972.
48. S. G. Parker, R. E. Johnson, Preparation of (Y, Yb, Er) F_3 phosphors for green light emission, J. Electrochem. Soc., **119**, 610—613, 1972.
49. Y. Seki i Y. Furukawa, Effect of Dy^{3+} on the luminescence of Y_3OCl_7:Yb^{3+}, Er^{3+}, Jap. J. Appl. Phys., **10**, 891—901, 1977.
50. T. Miyakawa, D. L. Dexter, Cooperative and stepwise excitation of luminescence: trivalent rare-earth ions in Yb^{3+} — sensitized crystals, Phys. Rev. B, **1**, 70—80, 1970.
51. T. Miyakawa, D. L. Dexter, Phonon siedebands, multiphonon relaxation of excited states and phonon-assisted energy transfer between ions in solids, Phys. Rev. B, **1**, 2961—2969, 1970.
52. J. D. Kingsley, Analysis of energy transfer and infrared-to-visible conversion in LaF_3:Yb, Er, J. Appl. Phys., **41**, 175—182, 1970.
53. H. J. Richter, R. K. Watts, Probabilistic treatment of stepwise energy transfer and infrared-to-visible conversion, Phys. Rev. B, **5**, 320—3, 1972.

54. B. H. Ahn i in., The dependence of emission spectrum and efficiency of GaAs — Si diodes on the silicon concentration, J. Electrochem. Soc., **118**, 1015—16, 1971.
55. A. G. Fischer, C. J. Nuese, Highly refractive glasses to improve electroluminescent diode efficiencies, J. Electrochem. Soc., **116**, 1718—22, 1969.
56. T. Kano, H. Yamamoto, Y. Otomo, $NaLnF_4:Yb^{3+}$, Er^{3+} (Ln = Y, Gd, La): efficient green-emitting infrared-excited phosphors, J. Electrochem. Soc., **119**, 1561—64, 1972.
57. Y. Mita, E. Nagasawa, Optimization of infrared-to-visible conversion phosphors, NEC Res. Dev., **26**, 140—149, 1972.
58. A. A. Bergh, B. H. Johnson, A solid future for telephone lamps, Bell. Lab. Rec., grudz. 1969, 320—326.
59. O. Doyle, The right numeric readout: a critical choice for designers, Electronics, **44**, 65—72, 1971.
60. B. Darek, P. Dąbrowski, B. Mroziewicz, B. Pastuszka, Studium aktualnego stanu technologii półprzewodnikowych wskaźników cyfrowych, Prace ITE, 1972.

J. WOJCIECHOWSKI

PHOTOLUMINESCENT CONVERTERS OF INFRARED-TO-VISIBLE RADIATION

Summary

The paper presents an analysis of conversion process of infrared to visible radiation by Yb^{3+}-Er^{3+}, Yb^{3+}-Ho^{3+} and Yb^{3+}-Tm^{3+} ions in crystals. This conversion process is based on multiphoton energy transfer from sensitizer to activator ions and on the excitation of anti-Stokes luminescence.

The conditions for the obtaining of great power-conversion efficiency are discussed. Problems concerning the application of the converter in light sources utilizing GaAs:Si electroluminescent diodes as infrared sources are also presented.

J. WOJCIECHOWSKI

PHOTOLUMINESZENZ-INFRAROTSTRAHLUNG-UMWANDLER FÜR DIE LICHTSTRAHLUNG

Zusammenfassung

In der vorliegenden Arbeit sind die Umwandlungsprozesse der infraroten Strahlung in Lichtstrahlung in Systemen: Yb^{3+}-Er^{3+}, Yb^{3+}-Ho^{3+} bezw. Yb^{3+}-Tm^{3+} analysiert worden. Diese Prozesse basieren auf der Multiphonenergieumwandlung vom Sensibilitator zum Ionenaktivator und auf der Erregung der Anti-Stokes-Lumineszenz. Es wurden Bedingungen für große Leistungsfähigkeit des Umwandlers diskutiert. Auswertungsmöglichkeiten des Umwandlers in Halbleiter-Lichtquellen mit einer GaAs:Si-Elektrolumineszenzdiode als Infrarot-Lichtquelle sind dargestellt worden.

Е. ВОЙЦЕХОВСКИ

ФОТОЛЮМИНЕСЦЕНЦИОННЫЕ ПРЕОБРАЗОВАТЕЛИ ИНФРАКРАСНОГО ИЗЛУЧЕНИЯ

Резюме

Темой работы является новый вид преобразователей энергии инфракрасного излучения на энергию излучения в диапазоне видимого электромагнитного спектра действующий по принципу многофотонного трансфера энергии с сенсибилизатора к активатору и возбуждения антистоксовой люминесценции в ионах элементов группы лантаноидов ($Yb^{3+}-Er^{3+}$, $Yb^{3+}-Ho^{3+}$ или $Yb^{3+}-Tm^{3+}$).

Обсужден принцип действия преобразователя и продискутированы условия, соблюдение которых гарантирует получение максимального кпд. Представлены проблемы применения преобразователя в конструкции полупроводниковых источников света, в которых источником инфракрасного излучения является электролюминесценционный диод GaAs:Si.

TREŚĆ

Z. Rydzewski: Zastosowanie metody rewersyjnej do analizy nieliniowych obwodów elektrycznych . . . 3
J. Zabrodzki, S. Budkowski: Metody wyznaczania obszarów sprawności 15
O. Choreń, M. Zientalski: Metody oceny niezawodności systemów telekomunikacyjnych . . . 41
L. J. Weiss: Metoda czasowego zrównoważenia mostka w pomiarach zmian rezystancji . . 59
K. Kuźmiński: Analiza dynamiki układów drugiego rzędu, w których tłumienie jest funkcją uchybu . . . 69
K. Kuźmiński: Analiza dynamiki układów drugiego rzędu, w których tłumienie jest funkcją uchybu i jego pochodnej . . . 83
K. Kuźmiński: Analiza dynamiki układu, w którym tłumienie i częstotliwość drgań własnych są funkcją uchybu i jego pochodnej . . . 95
S. Pogorzelski: Równanie transmisji między antenami w obszarze Fresnela . . . 105
R. Nowak: Impulsowe detektory częstotliwości . . . 123
J. Sosnowski: Sterowany cyfrowo stabilizator napięciowo-prądowy . . . 143
J. J. Zieliński: Badania modelowe uderzeń pioruna w czynne kominy lub inne źródła ciepła 153
T. Niewierowicz: Model cyfrowy dielektrycznego suszenia ścianki cylindrycznej . . . 163
F. Kamiński: Funkcje strat torów niejednorodnych, złożonych z odcinków jednorodnych, o charakterystyce typu Czebyszewa i Butterwortha . . . 171
J. Łyskanowski: Niektóre zagadnienia analizy technicznej konstrukcji przekaźników 195
W. Bobrowski: Niezawodność układów zabezpieczeń przekaźnikowych zwielokrotnionych 209
J. Wojciechowski: Fotoluminescencyjne przetworniki podczerwieni . . . 219

CONTENTS — TABLE DES MATIÉRES — INHALT

Z. Rydzewski: Application of the Method for the Investigation of Nonlinear Electrical Circuits . . . 12
Application de la méthode de reversion dans l'analyse des circuits éléctriques nonlineares . . 12
Anwendung der Reversions-Methode in Untersuchung von nichtlinearen elektrischen Kreisen 12
J. Zabrodzki, S. Budkowski: Methods for Determining a Region of the Proper Operation . . . 39
Méthodes de détermination de l'espace de bonne fonctionnement . . . 39
Bestimmungsmethoden für Leistungsfähigkeitsbereiche . . . 39
O. Choreń, M. Zientalski: Methods of Reliability Measure of Communication Network 58
Bewertungsmethoden der Zuverlässigkeit von Systemen der Fernmeldetechnik . . . 58
L. J. Weiss: The Method of Time-Balancing of a Bridge in Resistance Deviations Measurements . . . 68
Methode zum Zeitausgleich einer Brücke bei Messungen von Widerstandsänderungen . . . 68
K. Kuźmiński: Analysis of Dynamic Properties of Second Order Systems, in Which Attenuation is a Function of Error . . . 80
Analyse der Dynamik von Systemen 2. Ordnung worin Dämpfung Funktion der Abweichung ist . . . 80
K. Kuźmiński: Analysis of Dynamic Properties of Second Order Systems, in Which Attenuation is a Function of Error and Its Derivative . . . 92
Analyse der Dynamik von Systemen 2. Ordnung, worin Dämpfung Funktion der Abweichung und deren Ableitung ist . . . 93
K. Kuźmiński: Analysis of Dynamic Properties of a System, in Which Attenuation and Frequency of Characteristic Oscillations Are Functions of Error and Its Derivative . . . 102

Analyse der Dynamik von Systemen 2. Ordnung, worin Dämpfung und Frequenz Funktionen der Abweichung und deren Ableitung sind 103
S. Pogorzelski: Power Transmission Between Antennas in Fresnel Region 120
Gleichung für Leistungsübertragung zwischen Antennen im Fresnelgebiet 121
R. Nowak: FM Pulse Detectors 141
FM-Impulsfrequenzdetektoren 141
J. Sosnowski: Digitaly Programmed Voltage-Current Supply 152
Stabilisateur voltage-courant avec programmation digitale 152
Zahlgesteuerter Stromspannungsstabilisator 152
J. J. Zieliński: Model Tests of Lighting Strokes to Active Chimneys or Other Heat Sources 161
Essais modellant les coups de foudre sur hautes-cheminées actives ou sur autres sources de chaleurs 161
Modelluntersuchungen von Blitzschlägen in tätige Schornsteine oder in andere Warmquellen 162
T. Niewierowicz: Digital Model of Dielectric Drying of Cylindrical Wall 170
Zahlenmodell für dielektrisches Trocken einer zylindrischen Wand 170
F. Kamiński: Loss Functions of the Stepped Lines With Chebyshev and Butterworth Transmission Characteristic 192
Fonctions de l'affaiblissement effectif des lignes inhomogènes composées de troncons de ligne homogène à caractéristique de type de Tchebysheff et de Butterworth 192
Betriebsdämpfungsfunktionen der Leitungsketten mit Tschebyscheff und Butterworth-Verhalten 193
J. Łyskanowski: Some Problems of the Technical Analysis of the Construction of Relays 206
Einige Probleme technischer Analyse der Relaiskonstruktion 207
W. Bobrowski: Reliability of Reservated Protective Relay Schemes 216
Zuverlässigkeit von Sicherheitssystemen vervielfachter Relais 216
J. Wojciechowski: Photoluminescent Converters of Infrared-to-visible Radiation . . . 250
Photolumineszenz-Infrarotstrahl-Umwandler für die Lichtstrahlung 250

СОДЕРЖАНИЕ

З. Рыдзевски: Применение реверсивного метода к исследованию нелинейных электрических цепей 13
Я. Забродзки, С. Будковски: Методы определения области работоспособности 40
О. Хорень, М. Зентальски: Методы определения надежности систем дальней связи 58
Л. Я. Вайс: Метод временной балансировки мостовой схемы в измерениях изменений сопротивления 68
К. Кузьминьски: Анализ динамических свойств систем второго порядка с затуханием являющимся функцией погрешности 81
К. Кузьминьски: Анализ динамических свойств систем второго порядка с затуханием являющимся функцией погрешности и ее производной 93
К. Кузьминьски: Анализ динамических свойств систем с затуханием и частотой собственных колебаний являющимися функциями погрешности и ее производной 103
С. Погожельски: Уравнение трансмиссии между антеннами в области Френеля . 122
Р. Новак: Импульсные детекторы чм сигналов 141
Я. Сосновски: Цифрово программированный стабилизатор напряжения и тока . . 152
Е. Я. Зелиньски: Модельные опыты ударов молнии в действующие дымоотводы или другие источники тепла 162
Т. Неверович: Цифровая модель диэлектрической сушки цилиндрической стенки . 170
Ф. Каминьски: Функции рабочего затухания ступенчатых линий с характеристиками типа Чебышева и Баттерворта 193
Я. Лыскановски: Некоторые вопросы технического анализа конструкции реле . . 207
В. Бобровски: Надежность резервированных схем релейной защиты 217
Е. Войцеховски: Фотолюминесценционные преобразователи инфракрасного излучения 250

POLSKA AKADEMIA NAUK
KOMITET ELEKTROTECHNIKI

ROZPRAWY ELEKTROTECHNICZNE

TOM XX · ZESZYT 2

KWARTALNIK

PAŃSTWOWE WYDAWNICTWO NAUKOWE
WARSZAWA 1974

ADRES REDAKCJI

00-661 Warszawa, Politechnika, Plac Jedności Robotniczej 1
Instytut Teleelektroniki, Gmach Elektroniki
pokój 343, tel. 21-007-564 oraz 25-09-10

PAŃSTWOWE WYDAWNICTWO NAUKOWE
Warszawa, Miodowa 10

Nakład 700 (575+125)	Oddano do składania 10. XII. 1974 r.
Ark. wyd. 13,75 ark. druk. 10,75	Podpisano do druku w marcu 1974 r.
Papier druk. sat. kl. III, 80 g, 70×100	Druk ukończono w kwietniu 1974 r.
Zamówienie nr 1719/73 W—99	Cena zł 40.—

Drukarnia im. Rewolucji Październikowej. Warszawa

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 2, ss. 255—266

621.372.44
621.372.5

Obliczanie obwodów elektrycznych zawierających oporniki nieliniowe

MICHAŁ TADEUSIEWICZ (ŁÓDŹ)

Instytut Podstaw Elektrotechniki Politechniki Łódzkiej

Otrzymano 3.1.1973

Opierając się na pewnych pojęciach i twierdzeniach analizy funkcjonalnej opracowano dwa warianty metody numerycznej obliczania przebiegów czasowych w rozgałęzionych obwodach rezystancyjnych zawierających liniowe i nieliniowe oporniki. Wymuszeniami w rozpatrywanych obwodach są niezależne, zmienne w czasie napięcia i prądy źródłowe.

Wykazano następnie, że w celu znalezienia odpowiedzi czasowej pewnej klasy nieliniowych obwodów *RLC* można je zastąpić w procesie obliczeniowym odpowiednim obwodem rezystancyjnym i wykorzystać do obliczeń podaną w pracy metodę.

1. WSTĘP

W ramach ogólnej teorii nieliniowych obwodów elektrycznych szczególną rolę odgrywa analiza rozgałęzionych nieliniowych obwodów rezystancyjnych. Obwody te znajdują zastosowanie w różnych układach elektronicznych, a metody ich rozwiązywania mogą być bezpośrednio wykorzystane do obliczania obwodów *RLC* w stanie ustalonym przy wymuszeniach stałych.

Analiza nieliniowych obwodów rezystancyjnych prowadzona jest dwoma torami. Pierwszy kierunek badań dotyczy problemu istnienia i jednoznaczności rozwiązania równania opisującego obwód [1], [5], [13], [17—19], [21], drugi zaś polega na poszukiwaniu efektywnych metod obliczania obwodów [2—3], [6], [9—10], [17—21].

W rezultacie wielu prowadzonych w ostatnich latach prac nad wzmiankowanymi wyżej zagadnieniami można dziś przytoczyć obszerną literaturę z tej dziedziny.

Za szczególnie przydatne do rozwiązywania obwodów nieliniowych uważane są metody numeryczne pozwalające wykorzystać do obliczeń maszyny cyfrowe stanowiące coraz powszechniejsze narzędzie pracy naukowca i inżyniera.

Znaczenie metod analizy nieliniowych obwodów rezystancyjnych polega również na tym, że mogą one być zastosowane do wyznaczania przebiegów czasowych bardziej ogólnej klasy obwodów, a mianowicie zawierających również cewki i kondensatory. Tego rodzaju podejście wymaga zastąpienia, w procesie obliczeniowym, obwodu *RLC* pewnym równoważnym obwodem rezystancyjnym. Odpowiednie metody obliczania nieliniowych obwodów rezystancyjnych pozwalają tu otrzymać rozwiązania w sposób szybki, wygodny i przejrzysty.

2. ZALEŻNOŚCI PODSTAWOWE

Na wstępie zostaną rozpatrzone obwody zawierające liniowe i nieliniowe oporniki, w których wymuszeniami są napięcia i prądy źródłowe będące ciągłymi i różniczkowalnymi funkcjami czasu.

Przyjęto, że spełnione są podane w pracy [20] założenia dotyczące:

1. charakterystyk rezystancyjnych lub konduktancyjnych oporników nieliniowych,
2. właściwości topologicznych obwodu.

Z rozważań przeprowadzonych w pracy [20] wynika, że badany obwód może być opisany za pomocą równania

$$\mathbf{F}[\mathbf{X}(t)]+\mathbf{A}\mathbf{X}(t) = \mathbf{B}(t), \tag{2.1}$$

lub dla ustalonego $t = t_k$

$$\mathbf{F}(\mathbf{X})+\mathbf{A}\mathbf{X} = \mathbf{B}(t_k), \tag{2.2}$$

przy czym

$$\mathbf{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{oraz} \quad \mathbf{B}(t_k) = \begin{bmatrix} b_1(t_k) \\ \vdots \\ b_n(t_k) \end{bmatrix}$$

można uważać za punkty n-wymiarowej przestrzeni arytmetycznej m_n, zaś

$$\mathbf{F}(\mathbf{X}) = \begin{bmatrix} f_1(x_1) \\ \vdots \\ f_n(x_n) \end{bmatrix}$$

jest operacją nieliniową określoną na elementach przestrzeni m_n, o wartościach z przestrzeni m_n; $\mathbf{A}$ jest macierzą symetryczną stopnia n.

Należy podkreślić, że wprowadzona do rozważań przestrzeń m jest liniowa, unormowana i zupełna (typu Banacha), w której normę określa wzór [8]

$$||\mathbf{X}|| = \max_{1 \leqslant i \leqslant n} |x_i|.$$

3. ROZWIĄZANIE RÓWNANIA OBWODU

W poprzednim punkcie podano opis matematyczny badanego nieliniowego obwodu rezystancyjnego. Równanie (2.1) określa stan obwodu w dowolnej chwili t, zaś równanie (2.2) opisuje obwód w pewnej ustalonej chwili t_k.

Niech t_k będzie punktem przedziału $\langle t_0, t_N \rangle$, w którym należy znaleźć rozwiązanie równania (2.1). Jeżeli $\mathbf{X}(t_k) = \mathbf{X}_k$ spełnia równanie (2.2), to należy oczekiwać, że rozwiązanie równania

$$\mathbf{F}(\mathbf{X})+\mathbf{A}\mathbf{X} = \mathbf{B}(t_{k+1}), \tag{3.1}$$

określającego stan obwodu w chwili $t_{k+1} = t_k + \Delta t$ bliskiej chwili t_k będzie nieznacznie różniło się od $\mathbf{X}_k$. Przy ustalaniu tego faktu istotną rolę odgrywa twierdzenie „o zmianie

rozwiązania przy zmianie równania" podane w pracy [14] i omówione w dodatku. Przy przyjętych założeniach rozwiązanie $\mathbf{X}(t_{k+1}) = \mathbf{X}_{k+1}$ równania (3.1), w myśl wzmiankowanego wyżej twierdzenia, może być obliczone z przybliżonej zależności

$$\mathbf{X}_{k+1} = \mathbf{X}_k + [\mathbf{F}'_{\mathbf{x}_k} + \mathbf{A}]^{-1}(\mathbf{B}(t_{k+1}) - \mathbf{B}(t_k)), \tag{3.2}$$

gdzie $\mathbf{F}'_{\mathbf{x}_k}$ jest pochodną Frecheta nieliniowego operatora w punkcie $\mathbf{X}_k$ [20].

Startując z rozwiązania $\mathbf{X}_0$ równania $\mathbf{F}(\mathbf{X}) + \mathbf{A}\mathbf{X} = \mathbf{B}(t_0)$ (otrzymanego np. za pomocą metody przedstawionej w pracy [20]) oraz dzieląc przedział $\langle t_0, t_N \rangle$ na N równych części Δt można stosując zależność (3.2) określić rozwiązanie równania (2.1) kolejno we wszystkich wyodrębnionych punktach.

Jeżeli $\mathbf{X}_k$ jest przybliżonym rozwiązaniem równania

$$\mathbf{F}(\mathbf{X}) + \mathbf{A}\mathbf{X} - B(t_k) = \mathbf{\Theta}, \tag{3.3}$$

to podstawiając $\mathbf{X} = \mathbf{X}_k$ otrzyma się po prawej stronie tego równania pewien element $\boldsymbol{\epsilon}_k$ przestrzeni m_n, różny od elementu zerowego $\mathbf{\Theta}$, będący miarą niedokładności rozwiązania $\mathbf{X}_k$. Również obliczone ze wzoru (3.2) przybliżone rozwiązanie $\mathbf{X}_{k+1}$ równania

$$\mathbf{F}(\mathbf{X}) + \mathbf{A}\mathbf{X} - \mathbf{B}(t_{k+1}) = \mathbf{\Theta}, \tag{3.4}$$

podstawione w miejsce $\mathbf{X}$ spowoduje pojawienie się po prawej stronie tego równania pewnego elementu $\boldsymbol{\epsilon}_{k+1}$. Wprowadzając oznaczenia $\Delta\mathbf{X}_k = [\mathbf{F}'_{\mathbf{x}_k} + \mathbf{A}]^{-1}(\mathbf{B}(t_{k+1}) - \mathbf{B}(t_k))$ oraz $\Delta\boldsymbol{\epsilon}_k = \boldsymbol{\epsilon}_{k+1} - \boldsymbol{\epsilon}_k$ i $\Delta t = t_{k+1} - t_k$ otrzymuje się

$$\lim_{\Delta t \to 0} \frac{\Delta\boldsymbol{\epsilon}_k}{\Delta t} = \lim_{\Delta t \to 0} \frac{\mathbf{F}(\mathbf{X}_k + \Delta\mathbf{X}_k) - \mathbf{F}(\mathbf{X}_k) + \mathbf{A}\cdot\Delta\mathbf{X}_k - \mathbf{B}(t_k + \Delta t) + \mathbf{B}(t_k)}{\Delta t} =$$

$$= \mathbf{F}'_{\mathbf{x}_k}[\mathbf{F}'_{\mathbf{x}_k} + \mathbf{A}]^{-1}\mathbf{B}'_{t_k} + \mathbf{A}[\mathbf{F}'_{\mathbf{x}_k} + \mathbf{A}]^{-1}\mathbf{B}'_{t_k} - \mathbf{B}'_{t_k} = \mathbf{\Theta}.$$

Zatem $\lim_{\Delta t \to 0} \frac{\|\Delta\boldsymbol{\epsilon}_k\|}{\Delta t} = 0$, co oznacza, że

$$\|\Delta\boldsymbol{\epsilon}_k\| = S_k(\Delta t)\cdot\Delta t, \qquad \text{przy czym} \qquad \lim_{\Delta t \to 0} S_k(\Delta t) \to 0.$$

Błąd w końcowym punkcie przedziału $\langle t_0, t_N \rangle$ może być oszacowany w następujący sposób:

$$\|\boldsymbol{\epsilon}\| \leqslant \sum_{k=1}^{N} \|\Delta\boldsymbol{\epsilon}_k\| = \sum_{k=1}^{N} S_k(\Delta t)\cdot\Delta t \leqslant \max_k S_k(\Delta t)\cdot N\cdot\Delta t =$$

$$= \max_k S_k(\Delta t)\frac{t_N - t_0}{\Delta t}\cdot\Delta t = (t_N - t_0)\max_k S_k(\Delta t).$$

Wobec tego $\|\boldsymbol{\epsilon}\| \to 0$, gdy $\Delta t \to 0$, co wskazuje na możliwość otrzymania rozwiązania z dowolnym stopniem dokładności pod warunkiem, że odstęp Δt jest dostatecznie mały. W przypadku, gdy przyrost Δt nie jest wystarczająco mały należy liczyć się z możliwością kumulacji błędów i małą dokładnością rozwiązania w punktach końcowych.

Obecnie zostanie przedstawiony inny wariant metody numerycznej dający na ogół dokładniejsze rozwiązanie.

Na podstawie przeprowadzonych wyżej rozważań można stwierdzić, że jeżeli $\mathbf{X}_k$ jest rozwiązaniem równania (2.2), to rozwiązanie $\mathbf{X}_{k+1}$ równania (3.1) dla t_{k+1} bliskiego t_k

będzie mało różniło się od $\mathbf{X}_k$ i może być obliczone z przybliżonej zależności (3.2). Podobnie w przypadku, gdy t_{k-1} nieznacznie różni się od t_k można ze wzoru

$$\mathbf{X}_{k-1} = \mathbf{X}_k + [\mathbf{F}'_{\mathbf{x}_k} + \mathbf{A}]^{-1}(\mathbf{B}(t_{k-1}) - \mathbf{B}(t_k)) \tag{3.5}$$

znaleźć przybliżone rozwiązanie równania

$$\mathbf{F}(\mathbf{X}) + \mathbf{A}\mathbf{X} - \mathbf{B}(t_{k-1}) = \mathbf{\Theta}. \tag{3.6}$$

Z zależności (3.2) i (3.5) wynika następujący związek istniejący pomiędzy $\mathbf{X}_{k+1}$, $\mathbf{X}_k$ oraz $\mathbf{X}_{k-1}$:

$$\mathbf{X}_{k+1} = \mathbf{X}_{k-1} + [\mathbf{F}'_{\mathbf{x}_k} + \mathbf{A}]^{-1}(\mathbf{B}(t_{k+1}) - \mathbf{B}(t_{k-1})). \tag{3.7}$$

Wzór (3.7) określa rozwiązanie $\mathbf{X}_{k+1}$ równania (3.1) w zależności od rozwiązania $\mathbf{X}_k$ równania (2.2) oraz rozwiązania $\mathbf{X}_{k-1}$ równania (3.6) przy założeniu, że odstęp $\Delta t = t_{k+1} - t_k = t_k - t_{k-1}$ jest dostatecznie mały. W przypadku, gdy $\Delta t \to 0$, $\mathbf{F}'_{\mathbf{x}_k} \to \mathbf{F}'_{\mathbf{x}_{k-1}}$ i wzór (3.7) przyjmie postać (3.2).

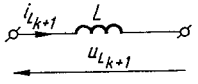
Przedstawione dwa warianty metody numerycznej rozwiązania równania (2.1) wynikające z zastosowania wzorów (3.2) i (3.7) do obliczania przebiegów czasowych w kolejnych punktach przedziału $\langle t_0, t_N \rangle$ charakteryzują się prostym algorytmem, co w wielu przypadkach pozwala na posługiwanie się w procesie obliczeniowym ogólnie dostępnymi środkami, takimi jak arytmometr lub nawet suwak logarytmiczny. W przypadkach bardziej skomplikowanych obliczenia należy wykonywać przy użyciu maszyn cyfrowych.

4. ANALIZA OBWODÓW ZAWIERAJĄCYCH LINIOWE I NIELINIOWE OPORNIKI ORAZ LINIOWE CEWKI I KONDENSATORY

Istnieje szereg metod analitycznych, graficznych i graficzno-analitycznych analizy nieliniowych obwodów *RLC*. Wiele z nich omówiono w pracach [4], [7], [11—12], [15—16].

Niżej zostanie wykazane, że znalezienie odpowiedzi czasowej badanej w niniejszej pracy klasy nieliniowach obwodów *RCL* może być sprowadzone do obliczenia odpowiedniego zastępczego obwodu rezystancyjnego.

Na wstępie zostaną rozpatrzone liniowe cewki i kondensatory stanowiące elementy omawianego obwodu.



Rys. 1. Cewka liniowa

W celu określenia napięcia na cewce w pewnej chwili t_{k+1} skorzystano ze znanego wzoru

$$u_L = L\frac{di_L}{dt} \tag{4.1}$$

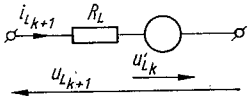
i zastąpiono w nim pochodną ilorazem różnicowym otrzymując w rezultacie przybliżoną zależność

$$u_{L_{k+1}} = L\frac{\Delta i_L}{\Delta t}, \tag{4.2}$$

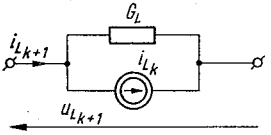
gdzie $\Delta i_L = i_{L_{k+1}} - i_{L_k}$, zaś $\Delta t = t_{k+1} - t_k$ jest małym przyrostem. Po wprowadzeniu oznaczeń $\frac{L}{\Delta t} = R_L$ oraz $\frac{L}{\Delta t} i_{L_k} = u'_{L_k}$ wzór (4.2) przyjmie postać

$$u_{L_{k+1}} = R_L i_{L_{k+1}} - u'_{L_k}. \tag{4.3}$$

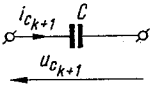
Z zależności (4.3) wynika, że cewka, w której płynie w chwili t_{k+1} prąd $i_{L_{k+1}}$ i na zaciskach której panuje napięcie $u_{L_{k+1}}$ (rys. 1) może być zastąpiona schematem równoważnym o postaci napięciowego źródła energii, przedstawionym na rys. 2.



Rys. 2. Napięciowy schemat zastępczy cewki liniowej



Rys. 3. Prądowy schemat zastępczy cewki liniowej



Rys. 4. Kondensator liniowy

Przyjmując oznaczenie $G_L = \frac{\Delta t}{L}$ można ze wzoru (4.2) otrzymać zależność

$$i_{L_{k+1}} = i_{L_k} + G_L u_{L_{k+1}}, \tag{4.4}$$

której odpowiada schemat zastępczy o postaci prądowego źródła energii, pokazany na rys. 3.

Łatwo zauważyć, że obydwa schematy są sobie równoważne i wynikają jeden z drugiego na podstawie ogólnych zasad dotyczących zamiany źródła napięciowego na prądowe i odwrotnie.

W podobny sposób może być rozpatrzony kondensator liniowy (rys. 4). W tym celu znany wzór $i_C = C\frac{du_C}{dt}$ zostanie zastąpiony przybliżoną zależnością pozwalającą określić prąd kondensatora w chwili t_{k+1}

$$i_{C_{k+1}} = C\frac{\Delta u_C}{\Delta t}, \tag{4.5}$$

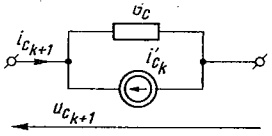
gdzie $\Delta u_C = u_{C_{k+1}} - u_{C_k}$, zaś $\Delta t = t_{k+1} - t_k$ jest niewielkim przyrostem. Wprowadzając oznaczenia $G_C = \frac{C}{\Delta t}$, $R_C = G_C^{-1}$ oraz $\frac{C}{\Delta t} u_{C_k} = i'_{C_k}$ otrzymuje się dwie równoważne zależności

$$i_{C_{k+1}} = G_C u_{C_{k+1}} - i'_{C_k} \tag{4.6}$$

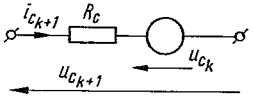
oraz

$$u_{C_{k+1}} = u_{C_k} + R_C i_{C_{k+1}}, \tag{4.7}$$

którym odpowiadają schematy zastępcze przedstawione na rys. 5 i 6.



Rys. 5. Prądowy schemat zastępczy kondensatora liniowego



Rys. 6. Napięciowy schemat zastępczy kondensatora liniowego

Podobnie jak to miało miejsce dla cewki liniowej, również i w rozpatrywanym przypadku schematy z rys. 5 i 6 są równoważne i mogą być otrzymane jeden z drugiego na podstawie ogólnych zasad dotyczących zamiany źródła napięciowego na prądowe i odwrotnie.

Rozpatrując zależności napięciowo-prądowe dla cewki korzystano z przybliżonego wzoru (4.2). W celu uzyskania większej dokładności można we wzorze (4.2) przyjąć średnią arytmetyczną napięć na krańcach przedziału $\langle t_k, t_{k+1} \rangle$ zamiast napięcia $u_{L_{k+1}}$. Otrzymuje się wówczas

$$\frac{1}{2}(u_{L_k} + u_{L_{k+1}}) = L \frac{\Delta i_L}{\Delta t}. \tag{4.8}$$

Podstawiając $G''_L = \frac{\Delta t}{2L}$ oraz $i''_{L_k} = i_{L_k} + G''_L u_{L_k}$ znajduje się

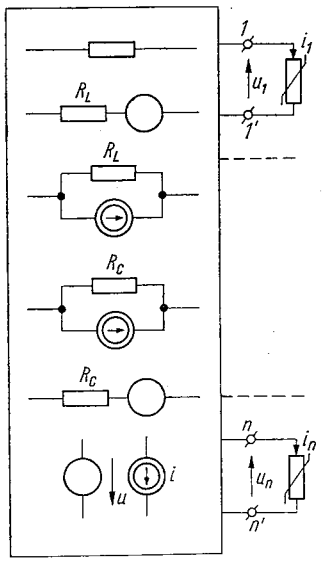
$$i_{L_{k+1}} = G''_L u_{L_{k+1}} + i''_{L_k} \tag{4.9}$$

lub

$$u_{L_{k+1}} = R''_L i_{L_{k+1}} - u''_{L_k}, \tag{4.10}$$

gdzie $R''_L = (G''_L)^{-1}$ oraz $u''_{L_k} = i''_{L_k} R''_L$.

Na podstawie zależności (4.9) i (4.10) można zbudować schematy zastępcze cewki liniowej o postaci prądowego lub napięciowego źródła energii.



Rys. 7. Obwód *RLC* po zastąpieniu cewek i kondensatorów ich schematami zastępczymi

Podobnie można postąpić w przypadku kondensatora liniowego przyjmując we wzorze (4.5) średnią arytmetyczną prądów na krańcach przedziału $\langle t_k, t_{k+1}\rangle$ zamiast prądu $i_{C_{k+1}}$. Otrzymuje się wówczas bardziej dokładne zależności

$$i_{C_{k+1}} = G''_C u_{C_{k+1}} - i''_{C_k} \tag{4.11}$$

lub

$$u_{C_{k+1}} = R''_C i_{C_{k+1}} + u''_{C_k}, \tag{4.12}$$

gdzie $R''_C = \frac{\Delta t}{2C}$, $G''_C = \frac{1}{R''_C}$, $u''_{C_k} = u_{C_k} + R''_C i_{C_k}$ oraz $i''_{C_k} = G''_C u''_{C_k}$, na podstawie których można zbudować odpowiednie schematy zastępcze kondensatora liniowego.

W celu uzyskania większej przejrzystości w dalszych rozważaniach będą wykorzystane prostsze zależności (4.2) i (4.5) oraz wynikające z nich schematy zastępcze cewki i kondensatora.

Z przeprowadzonych rozważań wynika, że obwód złożony z liniowych cewek i kondensatorów, liniowych i nieliniowych oporników oraz niezależnych źródeł napięciowych i prądowych może być zastąpiony obwodem rezystancyjnym w wyniku umieszczenia w miejscu cewek i kondensatorów ich schematów równoważnych. W obwodzie tym oprócz niezależnych wymuszeń pojawią się dodatkowe napięcia lub prądy źródłowe zmienne w czasie i zależne od stanu obwodu w chwili wcześniejszej o Δt od rozpatrywanej.

Wyodrębniając z otrzymanego w ten sposób obwodu rezystancyjnego wszystkie oporniki nieliniowe i umieszczając pozostałą jego część wewnątrz prostokąta otrzymuje się połączenie przedstawione na rys. 7.

Jeżeli obwód rezystancyjny z rys. 7 spełnia założenia podane w p. 2, to jego stan w dowolnej chwili t opisuje równanie o postaci (2.1). Należy jednak podkreślić, że w rozpatry-

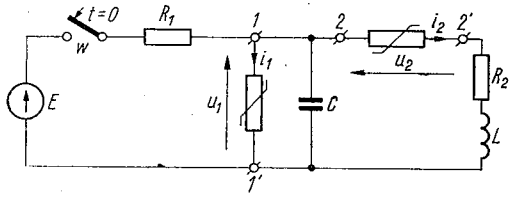
wanym przypadku elementy macierzy $\mathbf{B}(t)$ są sumą dwóch składników, z których pierwszy pochodzi od istniejących w obwodzie wyjściowym wymuszeń, drugi zaś od napięć i prądów źródłowych wprowadzonych do obwodu w wyniku zastąpienia cewek i kondensatorów ich schematami równoważnymi.

Wynika stąd, że w procesie obliczeniowym, który może być realizowany za pomocą metod omówionych w p. 3, należy dla każdej wybranej chwili, po znalezieniu napięć lub prądów w nieliniowych opornikach, wyznaczyć te wielkości, które determinują wprowadzone do obwodu napięcia i prądy źródłowe, to jest prądy w cewkach i napięcia na kondensatorach. Mogą one być określone w wyniku rozwiązania liniowego obwodu rezystancyjnego otrzymanego z połączenia przedstawionego na rys. 7 po zastosowaniu twierdzenia o kompensacji i umieszczeniu w miejscu oporników nieliniowych odpowiednich źródeł napięcia lub prądu.

5. PRZYKŁAD LICZBOWY

Zostanie rozpatrzony stan nieustalony w obwodzie przedstawionym na rys. 8 przy zerowych warunkach początkowych.

W chwili $t = 0$ nastąpiło zaburzenie w obwodzie w wyniku zamknięcia wyłącznika w. Należy obliczyć przebiegi czasowe napięć na elementach nieliniowych w przedziale $\langle 0; 4{,}5\rangle$ ms w punktach oddalonych od siebie o $\Delta t = 0{,}5$ ms.



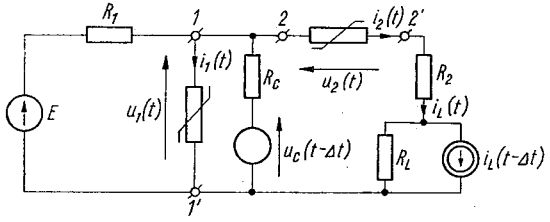
Rys. 8. Obwód zawierający dwa oporniki nieliniowe

Przyjęto, że $E = 150$ V, $R_1 = 4$ kΩ, $R_2 = 9{,}6$ kΩ, $L = 5$ H, $C = 0{,}5$ μF, zaś oporniki nieliniowe są opisane za pomocą charakterystyk konduktancyjnych:

1 1′: $i = 5 \cdot 10^{-3}u + 7{,}5 \cdot 10^{-5}u^3$,

2 2′: $i = 9{,}5 \operatorname{sh} 0{,}2u$,

przy czym prąd i jest wyrażony w mA, a napięcie u w woltach.

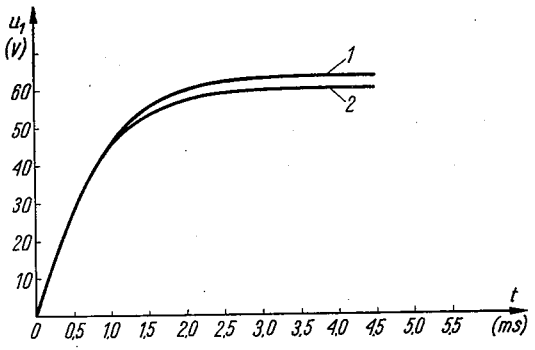


Rys. 9. Obliczeniowy schemat zastępczy obwodu z rys. 8

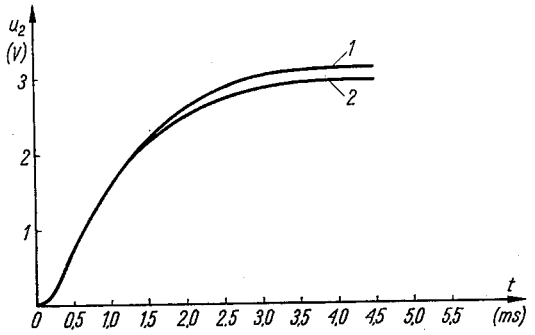
Umieszczając w miejscu cewki i kondensatora ich schematy zastępcze otrzymuje się nieliniowy obwód rezystancyjny przedstawiony na rys. 9, w którym $R_L = \frac{L}{\Delta t} = 10\ \text{k}\Omega$ oraz $R_C = \frac{\Delta t}{C} = 1\ \text{k}\Omega$.

Łatwo sprawdzić, że równanie opisujące obwód z rys. 9 w chwili $t_{k+1} = t_k + \Delta t$ jest następujące:

$$\begin{bmatrix} 5\cdot 10^{-3}(u_1)_{k+1} + 7{,}5\cdot 10^{-5}(u_1)^3_{k+1} \\ 9{,}5\ \text{sh}\ 0{,}2(u_2)_{k+1} \end{bmatrix} + \begin{bmatrix} 1{,}301 & -0{,}051 \\ -0{,}051 & 0{,}051 \end{bmatrix} \cdot \begin{bmatrix} (u_1)_{k+1} \\ (u_2)_{k+1} \end{bmatrix} = \begin{bmatrix} 37{,}50 + (u_1)_k - 4{,}845\ \text{sh}\ 0{,}2\ (u_2)_k \\ 4{,}845\ \text{sh}\ 0{,}2(u_2)_k \end{bmatrix}. \tag{5.1}$$



Rys. 10. Wykresy napięcia u_1 w zależności od czasu w obwodzie z rys. 8, otrzymane dwiema metodami



Rys. 11. Wykresy napięcia u_2 w zależności od czasu w obwodzie z rys. 8, otrzymane dwiema metodami

Obliczenia wykonano na arytmometrze wykorzystując obydwa warianty podanej w pracy metody. Odpowiednie przebiegi czasowe przedstawione zostały na rys. 10 i 11, przy czym krzywe *1* otrzymano na podstawie obliczeń prowadzonych za pomocą algorytmu wynikającego ze wzoru (3.2) ,zaś krzywe *2* powstały w rezultacie obliczeń dokonywanych na podstawie wzoru (3.7).

Gdyby w każdym punkcie obliczeniowym uściślać rozwiązanie (np. za pomocą metody podanej w pracy [20]), wówczas otrzymane przebiegi byłyby bardzo zbliżone do krzywych *2*.

DODATEK

Niżej przytoczono twierdzenie o zmianie rozwiązania przy zmianie równania zaczerpnięte z pracy [14].

Niech $D_{U,E}$ będzie przestrzenią funkcji $\mathbf{P}(\mathbf{X})$, określonych w pewnym obszarze U przestrzeni E i przyjmujących wartości z tejże przestrzeni: $\mathbf{X} \in U$, $\mathbf{P}(\mathbf{X}) \in E$. Założono, że funkcje $\mathbf{P}(\mathbf{X})$ są różniczkowalne, a funkcje $\mathbf{P}(\mathbf{X})$ i $\mathbf{P}'(\mathbf{X})$ są ciągłe i ograniczone w sensie normy.

Niech $||\mathbf{P}|| = \sup (||\mathbf{P}(\mathbf{X})|| + ||\mathbf{P}'(\mathbf{X})||)$.

Przestrzeń $D_{U,E}$ jest przestrzenią liniową, unormowaną.

Rozpatrzone zostanie równanie

$$\mathbf{P}(\mathbf{X}) = \mathbf{\Theta}; \quad \mathbf{X} \in U, \quad \mathbf{P} \in D_{U,E}.$$

Założono, że $\mathbf{P}(\mathbf{X}_k) = \mathbf{\Theta}$ dla pewnego $\mathbf{X}_k \in U$ oraz że operacja $\mathbf{P}'_{\mathbf{x}_k} \in (E \to E)$ ma operację odwrotną. Przy tych założeniach ma miejsce następujące twierdzenie.

Twierdzenie 1. Istnieją takie stałe $\tau > 0$, $\tau_1 > 0$, że dla dowolnego elementu $\mathbf{P}_1 \in D_{U,E}$ z otoczenia $||\mathbf{P}_1 - \mathbf{P}|| \leqslant \tau$ równanie $\mathbf{P}_1(\mathbf{X}) = \mathbf{\Theta}$ ma rozwiązanie $\mathbf{X} = \mathbf{X}_k + \Delta\mathbf{X}$, gdzie $||\Delta\mathbf{X}|| \leqslant \tau_1$; jeżeli $\mathbf{P}_1 \to \mathbf{P}$, to $\Delta\mathbf{X} \to \mathbf{\Theta}$.

Przyrost $\Delta\mathbf{X}$ może być obliczony z przybliżonej zależności

$$\Delta\mathbf{X} = -[\mathbf{P}'_{\mathbf{x}_k}]^{-1} \cdot \Delta\mathbf{P}(\mathbf{X}_k), \tag{d.1}$$

gdzie $\Delta\mathbf{P} = \mathbf{P}_1 - \mathbf{P}$.

Biorąc pod uwagę funkcję $\mathbf{P}(\mathbf{X}) = \mathbf{F}(\mathbf{X}) + \mathbf{A}\mathbf{X} - \mathbf{B}(t_k)$ można łatwo sprawdzić, że założenia twierdzenia 1 są spełnione. W szczególności $\mathbf{P}'_{\mathbf{x}_k} = [\mathbf{F}'_{\mathbf{x}_k} + \mathbf{A}]$ jest operacją liniową o postaci macierzy kwadratowej stopnia n, przy czym dla każdego $\mathbf{X}_k$ istnieje odwrotność tej macierzy [20]. Przyjmując $\mathbf{P}_1(\mathbf{X}) = \mathbf{F}(\mathbf{X}) + \mathbf{A}\mathbf{X} - \mathbf{B}(t_{k+1})$ i zakładając, że przyrost $\Delta t = t_{k+1} - t_k$ jest niewielki można obliczyć rozwiązanie $\mathbf{X}_{k+1}$ równania $\mathbf{P}_1(\mathbf{X}) = \mathbf{\Theta}$ ze wzoru $\mathbf{X}_{k+1} = \mathbf{X}_k + \Delta\mathbf{X}$, gdzie $\Delta\mathbf{X}$ dane jest zależnością (d.1), która w rozpatrywanym przypadku przyjmuje postać

$$\Delta\mathbf{X} = [\mathbf{F}'_{\mathbf{x}_k} + \mathbf{A}]^{-1}(\mathbf{B}(t_{k+1}) - \mathbf{B}(t_k)). \tag{d.2}$$

LITERATURA CYTOWANA W TEKŚCIE

1. G. Birkhoff, J.B. Diaz, Non-linear Network Problems, Quarterly of Applied Mathematics, nr 4, 1956.
2. M.A. Bondar, A.A. Łanne, Ob analizje odnogo klassa nieliniejnych rezistiwnych cepiej, Teoreticzeskaja Elektrotechnika, wypusk 11, Lwów 1971.
3. W.M. Bondarenko, Woprosy analiza nieliniejnych elektriczeskich cepiej, Naukowa dumka, Kijew 1967.
4. T. Cholewicki, Elektrotechnika teoretyczna, tom II, WNT, Warszawa 1971.
5. Ch. Desoer, J. Katzenelson, Nonlinear RLC Networks, BSTJ, nr 1, 1965.
6. A. Gerscho, Solving Nonlinear Network Equations Using Optimization Techniques, BSTJ, nr 9, 1969.
7. T. Kaczorek, Obliczanie rozgałęzionych obwodów nieliniowych przy użyciu maszyn cyfrowych, Archiwum Elektrotechniki, nr 4, 1963.
8. L.W. Kantorowicz, G.J. Akiłow, Funkcjonalnyj analiz w normirowannych prostranstwach, Fiz-mat-giz, Moskwa 1959.

9. J. Katzenelson, An Algorithm for Solving Nonlinear Resistor Networks, BSTJ, nr 8, 1965.
10. J. Katzenelson, L. H. Seitelman, An Iterative Method for Solution of Networks of Nonlinear Monotone Resistors, IEEE Transactions on CT, nr 2, 1966.
11. J. Kudrewicz, O przybliżeniu układu nieliniowego przez liniowy, Archiwum Elektrotechniki, nr 3, 1962.
12. J. Kudrewicz, O pewnej metodzie obliczania układów parametrycznych, Archiwum Elektrotechniki, nr 1, 1963.
13. E. S. Kuh, J. N. Hajj, Nonlinear Circuit Theory: Resistive Networks, Proceedings of the IEEE, nr 3, 1971.
14. L. A. Lusternik, W. I. Sobolew, Elementy analizy funkcjonalnej, PWN, Warszawa 1959.
15. P. Nowacki, Obliczanie nieliniowych obwodów elektrycznych i magnetycznych, PWN, Warszawa 1959.
16. I. W. Sandberg, On Truncation Techniques in the Approximate Analysis of Periodically Time-Varying Nonlinear Networks, IEEE Transactions on Circuit Theory, nr 6, 1964.
17. I. W. Sandberg, A. N. Willson, Some Theorems on Properties of DC Equations of Nonlinear Networks, BSTJ, nr 1, 1969.
18. I. W. Sandberg, A. N. Willson, Some Network-Theoretic Properties of Nonlinear DC Transistor Networks, BSTJ, nr 5, 1969.
19. I. W. Sandberg, Theorems on the Analysis of Nonlinear Transistor Networks, BSTJ, nr 1, 1970.
20. M. Tadeusiewicz, Analiza pewnej klasy nieliniowych obwodów rezystancyjnych w przestrzeni m_n, Rozprawy Elektrotechniczne, z. 2. 1973.
21. A. N. Willson, On the Solutions of Equations for Nonlinear Resistive Networks, BSTJ, nr 8, 1968.

M. TADEUSIEWICZ

CALCULATION OF ELECTRICAL NETWORKS WHICH INCLUDE NON-LINEAR RESISTORS

Summary

Basing on certain conceptions and theorems existing in the functional analysis there has been developed a numerical method for solving branched electrical networks composed of linear and non-linear resistors which are excited by time dependent voltage and current sources.

Furthermore, it has been proved that the presented method may be employed for determining transient response of a certain class of *RLC*-type non-linear networks after they are replaced first by adequate calculational resistor networks.

M. TADEUSIEWICZ

BERECHNUNG VON ELEKTRISCHEN KREISEN MIT NICHTLINEAREN WIDERSTÄNDEN

Zusammenfassung

In Anlehnung an gewisse Begriffe und Theoreme der Funktions analyse wurden zwei Varianten einer numerischen Berechnungsmethode von zeitlichen Vorgängen in verzweigten Resistanzkreisen mit linearen und nichtlinearen Widerständen bearbeitet. Erzwungene Werte sind in den in Rede stehenden Kreisen die während der Spannung veränderlichen Unabhängigen sowie Urströme.

Es wurde auch nachgewiesen, daß zwecks Ermittlung einer zeitlichen Antwort für eine gewisse Klasse nichtlinearer *RLC*-Kreise diese in den Berechnungsprozessen durch einen entsprechenden Resisstanzkreis vertreten werden können; sie können dann mittels der hier angeführten Methode zur Berechnung ausgenutzt werden.

М. ТАДЕУСЕВИЧ

РАСЧЕТ ЭЛЕКТРИЧЕСКИХ ЦЕПЕЙ С НЕЛИНЕЙНЫМИ РЕЗИСТОРАМИ

Р е з ю м е

Исходя из некоторых элементов функционального анализа разработаны два варианта численного метода решения разветвленных электрических цепей с линейными и нелинейными резисторами. Вынуждениями в этих цепях являются независимые, изменяющиеся во времени напряжения и токи источников.

Кроме того доказано, что метод пригоден также к решениям определенного класса нелинейных цепей *RLC* после замены их эквивалентной вычислительной резистивной цепью.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 2, ss. 267—294

621.372.57
621.374.5

Analiza warunków pracy tranzystorowego układu opóźniającego o dużych czasach opóźnienia

HENRYK MROCZEK (ŁÓDŹ)

Instytut Automatyki Politechniki Łódzkiej

Otrzymano 20,2.1973

W pracy przeprowadzono analizę układu opóźniającego przystosowanego do realizacji dużych opóźnień, złożonego z obwodu *RC*, pomocniczego źródła impulsów i tranzystorowego przerzutnika bistabilnego o szeregowo-równoległym sprzężeniu zwrotnym. Układ ten umożliwia zastosowanie dużych wartości rezystancji *R* (rzędu kilkudziesięciu MΩ), pomimo prądowego charakteru sterowania tranzystorów bipolarnych, wchodzących w skład przerzutnika.

Wyznaczono zależności wiążące czas opóźnienia z parametrami ciągu impulsów oraz statycznymi i dynamicznymi parametrami przerzutnika. W analizie tej wyodrębniono cztery etapy, w których przebieg zjawisk jest jakościowo różny, przy czym stwierdzono, że czasy trwania poszczególnych etapów zależne są od parametrów ciągu impulsów, obwodu *RC* oraz parametrów określających statyczną charakterystykę wejściową przerzutnika, a czas trwania etapu czwartego zależy ponadto od dynamicznych parametrów przerzutnika.

1. WSTĘP

W układach opóźniających o działaniu przekaźnikowym, opartych na wykorzystaniu inercyjnych własności obwodu *RC* a realizowanych przy użyciu tranzystorów bipolarnych, występują trudności w uzyskaniu dużych opóźnień. Wydłużenie czasu opóźnienia przez powiększenie rezystancji *R* ograniczone jest w tym przypadku prądowym charakterem sterowania tranzystorów, zaś przez powiększenie pojemności *C* — wymiarami kondensatorów o dostatecznie dużej rezystancji izolacji.

Graniczną wartość rezystancji *R* można jednak znacznie powiększyć (nawet do wartości rzędu kilkudziesięciu MΩ) przez zastosowanie odpowiedniego rozwiązania układu opóźniającego. Uzyskuje się to przez włączenie pomiędzy obwód *RC* i przerzutnik bistabilny o określonym progu zadziałania pomocniczego źródła impulsów oraz diody oddzielającej w taki sposób, aby umożliwić przepływ impulsów przez obwód wejściowy przerzutnika dopiero wtedy, gdy napięcie na kondensatorze osiągnie określoną wartość. Zadziałanie przerzutnika, pomimo dużej wartości rezystancji *R*, może nastąpić dzięki temu, że impedancja obwodu, w którym następuje przepływ impulsów, jest niewielka. Maksymalna osiągalna wartość czasu opóźnienia przy określonej dokładności działania zależy w tym

przypadku głównie od stałej czasowej obwodu RC, rezystancji izolacji kondensatora oraz od rezystancji wstecznej diody oddzielającej obwód RC od układu przerzutnika.

Układy opóźniające z tranzystorami bipolarnymi zrealizowane w oparciu o powyższą zasadę działania umożliwiają więc zastosowanie rezystancji R o wartości tego samego rzędu, co układy z tranzystorami unipolarnymi. Układy tego rodzaju są szeroko stosowane [3], [4], [6], [7].

Mimo szerokiego rozpowszechnienia wyżej wymienionej klasy układów opóźniających, w literaturze brak jest dotychczas ogólniejszej analizy zachodzących w nich zjawisk.

Badania przeprowadzone nad tymi układami wykazały ich dużą przydatność do realizacji dużych opóźnień zarówno z punktu widzenia maksymalnie osiągalnych wartości czasów opóźnienia, jak i dokładności działania. W wyniku tych badań nasunęło się jednak wiele problemów, których dokładne przeanalizowanie okazało się niezbędne dla zrozumienia zjawisk zachodzących przy pracy tego rodzaju układów. Problemy te scharakteryzować można następująco:

1° stwierdzono, że czas opóźnienia jest proporcjonalny do pojemności C, lecz nie jest proporcjonalny do rezystancji R;

2° stwierdzono, że czas ten zależny jest od stosunku okresu powtarzania impulsów do czasu ich trwania, przy czym istnieje pewna krytyczna wartość tego stosunku, przy której układ przestaje działać prawidłowo;

3° porównując dwa układy opóźniające różniące się tylko parametrami dynamicznymi tranzystorów zastosowanych w układzie przerzutnika stwierdzono, że przy mniejszej częstotliwości granicznej tranzystorów uzyskuje się dłuższe czasy opóźnienia.

Celem niniejszej pracy jest właśnie wyjaśnienie przyczyn powyższych zjawisk na podstawie jakościowej i ilościowej analizy warunków pracy rozważanej klasy układów opóźniających.

Wykaz ważniejszych oznaczeń

C — pojemność obwodu RC

E_k, E_p — napięcia zasilające przerzutnik i obwód RC

i, I — oznaczenia chwilowych i ustalonych wartości prądów

k — współczynnik nasycenia tranzystorów

K_i — stosunek okresu powtarzania impulsów τ_i do czasu ich trwania t_i

m — współczynnik wysterowania

r, R, ϱ — oznaczenia rezystancji

t — czas

t_i — czas trwania impulsu

T — stała czasowa obwodu RC

u, U — oznaczenia chwilowych i ustalonych wartości napięć

U_i — amplituda impulsów

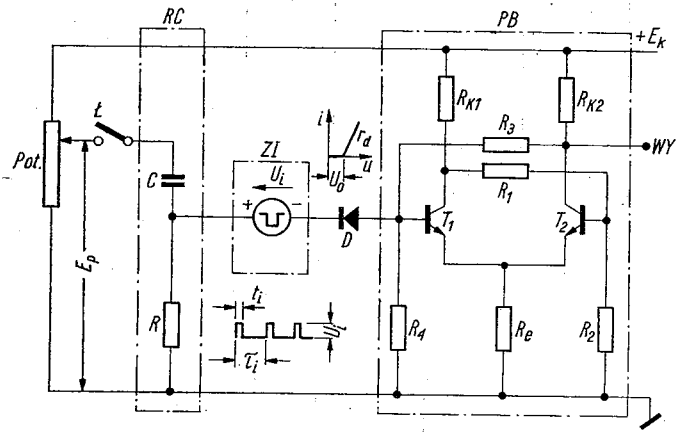
β — współczynnik wzmocnienia tranzystorów w układzie *OE*

τ — oznaczenia stałych czasowych charakteryzujących przerzutnik

τ_i — okres powtarzania impulsów

2. OGÓLNA CHARAKTERYSTYKA ANALIZOWANEGO UKŁADU OPÓŹNIAJĄCEGO

Analizowany w pracy układ opóźniający składa się z trzech zasadniczych członów: obwodu *RC*, źródła impulsów *ZI* i przerzutnika o własnościach bistabilnych *PB* (rys. 1).



Rys. 1. Schemat analizowanego układu opóźniającego

RC — obwód *RC*, *ZI* — źródło impulsów, *PB* — przerzutnik bistabilny, *D* — dioda, *Ł* — łącznik, *Pot.* — potencjometr

Do analizy przyjęto szeregowo-równoległy układ przerzutnika bistabilnego z następujących powodów:

1° przerzutnik taki jest najczęściej stosowany w rozważanych układach opóźniających, gdyż umożliwia zastosowanie tylko jednego źródła napięcia oraz zapewnia — praktycznie biorąc — niezależność czasu opóźnienia od zmian tego napięcia;

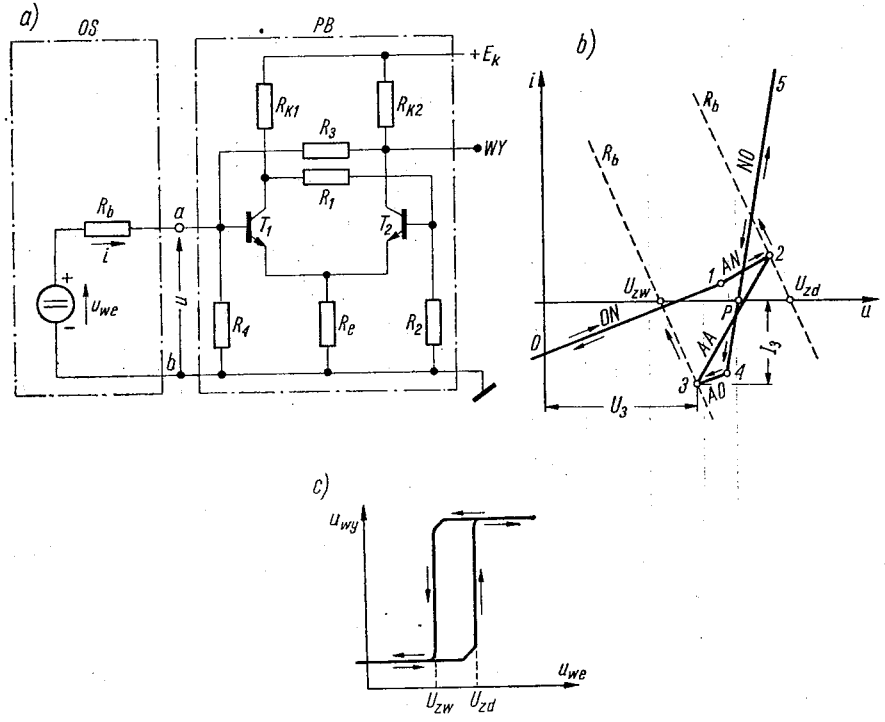
2° wyniki analizy przeprowadzone dla układu opóźniającego z przerzutnikiem szeregowo-równoległym mogą być również odniesione do układu z przerzutnikiem równoległym, który jest jego szczególnym przypadkiem.

Obwód *RC* zasilany jest potencjometrycznie, co umożliwia nastawienie różnych wartości czasu opóźnienia. Układ opóźniający obejmuje także dodatkowe człony nie uwidocznione na rys. 1, zapewniające rozładowanie kondensatora i wstępne ustawienie przerzutnika oraz człon, za pomocą którego przerzutnik steruje elementem wykonawczym (zazwyczaj przekaźnikiem elektromagnetycznym).

Stan przerzutnika *PB* ustalany jest wstępnie w taki sposób, aby tranzystor T_1 był w stanie nasycenia. Odmierzanie czasu opóźnienia rozpoczyna się z chwilą zamknięcia łącznika *Ł*, kończy zaś z chwilą, gdy suma napięcia na kondensatorze i napięcia wytwarzanego przez źródło impulsów *ZI* (np. o amplitudzie 0,1 E_k) osiągnie wartość wystarczającą do spowodowania zmiany stanu przerzutnika.

3. STATYCZNA CHARAKTERYSTYKA WEJŚCIOWA PRZERZUTNIKA

Statyczną charakterystykę wejściową przerzutnika o szeregowo-równoległym sprzężeniu zwrotnym zgodnie z rozważaniami przeprowadzonymi w pracy [5], przedstawić można w postaci linii łamanej 01—12—23—34—45 (rys. 2). Poszczególne odcinki tej cha-



Rys. 2. Charakterystyki statyczne przerzutnika o szeregowo-równoległym sprzężeniu zwrotnym
a) schemat układu pomiarowego do wyznaczania charakterystyk: *OS* — układ sterujący, *PB* — przerzutnik, b) charakterystyka wejściowa, c) charakterystyka wyjściowa

rakterystyki odpowiadają następującym stanom przerzutnika: *ON—AN—AA—AO—NO*, gdzie symbole *A*, *O*, *N* oznaczają odpowiednio stan aktywny, odcięcia i nasycenia tranzystora, przy czym litera na pierwszym miejscu oznacza stan tranzystora T_1, a na drugim — stan tranzystora T_2.

Wartości napięć wejściowych, przy których charakterystyki obwodu sterującego (proste zaznaczone na rys. 2b liniami przerywanymi) przechodzą przez punkty 2 i 3, odpowiadają progom zadziałania U_{zd} i zwalniania U_{zw} przerzutnika.

Dzięki temu, że w stanach stabilnych tranzystor T_2 może być tylko odcięty lub nasycony, uzyskuje się przekaźnikowy przebieg charakterystyki wyjściowej (rys. 2c).

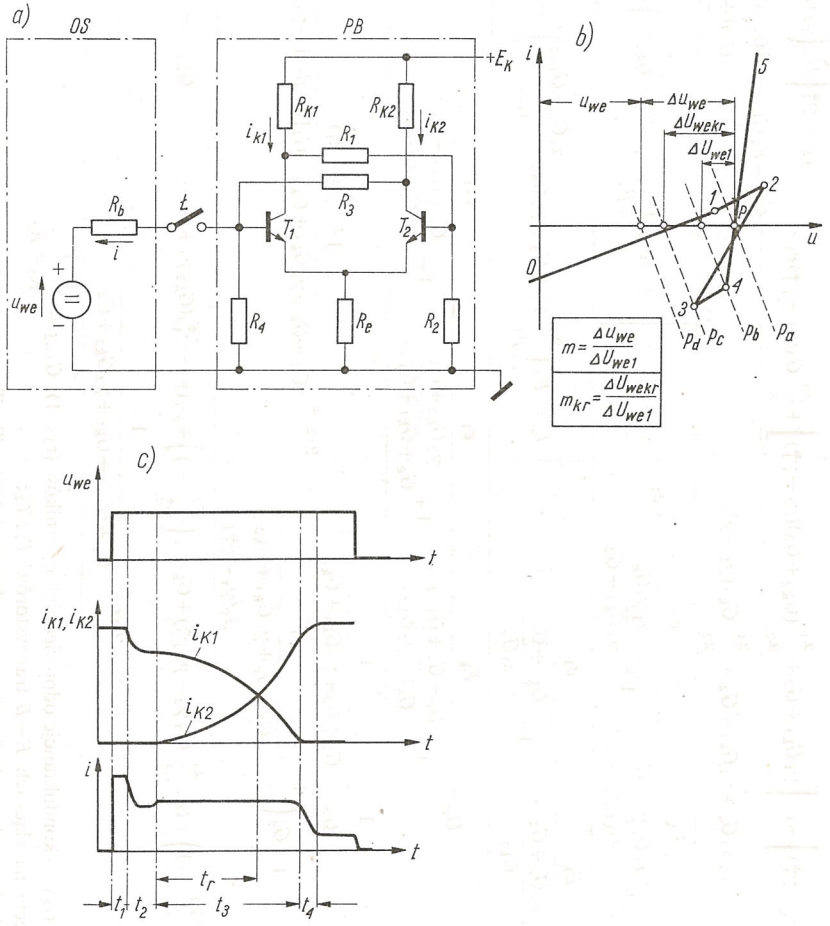
Wartości współrzędnych podstawowych punktów charakterystyki wejściowej, jak również wartości rezystancji dynamicznych odpowiadających poszczególnym odcinkom tej charakterystyki zestawiono w tablicy. Wzory te mogą być wykorzystane do celów analizy ilościowej procesu odmierzania czasu.

Parametry określające statyczną charakterystykę wejściową przerzutnika Tablica 1

współrzędne p. 1	$U_1 = E_k\frac{G_{k2}+\gamma_1 G_{k1}}{G_e+G_{k2}+\gamma_1 G_{k1}+G_2}+\frac{e_1G_3-e_2(G_{k1}\gamma_1+G_2)-e_{s2}(G_3+G_{k2})}{G_e+G_{k2}+\gamma_1 G_{k1}+G_2}+e_1$	$I_1 = U_1G_4+(e_1-e_{s2})G_3$
współrzędne p. 2	$U_2 = \frac{E_k\left[G_{k1}\gamma_1+\frac{\alpha_1}{\alpha_2}G_{k2}(\gamma_1-\gamma_1^*)\right]-e_2\left[\gamma_1G_{k1}+G_2+\frac{\alpha_1}{\alpha_2}(G_{k2}+G_3)(\gamma_1-\gamma_1^*)\right]+\frac{\alpha_1}{\alpha_2}G_3(\gamma_1-\gamma_1^*)e_1}{\gamma_1\alpha_1G_e+\gamma_1G_{k1}+G_2+\frac{\alpha_1}{\alpha_2}G_{k2}(\gamma_1-\gamma_1^*)}+e_1$	$I_2 = \gamma_2^*\left[U_2\left(\alpha_2G_e+G_{k2}+\frac{1}{\gamma_2^*}G_4\right)-E_kG_{k2}-e_1(G_e\alpha_2+G_{k2}+G_3)+e_2(G_{k2}+G_3)\right]+(e_1-e_2)G_3$
współrzędne p. 3	$U_3 = \frac{E_k}{1+\frac{\alpha_1G_e\gamma_1+G_2}{G_{k1}\gamma_1}}-\frac{e_2}{1+\frac{\alpha_1\gamma_1G_e}{\gamma_1G_{k1}+G_2}}+e_1$	$I_3 = U_3\left(\frac{\alpha_1}{\beta_1}G_e+G_4+G_{k2}\gamma_2\right)-E_kG_{k2}\gamma_2-e_1\frac{\alpha_1}{\beta_1}G_e$
współrzędne p. 4	$U_4 = \frac{E_k}{1+\frac{\alpha_1G_e+G_{1-2}}{G_{k1}}}+\frac{e_1}{1+\frac{G_{k1}+G_{1-2}}{\alpha_1G_e}}$	$I_4 = U_4\left(\frac{\alpha_1}{\beta_1}G_e+G_4-\gamma_2G_{k2}\frac{\alpha_1G_e+G_{1-2}}{G_{k1}}\right)+e_1\alpha_1G_e\left(\gamma_2\frac{G_{k2}}{G_{k1}}-\frac{1}{\beta_1}\right)$
napięcie U_p	$U_p = \frac{E_k}{1+\frac{G_e+G_4+G_{1-2}}{G_{k1}+\gamma_2G_{k2}}}+\frac{e_1}{1+\frac{\gamma_2G_{k2}+G_4}{G_e+G_{k1}+G_{1-2}}}-\frac{e_{s1}}{1+\frac{G_e+G_{k1}+G_4}{\gamma_2G_{k2}+G_{1-2}}}$	
rezystancje odpowiadające poszczególnym odcinkom charakterystyki	$\varrho_{01} = \frac{\frac{1}{G_3}+\frac{1}{G_2+G_e+\gamma_1G_{k1}+G_{k2}}}{1+G_4\left(\frac{1}{G_3}+\frac{1}{G_2+G_e+\gamma_1G_{k1}+G_{k2}}\right)}$, $\varrho_{12} = \frac{1+\beta_1(1-\gamma_1)}{G_2+G_e+\gamma_1G_{k1}+G_{k2}+G_4[1+\beta_1(1-\gamma_1)]}$, $\varrho_{23} = \frac{\beta_1(\gamma_1-\gamma_1^*)}{G_2\left(\frac{\gamma_2}{\gamma_2^*}-1\right)+G_e\frac{\alpha_1}{\gamma_2^*}(\gamma_1\gamma_2-\gamma_1^*\gamma_2^*)+G_{k1}\gamma_1\left(\frac{\gamma_2}{\gamma_2^*}-1\right)+\beta_1(\gamma_1-\gamma_1^*)(G_{k2}\gamma_2+G_4)}$, $\varrho_{34} = \frac{1}{G_{k2}\gamma_2+\frac{\alpha_1}{\beta_1}G_e+G_4}$, $\varrho_{45} = \frac{1}{G_e+G_{1-2}+G_{k1}+\gamma_2G_{k2}+G_4}$	
oznaczenia	$G_1, G_2, G_3, G_4, G_{k1}, G_{k2}$ — konduktancje odpowiednich oporników (rys. 1); $G_{1-2} = \frac{1}{R_1+R_2}$; e_1, e_2 — spadki napięcia na złączach $E-B$ tranzystorów T_1, T_2; e_{s1}, e_{s2} — spadki napięcia na złączach $E-K$ nasyconych tranzystorów T_1, T_2; $\alpha_1, \alpha_2, \beta_1, \beta_2$ — współczynniki wzmocnienia prądowego w składzie OB i OE tranzystorów T_1, T_2; współczynniki $\gamma_1, \gamma_2, \gamma_1^*, \gamma_2^*$ zdefiniowane są wzorami: $\gamma_1 = \frac{G_1}{G_1+G_{k1}}$, $\gamma_2 = \frac{G_3}{G_3+G_{k2}}$, $\gamma_1^* = \frac{\alpha_2}{\alpha_1\beta_2}$, $\gamma_2^* = \frac{\alpha_1}{\alpha_2\beta_1}$	

4. PROCESY PRZEJŚCIOWE ZACHODZĄCE W PRZERZUTNIKU

Procesy przejściowe zachodzące w przerzutnikach bistabilnych były przedmiotem wielu prac, np. [2], jednakże wyniki tych prac nie mogą być w tym przypadku bezpośrednio wykorzystane ze względu na odmienne warunki sterowania przerzutnika. Niżej przeanali-



Rys. 3. Przebiegi przejściowe w przerzutniku

a) schemat układu do wyznaczania przebiegów przejściowych, b) ilustracja sposobu zdefiniowania współczynnika wysterowania m, c) przebiegi prądów: i_{K1}, i_{K2}, i przy zmianie stanu przerzutnika z *NO* do stanu *ON*, $t_1 \ldots t_4$ — czasy trwania czterech kolejnych faz procesów przejściowych

zowano procesy przejściowe zachodzące w przerzutniku, zakładając zgodnie z warunkami sterowania występującymi w rozważanym układzie opóźniającym, że do wejścia przerzutnika zostaje doprowadzony impuls napięcia ze źródła o rezystancji R_b (rys. 3a). Zakłada się przy tym, że tranzystor T_1 jest nasycony, T_2 — zablokowany oraz że amplituda napięcia wejściowego u_{we} jest mniejsza od progu zwalniania U_{zw} (rys. 2b), tzn. że spełniony jest warunek: $u_{we} \leqslant U_3 - |I_3| R_b$.

Przy zmianie stanu przerzutnika (z *NO* do *ON*) można wyodrębnić cztery fazy procesów przejściowych (rys. 3c).

W fazie pierwszej odbywa się wysysanie nagromadzonych w bazie tranzystora T_1 nośników mniejszościowych przy odciętym tranzystorze T_2. Z chwilą usunięcia z bazy tranzystora T_1 całego ładunku nadmiarowego, tranzystor ten wchodzi w obszar aktywny i rozpoczyna się faza druga, w której następuje zmniejszanie prądu kolektorowego i_{k1} tranzystora T_1 aż do wartości, przy której w obszar aktywny wchodzi zablokowany dotychczas tranzystor T_2.

W fazie trzeciej obydwa tranzystory wykazują własności wzmacniające. Prąd i_{k1} maleje, zaś i_{k2} rośnie, przy czym przebiegi tych prądów mają charakter wykładniczy.

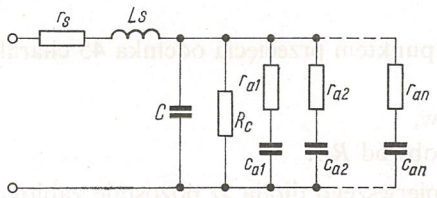
Koniec fazy trzeciej następuje z chwilą, gdy prąd kolektorowy tranzystora T_1 zmaleje do zera. Od tej chwili (faza czwarta) zmienia się jedynie prąd kolektorowy tranzystora T_2, aż do osiągnięcia stanu nasycenia.

W dodatku przedstawiono wyniki uproszczonej analizy ilościowej. W analizie tej posługiwano się tzw. współczynnikiem wysterowania m zdefiniowanym za pomocą rys. 3b, zamiast bezwzględną wartością napięcia u_{we}. W wyniku tej analizy stwierdzono, że dla spowodowania zmiany stanu przerzutnika konieczna jest obecność impulsu wejściowego przez czas t_r (rys. 3c), a więc w pierwszej, drugiej i takiej części fazy trzeciej, po upływie której zostanie spełniony warunek: $i_{k1} = i_{k2}$.

Zależności przedstawione w dodatku pozwalają na wyznaczenie wartości czasów: t_1, t_2, t_r oraz na określenie prądu i pobieranego przez obwód wejściowy przerzutnika. Z zależności tych wynika, że zmiana stanu przerzutnika możliwa jest przy spełnieniu warunku: $m > m_{kr}$ oraz że wraz ze wzrostem współczynnika wysterowania m wzrasta szybkość z jaką przebiegają procesy przejściowe, ale równocześnie wzrasta średnia wartość prądu pobieranego przez przerzutnik ze źródła sterującego.

5. ANALIZA JAKOŚCIOWA PROCESU ODMIERZANIA CZASU

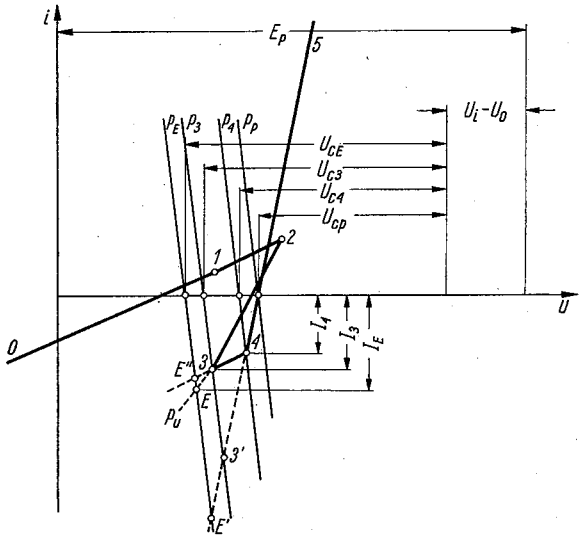
Przedstawiona niżej analiza dotyczy układu opóźniającego zamieszczonego na rys. 1. Rezystancję wyjściową potencjometru, z którego zasilany jest obwód RC oznaczono przez r_p oraz założono, że statyczną charakterystykę diody D można opisać zależnością: $u = U_o + ir_d$, gdzie U_o oznacza napięcie progowe, zaś r_d — rezystancję dynamiczną.



Rys. 4. Schemat zastępczy kondensatora

C — pojemność, R_c — rezystancja równoległa uwzględniająca upływność; r_s, L_s — rezystancja i indukcyjność szeregowa; r_a, c_a — obwody uwzględniające absorpcję dielektryczną

Przy analizie nie uwzględniono niektórych elementów schematu zastępczego kondensatora (rys. 4), a mianowicie rezystancji równoległej R_c, rezystancji i indukcyjności szeregowej oraz obwodów r_a, c_a, których obecność w schemacie wynika ze zjawisk absorpcyj-



Rys. 5. Proces odmierzania czasu na tle charakterysytki wejściowej przerzutnika

nych zachodzących w dielektryku. Wpływ tych parametrów na proces odmierzania czasu jest pomijalnie mały w przypadku doboru odpowiednich kondensatorów (np. poliestrowych).

Zjawiska zachodzące w układzie opóźniającym można przeanalizować posługując się charakterystyką wejściową przerzutnika (rys. 5).

Cały proces odmierzania czasu można podzielić na cztery etapy, w których przebieg zjawisk jest jakościowo odmienny. Położenie prostych p_p, p_4, p_3, p_E na rys. 5 odpowiada końcowym chwilom kolejnych etapów, zaś nachylenie zależne jest od sumy rezystancji (r_p+r_d). Przez U_{cp}, U_{c4}, U_{c3}, U_{cE} oznaczono napięcie na kondensatorze w końcowych chwilach kolejnych etapów.

Etap pierwszy rozpoczyna się z chwilą zamknięcia łącznika Ł, kończy zaś wtedy, gdy napięcie na kondensatorze osiągnie wartość U_{cp} określoną wzorem:

$$U_{cp} = E_p + U_o - U_i - U_p, \tag{1}$$

gdzie:

U_p — napięcie określone punktem przecięcia odcinka 45 charakterystyki wejściowej z osią odciętych,

U_i — amplituda impulsów,

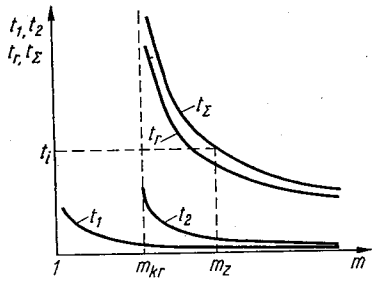
E_p — napięcie zasilające obwód RC.

W czasie trwania etapu pierwszego dioda D pozostaje zablokowana.

Etap drugi wyznaczony jest przez chwile, w których napięcie na kondensatorze osiąga wartości odpowiednio U_{cp} i U_{c4} (współczynnik wysterowania m zmienia się przy tym w granicach 0 ... 1). W etapie tym następuje przepływ prądu przez obwód wejściowy przerzutnika w czasie trwania impulsu. Prąd ten zachowuje stałą wartość przez cały czas trwania impulsu, przy czym w miarę przepływu kolejnych impulsów prąd wzrasta od zera (przy $u_c = U_{cp}$) do wartości I_4 (przy $u_c = U_{c4}$).

W etapie trzecim napięcie na kondensatorze zmienia się w zakresie od U_{c4} do U_{c3}, któremu to zakresowi odpowiada zmiana współczynnika wysterowania m w zakresie: 1 ... m_{kr} (rys. 3b). Prąd wejściowy w chwili pojawienia się impulsu przyjmuje wartość odpowiadającą punktowi przecięcia odpowiedniej prostej równoległej do p_p, p_4, p_3, p_E z przedłużeniem odcinka 45, przy czym wartość ta pozostaje niezmienna przez czas t_1 będący czasem trwania pierwszej fazy procesów przejściowych. Następnie prąd wejściowy maleje wykładniczo ze stałą czasową τ_z (wzór D8) aż do osiągnięcia wartości ustalonej, odpowiadającej punktowi przecięcia tejże prostej z odcinkiem 34 charakterystyki wejściowej.

Etap czwarty rozpoczyna się z chwilą, gdy napięcie na kondensatorze osiąga wartość U_{c3}, której odpowiada współczynnik wysterowania m_{kr}, kończy zaś przy takiej wartości U_{cE}, przy której współczynnik m osiąga wartość wystarczającą do spowodowania zmiany stanu przerzutnika, pomimo skończonego czasu trwania impulsu t_i.



Rys. 6. Charakterystyki $t_1, t_2, t_r, t_\Sigma = f(m)$

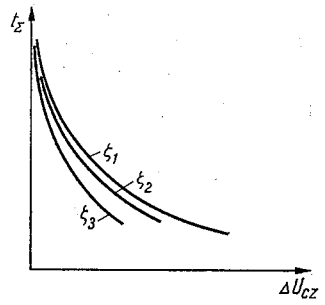
Zgodnie z rozważaniami przeprowadzonymi w p. 4, przerzut nastąpi przy takiej wartości napięcia na kondensatorze, przy której zostanie spełniony warunek:

$$t_i = t_\Sigma, \tag{2}$$

gdzie:

$$t_\Sigma = t_1 + t_2 + t_r. \tag{3}$$

Wartości czasów: t_1, t_2, t_r wyznaczyć można ze wzorów (D5), (D7), (D17), przy czym w celu wyznaczenia występującego tam współczynnika ζ należy uwzględnić, że rezystancja obwodu sterującego przerzutnikiem wynosi: $R_b = r_p + r_d$. Na podstawie powyższych zależności można wyznaczyć charakterystykę $t_\Sigma = f(m)$ (rys. 6). Charakterystyka ta umożliwia — przy zadanym czasie trwania impulsu — wyznaczenie takiej wartości m_z, po osiągnięciu której przerzutnik zmienia stan. Różnicy współczynników wysterowania: $\Delta m = m_z - m_{kr}$ odpowiada zmiana napięcia na kondensatorze o wartość ΔU_{cz}, którą można wyznaczyć ze wzoru (D4) przy podstawieniu $m = m_z$. Zatem ΔU_{cz} określa, o ile musi wzrosnąć dodatkowo napięcie na kondensatorze w porównaniu z wartością odpowiadającą statycznemu progowi zwalniania przerzutnika, aby przy zadanym czasie trwania impulsu t_i nastąpiła zmiana stanu przerzutnika. Wykorzystując zależność (D4) oraz charakterystykę z rys. 6, zbudowano charakterystykę $t_\Sigma = f(\Delta U_{cz})$ pozwalającą na wy-



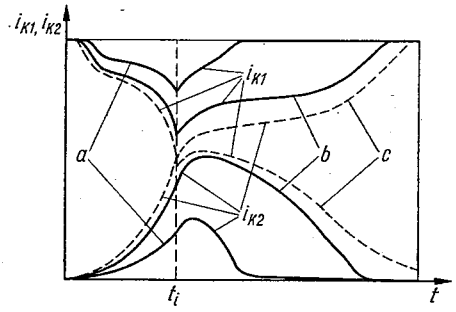
Rys. 7. Charakterystyki $t_\Sigma = f(\Delta U_{cz})$ przy różnych wartościach parametru ζ: $\zeta_1 > \zeta_2 > \zeta_3$

znaczenie napięcia ΔU_{cz} przy różnych czasach trwania impulsu t_i (rys. 7). Jak widać, przy zadanej wartości t_i napięcie to zwiększa się wraz ze wzrostem rezystancji (a więc i współczynnika ζ) obwodu, w którym następuje przepływ impulsów.

Typowy przebieg prądu przepływającego przez obwód wejściowy przerzutnika w rozważanym czwartym etapie przedstawiono na rys. 3. Bezpośrednio po pojawieniu się impulsu wejściowego prąd ten osiąga wartość określoną punktem przecięcia odpowiedniej prostej (np. p_E) z przedłużeniem odcinka 45 charakterystyki wejściowej (rys. 5, p. E'), zachowuje stałą wartość w czasie t_1 a następnie maleje ze stałą czasową τ_z do wartości ustalonej (p. E'') i pozostaje niezmieniony aż do zakończenia fazy drugiej. W fazie trzeciej prąd ten wzrasta ze stałą czasową τ_2 osiągając (p. E) wartość ustaloną określoną wzorem (D18) i pozostaje niezmieniony aż do końca czasu trwania impulsu.

Ustalone wartości prądu wejściowego w trzeciej fazie procesów przejściowych leżą na prostej p_u, którą można wyznaczyć na podstawie wzoru (D18) dla określonej wartości współczynnika ζ. Opisany wyżej przebieg zmian prądu wejściowego występuje wtedy, gdy na przebiegi przejściowe w fazie drugiej wywierają dominujący wpływ pojemności złączowe tranzystorów. W przypadku, gdy wpływ ten jest do pominięcia, prąd wejściowy osiąga wartość ustaloną dopiero w fazie trzeciej. Na charakterystyce wejściowej odpowiada to przejściu z punktu E' bezpośrednio do punktu E.

Typowe przebiegi prądów kolektorowych i_{k1}, i_{k2} przedstawiono na rys. 8; są one opisane przez zależności (D11), (D12) w czasie trwania impulsu, a po jego zaniku przez



Rys. 8. Typowe przebiegi prądów kolektorowych i_{k1}, i_{k2} w czasie trwania impulsu i tuż po jego zaniknięciu
a, b — dla $m_{kr} < m < m_z$, c — dla $m > m_z$

wzory (D14), (D15). Z rysunku tego wynika, że dla przypadków *a*, *b* prąd i_{k2} po zaniku impulsu zwiększa się, osiąga maksimum i następnie maleje. Występowanie maksimum można wyjaśnić wpływem członu z wykładnikiem ujemnym we wzorze (D15).

Przeprowadzona dotychczas analiza jakościowa procesu odmierzania czasu pozwala na wyznaczenie minimalnej amplitudy impulsu U_i, niezbędnej do prawidłowego działania układu. Mianowicie, w zakresie zmian współczynnika m od zera do m_z, po zaniku impulsu dioda D powinna pozostawać zablokowana. Zgodnie z rys. 7 odpowiada to spełnieniu warunku:

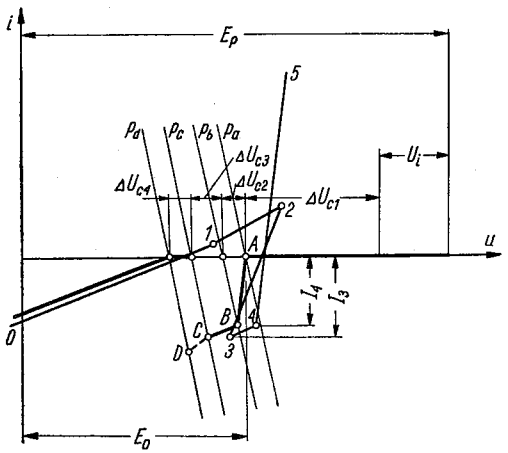
$$U_i \geqslant U_p - [U_3 - (r_p + r_d)|I_3| - \Delta U_{cz}]. \tag{4}$$

Wartości napięć U_3, U_p oraz prądu I_3 określone są odpowiednimi zależnościami w tablicy 1.

6. ANALIZA ILOŚCIOWA PROCESU ODMIERZANIA CZASU

Analizę ilościową procesu odmierzania czasu przeprowadzono wykorzystując charakterystykę wejściową układu: przerzutnik — dioda D (rys. 9). W celu dokonania tej analizy wprowadzono następujące oznaczenia:

E_o — wartość napięcia odpowiadającego punktowi A (rys. 9),
r, R_d — rezystancje dynamiczne odpowiadające odcinkom AB i BC,
I_3, I_4 — bezwzględne wartości prądów odpowiadające punktom C i B,
E_p — napięcie zasilające obwód RC,
r_p — rezystancja wyjściowa potencjometru zasilającego obwód RC,
r_d, U_o — rezystancja dynamiczna i napięcie progowe diody D,
U_i, t_i, τ_i — amplituda impulsów, czas trwania i odstęp między nimi.



Rys. 9. Proces odmierzania czasu na tle charakterystyki wejściowej układu: przerzutnik — dioda D
—— charakterystyka wejściowa przerzutnika, ▬▬ charakterystyka wejściowa układu: przerzutnik — dioda D; ΔU_{c1}, ΔU_{c2}, ΔU_{c3}, ΔU_{c4} — zmiany napięcia na kondensatorze w 1, 2, 3, 4-tym etapie procesu odmierzania czasu; p_a, p_b, p_c, p_d — proste o nachyleniu odpowiadającym rezystancji r_p

Wartość napięcia E_o oraz rezystancje r, R_d określone są wzorami:

$$E_o = U_p - U_o, \tag{5}$$

$$r = r_d + \varrho_{45}, \tag{6}$$

$$R_d = r_d + \varrho_{34}, \tag{7}$$

gdzie: U_p, ϱ_{34}, ϱ_{45} — parametry opisujące charakterystykę wejściową przerzutnika (tablica).

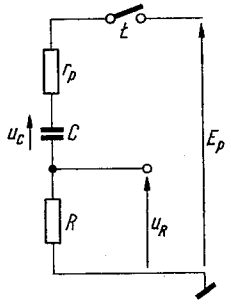
Ponadto wprowadzono współczynnik K_i zdefiniowany wzorem:

$$K_i = \frac{\tau_i}{t_i}. \tag{8}$$

6.1. Analiza etapu pierwszego

W schemacie zastępczym przedstawionym na rys. 10 przy zerowej wartości początkowej napięcia na kondensatorze, przebieg zmian napięcia u_R określony jest wzorem:

$$u_R = E_p \frac{R}{R+r_p} e^{-\frac{t}{(R+r_p)C}}. \tag{9}$$



Rys. 10. Schemat zastępczy układu opóźniającego przyjęty do analizy etapu pierwszego

Czas trwania etapu pierwszego można wyznaczyć z powyższego wzoru przyjmując: $t = t_{o1}$, $u_R = E_o + U_i$, Otrzymuje się wtedy:

$$t_{o1} = (R+r_p)C \ln \frac{E_p \cdot \frac{R}{R+r_p}}{E_o + U_i}, \tag{10}$$

lub po uwzględnieniu, że: $r_p \ll R$

$$t_{o1} = RC \ln \frac{E_p}{E_o + U_i}. \tag{11}$$

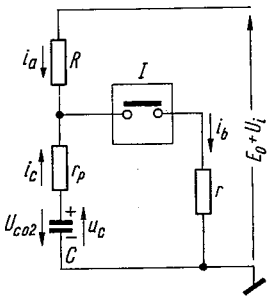
6.2. Analiza etapu drugiego

Analizę etapu drugiego można przeprowadzić w oparciu o schemat zastępczy przedstawiony na rys. 11. Zakłada się, że czas zamknięcia zestyków impulsatora I wynosi t_i,

zaś odstęp między kolejnymi zadziałaniami wynosi τ_i. Początkowa wartość napięcia na kondensatorze wynosi:

$$U_{co2} = \frac{r_p}{R}(E_o + U_i), \tag{12}$$

przy czym znak tego napięcia jest taki, jak zaznaczono na rys 11.



Rys. 11. Schemat zastępczy układu opóźniającego przyjęty do analizy etapu drugiego
I — impulsator

Oznaczając przez n liczbę zadziałań impulsatora I oraz wprowadzając następujące oznaczenia:

$$\left.\begin{aligned} a &= \frac{\tau_i - t_i}{\tau_i}, \qquad t' = \frac{t}{\tau_i}, \qquad T = C(R+r_p), \\ T_{c2} &= \left(r_p + \frac{Rr}{R+r}\right)C, \qquad A_2 = \frac{C(R+r_p)}{\tau_i}, \qquad B_2 = \frac{T_{c2}}{\tau_i}, \end{aligned}\right\} \tag{13}$$

można dla rozważanego układu przy uwzględnieniu warunków początkowych uzyskać następujące zależności:

$$u_c(t') = E_o + U_i + [u_c(n) - (E_o + U_i)]e^{-\frac{t'-n}{A_2}} \qquad \text{dla:} \quad n \leqslant t' \leqslant n+a, \tag{14}$$

$$u_c(t') = \left\{E_o + U_i + [u_c(n) - (E_o + U_i)]e^{-\frac{a}{A_2}}\right\} e^{-\frac{t'-n-a}{B_2}} + \left(1 - e^{-\frac{t'-n-a}{B_2}}\right)\frac{r}{R+r}(E_o + U_i) \tag{15}$$

$$\text{dla: } n+a \leqslant t' \leqslant n+1.$$

Ze wzorów tych otrzymuje się następujące równanie różnicowe:

$$u_c[n+1] - u_c[n]e^{-b_2} = D_2, \tag{16}$$

gdzie:

$$\left.\begin{aligned} b_2 &= \frac{a}{A_2} + \frac{1-a}{B_2}, \\ D_2 &= (E_o + U_i)\left(1 - e^{-\frac{a}{A_2}}\right)e^{-\frac{1-a}{B_2}} + \left(1 - e^{-\frac{1-a}{B_2}}\right)\frac{r}{R+r}(E_o + U_i). \end{aligned}\right\} \tag{17}$$

Rozwiązanie równania (16) pozwala na wyznaczenie $U_c[n]$, a następnie na wyznaczenie przebiegu zmian napięcia na kondensatorze jako funkcji liczby impulsów n. Otrzymuje się wtedy:

$$u_c(t') = E_o + U_i + \left[D_2 \frac{1-e^{-b_2 n}}{1-e^{-b_2}} - \frac{r_p}{R}(E_o+U_i)e^{-b_2 n} - (E_o+U_i) \right] e^{-\frac{t'-n}{A_2}} \quad \text{dla:} \quad n \leqslant t' \leqslant n+a, \tag{18}$$

$$u_c(t') = \left\{ E_o + U_i + \left[D_2 \frac{1-e^{-b_2 n}}{1-e^{-b_2}} - \frac{r_p}{R}(E_o+U_i)e^{-b_2 n} - (E_o+U_i) \right] e^{-\frac{a}{A_2}} \right\} e^{-\frac{t'-n-a}{B_2}} + \left(1-e^{-\frac{t'-n-a}{B_2}}\right) \frac{r}{R+r}(E_o+U_i) \quad \text{dla:} \quad n+a \leqslant t' \leqslant n+1. \tag{19}$$

Drugi etap odmierzania czasu kończy się z chwilą, gdy napięcie na kondensatorze osiągnie taką wartość, że prąd rozładowania kondensatora w czasie trwania impulsu, a więc w chwili: $n+a \leqslant t' \leqslant n+1$ osiągnie wartość I_4. Zgodnie ze wzorem (19) napięcie u_c w chwili $[n+a]$ ma wartość większą, niż w chwili $[n+1]$, dlatego słuszniej będzie przyjąć, że prąd I_4 zostanie osiągnięty w chwili $[n+1]$. Wartość tego napięcia można określić ze wzoru (19) podstawiając: $t' = n+1$.

Dla rozpatrywanej chwili odpowiadającej końcowi drugiego etapu słuszna jest następująca zależność wynikająca bezpośrednio ze schematu zastępczego (rys. 11) przy założeniu: $i_b = I_4$:

$$u_c[n+1] = I_4\left(r_p + r + \frac{r r_p}{R}\right) - \frac{E_o+U_i}{R} r_p. \tag{20}$$

Czas trwania drugiego etapu związany jest z liczbą impulsów n zależnością:

$$t_{o2} = (n+1)\tau_i.$$

Wykorzystując zależności (20), (21) oraz (19) dla warunku: $t' = n+1$, otrzymuje się wyrażenie określające czas trwania drugiego etapu:

$$t_{o2} = \tau_i + \frac{\tau_i}{\frac{a}{A_2} + \frac{1-a}{B_2}} \ln \times$$

$$\times \frac{1}{1-\left(e^{\frac{a}{A_2}+\frac{1-a}{B_2}}-1\right)\left[\frac{\frac{I_4\left[r+r_p\left(1+\frac{r}{R}\right)\right]}{E_o+U_i}}{\left(1-e^{-\frac{a}{A_2}}\right)e^{-\frac{1-a}{B_2}} + \frac{r}{R+r}\left(1-e^{-\frac{1-a}{B_2}}\right) + \frac{r_p}{R}\left(1-e^{-\frac{a}{A_2}-\frac{1-a}{B_2}}\right)} - 1 \right]}. \tag{22}$$

Zazwyczaj spełnione są następujące warunki:

$$T \gg \tau_i, \quad T_{c2} \gg t_i; \quad \tau_i \gg t_i; \quad T \gg T_{c2} \tag{23}$$

i wtedy wzór (22) można uprościć do postaci:

$$t_{o2} = \tau_i + \frac{K_i T_{c2}}{1+K_i\frac{T_{c2}}{T}} \ln \frac{1}{1-\left(1+K_i\frac{T_{c2}}{T}\right)\left[\frac{I_4 R}{K_i(E_e+U_i)} - \frac{t_i}{T_{c2}}\right]}. \tag{24}$$

Wzór ten można poddać dalszemu uproszczeniu uwzględniając, że zwykle:

$$\frac{I_4 R}{K_i(E_o+U_i)} \gg \frac{t_i}{T_{c2}}, \quad t_{o2} \gg \tau_i; \tag{25}$$

wtedy:

$$t_{o2} = \frac{K_i T_{c2}}{1+K_i\frac{T_{c2}}{T}} \ln \frac{1}{1-\frac{I_4 R}{K_i(E_o+U_i)}\left(1+K_i\frac{T_{c2}}{T}\right)}. \tag{26}$$

Należy tu zaznaczyć, że założenia, przy których został wyprowadzony wzór (24) spełnione są praktycznie biorąc — zawsze.

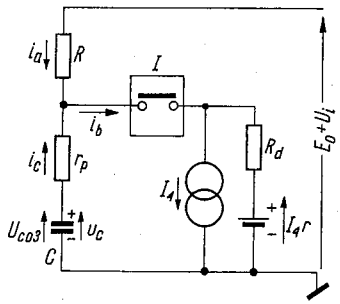
Przy dużych wartościach $K_i\left(K_i \gg \frac{I_4 R}{E_o+U_i}\right)$ oraz małych wartościach stałej czasowej T_{c2} wartość czasu t_{o2} może być niewielka i porównywalna z wartością τ_i. W takim przypadku nie są spełnione warunki (25) i korzystanie ze wzoru (26) nie jest dopuszczalne. Wzory (22), (24) wyprowadzone zostały przy założeniu, że pierwszy impuls (zamknięcie zestyków impulsatora I na rys. 11) pojawia się po czasie $(\tau_i - t_i)$ od chwili, w której rozpoczyna się liczenie czasu t_{o2}. W rzeczywistości czas ten może zawierać się w zakresie $0 \ldots (\tau_i - t_i)$. Dlatego czas t_{o2} określony jest za pomocą powyższych wzorów z dokładnością do $\pm\tau_i$.

Ze wzoru (26) wynika, że istnieje pewna krytyczna wartość współczynnika K_i, przy której uzyskuje się nieprawidłową pracę układu opóźniającego ($t_{o2} \to \infty$). Wartość ta — jak wynika ze wzoru (26) — wynosi:

$$K_{ikr2} = \frac{I_4 R}{E_o+U_i-I_4(r+r_p)}. \tag{27}$$

6.3. Analiza etapu trzeciego

Schemat zastępczy układu przyjęty do analizy etapu trzeciego przedstawiono na rys. 12. Schemat ten narysowano zakładając, że wystarczy uwzględnić tylko składową ustaloną prądu wejściowego, tzn. spełniając warunek: $t_i \gg \tau_z$, gdzie stała czasowa τ_z określona jest wzorem (D8). Początkowe napięcie na kondensatorze U_{co3} ma taką samą wartość, jak przy końcu etapu drugiego (wzór 20).



Rys. 12. Schemat zastępczy układu opóźniającego przyjęty do analizy etapu trzeciego
I — impulsator

Przeprowadzając dla schematu przedstawionego na rys. 12 analizę analogiczną jak dla etapu drugiego, otrzymuje się następujące zależności określające przebieg zmian napięcia na kondensatorze.

$$u_c(t') = E_o + U_i + \left[D_3 \frac{1-e^{-b_3 n}}{1-e^{-b_3}} + U_{co3} e^{-b_3 n} - (E_o + U_i) \right] e^{-\frac{t'-n}{A_3}} \quad \text{dla:} \quad n \leqslant t' \leqslant n+a, \tag{28}$$

$$u_c(t') = \left\{ E_o + U_i + \left[D_3 \frac{1-e^{-b_3 n}}{1-e^{-b_3}} + U_{co3} e^{-b_3 n} - (E_o + U_i) \right] e^{-\frac{a}{A_3}} \right\} e^{-\frac{t'-n-a}{B_3}} + \\ + \left(1 - e^{-\frac{t'-n-a}{B_3}}\right) \left[(E_o + U_i) \frac{R_d}{R+R_d} - I_4 \frac{R(R_d - r)}{R+R_d} \right] \quad \text{dla:} \quad n+a \leqslant t' \leqslant n+1, \tag{29}$$

w których przyjęto następujące oznaczenia:

$$\left.\begin{aligned} t' &= \frac{t}{\tau_i}, \quad T = C(R+r_p), \quad T_{c3} = \left(r_p + \frac{R R_d}{R+R_d}\right) C, \\ A_3 &= \frac{T}{\tau_i}, \quad B_3 = \frac{T_{c3}}{\tau_i}, \quad b_3 = \frac{a}{A_3} + \frac{1-a}{B_3}, \\ D_3 &= (E_o + U_i)\left(1 - e^{-\frac{a}{A_3}}\right) e^{-\frac{1-a}{B_3}} + \left(1 - e^{-\frac{1-a}{B_3}}\right) \left[(E_o + U_i) \frac{R_d}{R+R_d} - I_4 \frac{R(R_d-r)}{R+R_d} \right]. \end{aligned}\right\} \tag{30}$$

Etap trzeci kończy się z chwilą, gdy napięcie na kondensatorze osiągnie taką wartość, że prąd rozładowania kondensatora w czasie trwania impulsu osiągnie wartość I_3. Zakładając, że prąd I_3 zostaje osiągnięty w chwili $[n+1]$, można wyznaczyć wartość napięcia na kondensatorze ze wzoru (29) podstawiając $t' = n+1$.

Dla rozpatrywanej chwili odpowiadającej końcowi etapu trzeciego słuszna jest zależność, wynikająca bezpośrednio ze schematu zastępczego (rys. 12) przy założeniu $i_b = I_3$.

$$U_c[n+1] = I_3 \left(r_p + \frac{r_p R_d}{R} + R_d \right) - I_4 \left(R_d - r + \frac{r_p R_d}{R} - \frac{r r_p}{R} \right) - (E_o + U_i) \frac{r_p}{R}. \tag{31}$$

Czas trwania etapu trzeciego związany jest z liczbą impulsów n następującą zależnością:

$$t_{o3} = (n+1)\,\tau_i. \tag{32}$$

Biorąc pod uwagę zależności (31), (32) oraz (29) dla warunku: $t' = n+1$, otrzymuje się wyrażenie określające czas trwania etapu trzeciego:

$$t_{o3} = \tau_i + \frac{\tau_i}{\frac{a}{A_3} + \frac{1-a}{B_3}} \ln \times$$

$$\times \frac{1}{1-\left(e^{\frac{a}{A_3}+\frac{1-a}{B_3}}-1\right)\left[\frac{\frac{(I_3-I_4)\left[r_p+\frac{r_p R_d}{R}+R_d\right]}{(E_o+U_i)\left(1+\frac{r_p}{R}\right)-I_4\left(r_p+r+\frac{r_p r}{R}\right)}}{1-e^{-\frac{a}{A_3}-\frac{1-a}{B_3}}-\left(1-e^{-\frac{1-a}{B_3}}\right)\frac{(E_o+U_i)\frac{R}{R+R_d}+I_4\frac{R(R_d-r)}{R+R_d}}{(E_o+U_i)\left(1+\frac{r_p}{R}\right)-I_4\left(r_p+r+\frac{r_p r}{R}\right)}}-1\right]} \tag{33}$$

Przy spełnieniu warunków (23) wzór powyższy upraszcza się do postaci:

$$t_{o3} = \tau_i + \frac{K_i T_{c3}}{1+K_i\frac{T_{c3}}{T}} \ln \frac{1}{1-\left(1+K_i\frac{T_{c3}}{T}\right)\left[\frac{\frac{(I_3-I_4)(r_p+R_d)}{E_o+U_i-I_4(r_p+r)}}{\left(1+K_i\frac{T_{c3}}{T}\right)-\frac{E_o+U_i+I_4(R_d-r)}{E_o+U_i-I_4(r_p+r)}}-\frac{t_i}{T_{c3}}\right]}. \tag{34}$$

Jeśli dodatkowo spełnione są warunki:

$$t_{o3} \gg \tau_i, \quad \frac{\frac{(I_3-I_4)(r_p+R_d)}{E_o+U_i-I_4(r_p+r)}}{1+K_i\frac{T_{c3}}{T}-\frac{E_o+U_i+I_4(R_d-r)}{E_o+U_i-I_4(r_p+r)}} \gg \frac{t_i}{T_{o3}}, \tag{35}$$

które są analogiczne do warunków (25) i spełnione przy tych samych założeniach, wzór (34) przyjmuje postać:

$$t_{o3} = \frac{K_i T_{c3}}{1+K_i\frac{T_{c3}}{T}} \ln \frac{1}{1-\left(1+K_i\frac{T_{c3}}{T}\right)\frac{I_3-I_4}{\frac{K_i}{R}[E_o+U_i-I_4(r+r_p)]-I_4}}, \tag{36}$$

Krytyczna wartość współczynnika K_i dla etapu trzeciego określona jest następującym wzorem:

$$K_{ikr3} = \frac{I_3 R}{E_o+U_i-(I_3-I_4)(r_p+R_d)-I_4(r_p+r)}. \tag{37}$$

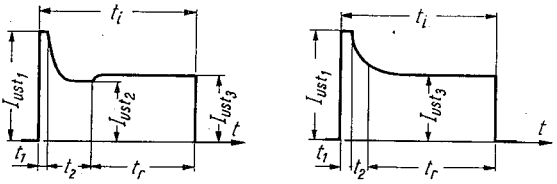
Schemat zastępczy przedstawiony na rys. 12, na podstawie którego wyprowadzony został wzór (33), narysowano w założeniu, że można pominąć składową nieustaloną prądu przepływającego przez obwód wejściowy przerzutnika w czasie trwania impulsu. Jeśli założenie to nie jest spełnione, wówczas można podać inny schemat zastępczy, w którym przerzutnik jest zastąpiony przez obwód pobierający określone porcje ładunku Q będącego funkcją współczynnika wysterowania m, a więc i napięcia na kondensatorze. Zależność $Q(m)$ wyznaczyć można wykorzystując wzory przytoczone w dodatku . Wzór uzyskany na podstawie takiego schematu zastępczego byłby jednak bardziej złożony od wzoru (33).

W praktycznych rozwiązaniach układów opóźniających ze względu na dokładność ich działania dąży się do tego, aby czwarty etap odmierzania czasu był pomijalnie mały w porównaniu z łącznym czasem trwania pierwszych trzech etapów. Osiągnąć to można przez dobór odpowiedniego (dostatecznie długiego) czasu trwania impulsu t_i (rys. 7 i wzór 2), a wtedy spełnione są już założenia, przy których został podany schemat zastępczy przedstawiony na rys. 12.

6.4. Analiza etapu czwartego

Typowe przebiegi prądu pobieranego przez obwód wejściowy przerzutnika dla etapu czwartego w czasie trwania impulsu przedstawiono na rys. 13. Wartość ładunku pobieranego przez przerzutnik w przypadku przebiegów jak na rys. 13a określa zależność:

$$Q_a = I_{ust1}\,t_1 + \\ + \int_0^{t_2} (I_{ust1} - I_{ust2}) e^{-\frac{t}{\tau_z}}\,dt + I_{ust2}\,t_2 + I_{ust3}(t_i - t_1 - t_2) - \int_0^{t_i - t_1 - t_2} (I_{ust3} - I_{ust2}) e^{-\frac{t}{\tau_2}}\,dt. \qquad (38)$$



Rys. 13. Przebiegi prądu pobieranego przez przerzutnik
a) przy dominującym wpływie pojemności złączonych tranzystorów, b) gdy wpływ tych pojemności jest pomijalnie mały

Jeśli przebieg prądu wejściowego jest taki, jak przedstawiono na rys. 13b, to zmiany prądu wejściowego zachodzą ze stałą czasową τ_z' (wzór D10) w fazie drugiej i ze stałą czasową τ_2 w fazie trzeciej (wzór D13). Ze wzorów tych wynika, że w rozważanym przypadku ($C_{Tc}R_k \ll \tau_c$) dla niedużych wartości ζ ($\zeta \leqslant 0{,}7$) różnice między tymi stałymi czasowymi są niewielkie. Wartość ładunku można więc określić zależnością:

$$Q_b = I_{ust1}\,t_1 + \int_0^{t_2 + t_r} (I_{ust1} - I_{ust3}) e^{-\frac{t}{\tau_z}}\,dt + I_{ust3}(t_i - t_1). \qquad (39)$$

Różnice między ustalonymi wartościami prądów I_{ust2} i I_{ust3} są zwykle pomijalnie małe. Przyjmując $I_{ust2} = I_{ust3}$, otrzymuje się jednakowe postacie obydwu wzorów (38), (39). Wzory powyższe po wykonaniu całkowania przyjmują postać:

$$Q(m) = (t_1 + \tau_z) I_{ust1}(m) + (t_i - t_1 - \tau_z) I_{ust3}(m). \tag{40}$$

Wyznaczenie czasu opóźnienia w oparciu o schemat zastępczy uwzględniający zmienność ładunku pobieranego przez przerzutnik przy zmianach napięcia na kondensatorze byłoby jednak kłopotliwe. Dlatego wygodniej jest obliczyć średnią wartość ładunku $Q_{śr}$ odpowiadającą zmienności współczynnika wysterowania m w zakresie $m_{kr} \dots m_z$ i na tej podstawie wyznaczyć zastępczą wartość prądu I_z pobieranego przez obwód wejściowy przerzutnika:

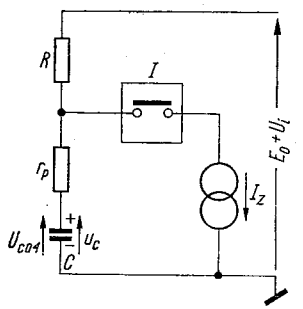
$$Q_{śr} = \frac{1}{m_z - m_{kr}} \int_{m_{kr}}^{m_z} Q(m)\,dm, \tag{41}$$

$$I_z = \frac{Q_{śr}}{t_i}. \tag{42}$$

Ze wzorów (40), (41), (42) uwzględniając zależności (D5), (D6), (D18) otrzymuje się następujące wyrażenie dla prądu I_z:

$$I_z = I_3 + \frac{\tau_s}{t_i}\left[\frac{I_{cn}}{\beta}(k-1)\frac{1}{2}\left(1 - \frac{1-d+\zeta d}{1-d}\cdot\frac{k+1-d+kd}{1+k\lambda}\right)\left(1 + \frac{1}{m_z - m_{kr}}\ln\frac{m_z-1}{m_{kr}-1} + \right.\right.$$
$$+ \frac{m_z^2}{m_z - m_{kr}}\ln\frac{m_z}{m_z-1} - \frac{m_{kr}^2}{m_z - m_{kr}}\ln\frac{m_{kr}}{m_{kr}-1}\Big) - \frac{1}{m_z - m_{kr}}\left(I_3 - \frac{I_{cn}}{\beta}(k-1)m_{kr}\times\right.$$
$$\left.\left.\times\frac{1-d+\zeta d}{1-d}\cdot\frac{k+1-d+kd}{1+k\lambda}\right)\left(m_z\ln\frac{m_z}{m_z-1} - m_{kr}\ln\frac{m_{kr}}{m_{kr}-1} + \ln\frac{m_z-1}{m_{kr}-1}\right)\right] - \frac{\tau_z}{t_i}I_3 +$$
$$+ \frac{1}{2}\frac{\tau_z}{t_i}(m_z + m_{kr})\frac{I_{cn}}{\beta}(k-1) + \frac{I_{cn}}{\beta}(k-1)\frac{m_z - m_{kr}}{2}\left(1 - \frac{\tau_z}{t_i}\right)\frac{1-d+\zeta d}{1-d}\cdot\frac{k+1-d+kd}{1+k\lambda}. \tag{43}$$

Czas trwania etapu czwartego można wyznaczyć posługując się schematem zastępczym przedstawionym na rys. 14. Schemat ten można uważać za szczególny przypadek



Rys. 14. Schemat zastępczy układu opóźniającego przyjęty do analizy etapu czwartego
I — impulsator

schematu z rys. 12 i w związku z tym można skorzystać z zależności (28), (29), (30) podstawiając w nich $R_d = \infty$ oraz I_z zamiast I_4 i U_{co4} zamiast U_{co3}. Otrzymuje się wtedy:

$$U_c(t') = E_o + U_i + \left[D_4 \frac{1-e^{-b_4 n}}{1-e^{-b_4}} + U_{co4} e^{-b_4 n} - (E_o + U_i) \right] e^{-\frac{t'-n}{A_4}}$$

$$\text{dla:} \quad n \leqslant t' \leqslant n+a, \tag{44}$$

$$u_c(t') = \left\{ E_o + U_i + \left[D_4 \frac{1-e^{-b_4 n}}{1-e^{-b_4}} + U_{co4} e^{-b_4 n} - (E_o + U_i) \right] e^{-\frac{a}{A_4}} \right\} e^{-\frac{t'-n-a}{A_4}} +$$

$$+ \left(1 - e^{-\frac{t'-n-a}{A_4}}\right)(E_o + U_i - I_z R) \quad \text{dla:} \quad n+a \leqslant t' \leqslant n+1, \tag{45}$$

przy czym:

$$\left.\begin{aligned} & t' = \frac{t}{\tau_i}, \quad T = C(R+r_p), \quad T_{c4} = T, \\ & A_4 = \frac{T}{\tau_i}, \quad B_4 = A_4, \quad b_4 = \frac{1}{A_4}, \\ & D_4 = (E_o + U_i)\left(1 - e^{-\frac{a}{A_4}}\right) e^{-\frac{1-a}{A_4}} + \left(1 - e^{-\frac{1-a}{A_4}}\right)(E_o + U_i - I_z R). \end{aligned}\right\} \tag{46}$$

Zgodnie z rozważaniami przeprowadzonymi w p. 5, etap czwarty kończy się z chwilą zadziałania przerzutnika, przy czym napięcie na kondensatorze musi zwiększyć się w tym etapie o wartość ΔU_{cz}. Wyznaczając napięcie $U_c[n+1]$ ze wzoru (45) — przez podstawienie $t' = n+1$ oraz uwzględniając dodatkowo zależności:

$$\Delta U_{cz} = u_c(n+1) - U_{co4}, \tag{47}$$

$$t_{o4} = (n+1)\tau_i, \tag{48}$$

otrzymuje się następujący wzór określający czas trwania etapu czwartego:

$$t_{o4} = \tau_i + A_4 \tau_i \ln \frac{1}{1 - \left(e^{\frac{1}{A_4}} - 1\right)\left[\dfrac{\Delta U_{cz}}{(E_o + U_i - U_{co4})\left(1 - e^{-\frac{1}{A_4}}\right) - I_z R\left(e^{\frac{1-a}{A_4}} - 1\right)} - 1 \right]}. \tag{49}$$

Przy spełnieniu warunków (23) wzór ten upraszcza się do postaci:

$$t_{o4} = \tau_i + T \ln \frac{1}{1 - \left[\dfrac{\Delta U_{cz}}{E_o + U_i - (I_3 - I_4)(R_d + r_p) - I_4(r + r_p) - \dfrac{I_z R}{K_i}} - \dfrac{\tau_i}{T} \right]}. \tag{50}$$

Wzór ten, podobnie jak poprzednie, wyznaczony został z dokładnością do $\pm \tau_i$ i można korzystać z niego przy założeniu, że $t_{o4} \gg \tau_i$. Pomijając ponadto człon $\frac{\tau_i}{T}$ w mianowniku jako mały, otrzymuje się następującą zależność przybliżoną:

$$t_{o4} = T \ln \frac{1}{1 - \dfrac{\Delta U_{cz}}{E_o + U_i - (I_3 - I_4)(R_d + r_p) - I_4(r + r_p) - \dfrac{I_z R}{K_i}}}. \tag{51}$$

Krytyczna wartość współczynnika K_i wynosi w tym przypadku:

$$K_{ikr4} = \frac{I_z R}{E_o + U_i - (I_3 - I_4)(R_d + r_p) - I_4(r + r_p) - \Delta U_{cz}}. \tag{52}$$

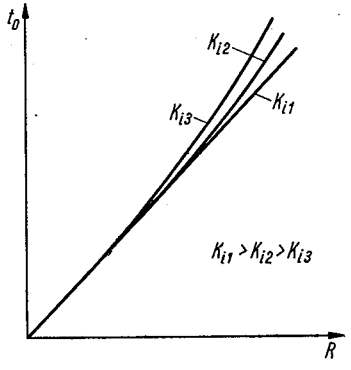
7. WPŁYW PARAMETRÓW UKŁADU NA PROCES ODMIERZANIA CZASU

Całkowity czas opóźniania analizowanego układu jest sumą czasów trwania wszystkich czterech etapów:

$$t_o = t_{o1} + t_{o2} + t_{o3} + t_{o4}. \tag{53}$$

Jak wynika z wyprowadzonych wyżej zależności, wartość tego czasu zależy od stałej czasowej obwodu RC, parametrów źródła impulsów, statycznych i dynamicznych parametrów przerzutnika, wartości napięcia E_p zasilającego obwód RC oraz od rezystancji obwodu, w którym następuje przepływ impulsów,

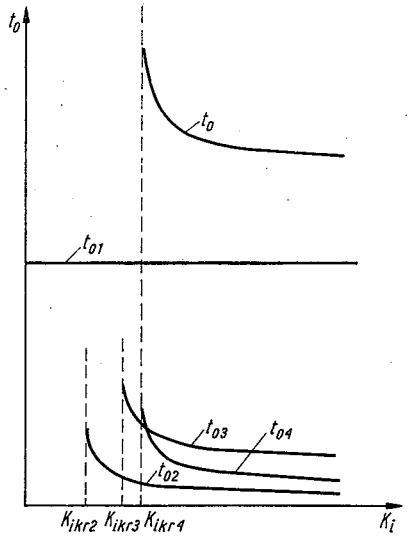
Zmiany pojemności C wywołują proporcjonalne zmiany czasu trwania każdego z etapów, natomiast zmiany rezystancji R wywołują proporcjonalne zmiany tylko czasu trwa-



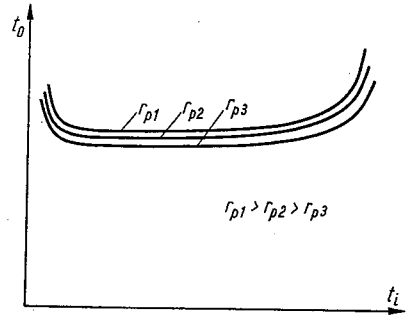
Rys. 15. Zależność czasu opóźnienia układu od rezystancji R przy różnych wartościach współczynnika K_i

nia etapu pierwszego. Ze wzrostem tej rezystancji czas trwania każdego z trzech pozostałych etapów (a więc i całkowity czas opóźnienia) wzrasta szybciej, niż proporcjonalnie (rys. 15), co wyjaśnić można oddziaływaniem przerzutnika i źródła impulsów na proces ładowania kondensatora. Oddziaływanie to jest tym silniejsze, im wartość współczynnika K_i jest bliższa krytycznej.

Zmniejszanie wartości współczynnika K_i powoduje wydłużanie czasu trwania trzech ostatnich etapów, a przy osiągnięciu wartości krytycznej nie następuje zadziałanie przerzutnika (rys. 16), wtedy bowiem ustala się równowaga między ładunkiem dostarczanym do kondensatora poprzez opornik R ze źródła zasilającego a ładunkiem odprowadzanym do obwodu wejściowego przerzutnika w czasie trwania impulsu. Ponieważ prąd ładowania kondensatora maleje w miarę wzrostu napięcia na jego okładzinach, zatem krytyczna wartość współczynnika K_i jest najmniejsza dla etapu drugiego, zaś największa



Rys. 16. Zależność czasu opóźnienia od wartości współczynnika K_i



Rys. 17. Zależność czasu opóźnienia t_o od czasu trwania impulsu t_i przy $\tau_i = \text{const}$

dla etapu czwartego. Krytyczna wartość K_i układu opóźniającego równa jest współczynnikowi K_{ikr4} określonemu przez wzór (52).

Skracając czas trwania impulsu t_i przy zachowaniu stałej wartości współczynnika K_i uzyskuje się stałe wartości czasów trwania trzech pierwszych etapów oraz wzrost czasu trwania etapu czwartego. W przypadku, gdy czas trwania impulsu jest dostatecznie długi (np. większy od $3 \div 4$ stałych czasowych τ_1 — wzór D13), etap czwarty trwa pomijalnie krótko. Przy skracaniu impulsu (np. do zakresu $1 \div 2$ stałych czasowych τ_1) wartość czasu t_{o4} może osiągnąć kilka do kilkunastu procent całkowitego czasu opóźnienia.

Na rys. 17 przedstawiono zależność całkowitego czasu opóźnienia od czasu trwania impulsu przy zachowaniu stałego okresu powtarzania impulsów.

Jak wynika z powyższego wykresu, skracanie impulsu w tym zakresie czasów, w którym wpływ dynamiki przerzutnika jest pomijalnie mały, powoduje zmniejszanie czasu

opóźnienia (przy τ_i = const. wzrasta K_i). Natomiast skracanie impulsu w zakresie, w którym wpływ dynamiki przerzutnika jest już dostatecznie duży — powoduje wzrost czasu opóźnienia. Występuje więc minimum funkcji $t_o = f(t_i)$, którego wyznaczenie w sposób analityczny nie jest jednak możliwe.

Amplituda impulsów wywiera wpływ na czasy trwania wszystkich etapów. Przy określonych parametrach układu uzyskuje się dłuższy czas opóźnienia w przypadku mniejszej amplitudy impulsów. Należy pamiętać jednak o tym, że dla zapewnienia prawidłowego działania układu niezbędne jest spełnienie warunku określonego wzorem (4).

Wartość napięcia E_p zasilającego obwód RC wpływa tylko na czas trwania etapu pierwszego. Nastawianie różnych wartości czasu opóźnienia przez zmianę tego napięcia odbywa się więc tylko na skutek zmiany czasu trwania pierwszego etapu.

Wzrost rezystancji obwodu, w którym następuje przepływ impulsów, powoduje wydłużenie czasu trwania trzech ostatnich etapów, a więc i całkowitego czasu opóźnienia (rys. 17).

Przeprowadzona dotychczas analiza pozwala na jakościową ocenę stabilności czasu opóźnienia.

Stabilność czasu trwania etapu pierwszego zależna jest głównie od stabilności (czasowej, temperaturowej itp.) elementów biernych, a więc elementów RC i diody oddzielającej ten obwód od przerzutnika. Stabilności czasów trwania pozostałych etapów zależne są ponadto od parametrów przerzutnika i źródła impulsów i są tym mniejsze, im wartość współczynnika K_i jest bliższa krytycznej. Niestabilność odmierzania czasu w etapach: drugim, trzecim i czwartym jest więc znacznie większa, niż w etapie pierwszym. W celu uzyskania możliwie dużych stabilności całkowitego czasu opóźnienia wydaje się konieczne spełnienie następujących warunków:

— dobór parametrów układu w ten sposób, aby suma czasów trwania etapów: drugiego, trzeciego i czwartego była możliwie mała w porównaniu z czasem trwania etapu pierwszego,

— zapewnienie wystarczająco dużej wartości współczynnika K_i, przynajmniej: $(1{,}5 \ldots 2{,}0) K_{ikr4}$,

— zapewnienie dużej stabilności czasu trwania etapu pierwszego przez zastosowanie wysokostabilnych elementów RC i diody o małym prądzie wstecznym.

WNIOSKI

W wyniku przeprowadzonej analizy pracy układu opóźniającego wyodrębniono cztery etapy procesu odmierzania czasu. Podział na etapy wyniknął z uwzględnienia statycznej charakterystyki wejściowej przerzutnika (pkt 3) oraz procesów przejściowych zachodzących w przerzutniku (pkt 4).

Analiza ilościowa procesu odmierzania czasu, której dokonano stosując schematy zastępcze (pkt 6) zbudowane dla poszczególnych etapów, doprowadziła do podstawowych zależności (11), (26), (36), (51) określających czasy trwania tych etapów.

W wyniku analizy powyższych zależności znaleziono rozwiązanie zagadnień (pkt 1), stanowiących podstawowy cel pracy, a w szczególności stwierdzono, co następuje.

1. Do rezystancji R obwodu RC proporcjonalny jest tylko czas trwania etapu pierwszego. Czasy trwania trzech pozostałych etapów (a więc i całkowity czas opóźnienia) nie są proporcjonalne do rezystancji R. Ze wzrostem tej rezystancji następuje więcej niż proporcjonalny wzrost czasu opóźnienia, co wynika z oddziaływania przerzutnika i źródła impulsów na proces ładowania kondensatora.

2. Czasy trwania trzech ostatnich etapów zależne są od wartości współczynnika K_i (stosunku okresu powtarzania impulsów do czasu ich trwania), przy czym dla każdego z tych etapów istnieje pewna krytyczna wartość tego współczynnika — wzory (27), (37), (52). Wartość krytyczna dla etapu czwartego K_{ikr4} jest jednocześnie wartością krytyczną dla całego układu. Przy $K_i \leqslant K_{ikr4}$ nie jest możliwa zmiana stanu przerzutnika, a więc układ opóźniający przestaje wtedy działać prawidłowo.

3. Czas trwania impulsu t_i (przy $K_i = \text{const}$) ma również wpływ na wartość czasu opóźnienia, przy czym wpływ ten wywierany jest tylko na czas trwania etapu czwartego.

DODATEK
DYNAMICZNE PARAMETRY PRZERZUTNIKA

Analizę procesów przejściowych zachodzących w przerzutniku o szeregowo-równoległym sprzężeniu zwrotnym wykonano na podstawie zastępczego schematu ładunkowego tranzystorów, wykorzystując następujące parametry tego schematu:

τ_c — stała czasowa kolektora

τ_s — stała czasu nasycenia,

C_{Tc}, C_{Te} — pojemności złącza kolektorowego i emiterowego,

β — stałoprądowy współczynnik wzmocnienia w układzie OE.

Równocześnie wprowadzono następujące założenia upraszczające:

a) założono pełną symetrię układu przerzutnika oraz jednakowe statyczne i dynamiczne parametry obu tranzystorów ($R_{k1} = R_{k2} = R_k$, $R_1 = R_3$, $R_2 = R_4$, $\beta_1 = \beta_2 = \beta$, $\tau_{c1} = \tau_{c2} = \tau_c$, $C_{Tc1} = C_{Tc2} = C_{Tc}$, $C_{Te1} = C_{Te2} = C_{Te}$),

b) pominięto prądy zerowe a także spadki napięć na złączach $E-K$ i $E-B$ tranzystorów nasyconych oraz na złączach $E-B$ tranzystorów pracujących w obszarze aktywnym,

c) założono, że prąd płynący przez obwód bazy nasyconego tranzystora jest znacznie większy od prądu płynącego przez opornik R_2 (lub R_4), oraz że: R_{k1}, R_{k2}, $R_e \ll R_1$, R_2, R_3, R_4.

Założenia powyższe nie wywierają większego wpływu na jakościowy przebieg zjawisk, a umożliwiają uzyskanie stosunkowo prostych zależności, pozwalających na przedyskutowanie wpływu różnych parametrów układu przerzutnika na jego własności dynamiczne.

Ze względu na to, że wyprowadzenia odnośnych zależności są zbyt długie, niżej przedstawiono jedynie wyniki dokonanej analizy. Zostały one potwierdzone na drodze eksperymentalnej, przy czym uzyskano zgodność wyników obliczeń i doświadczeń w granicach 20...30%.

W celu ilościowego opisu procesów przejściowych wprowadzono następujące definicje:

$$I_{cn} = \frac{E_k}{R_k + R_e}, \quad d = \frac{R_e}{R_k + R_e}, \quad k = \frac{\beta I_b}{I_{cn}}, \quad \zeta = \frac{1}{1 + \dfrac{R_e}{R_b}}, \quad \psi = \frac{R_1}{R_2}, \tag{D1}$$

gdzie I_{cn}, I_b oznaczają prąd kolektorowy i prąd bazy nasyconego tranzystora, k — współczynnik nasycenia, zaś ζ jest współczynnikiem uwzględniającym rezystancję źródła sterującego.

Ponadto wprowadzono współczynnik wysterowania m zdefiniowany zgodnie z rys. 3, jako:

$$m = \frac{\Delta u_{we}}{\Delta U_{we1}}. \tag{D2}$$

Wartość napięcia ΔU_{we1} oraz krytyczną wartość współczynnika m wyznaczyć można ze wzorów:

$$\Delta U_{we1} = \frac{I_{cn}}{\beta}(k-1)(R_b+R_k d), \quad m_{kr} = 1+\frac{\beta(1-\zeta)\psi d}{(k-1)(1-d+\zeta d)}. \tag{D3}$$

W celu przeliczenia wartości współczynnika m na odpowiadającą mu różnicę Δu_{we} pomiędzy rzeczywistą wartością napięcia wejściowego a wartością odpowiadającą statycznemu progowi zwalniania przerzutnika można skorzystać ze wzoru:

$$\Delta u_{we} = (m-m_{kr})\frac{E_k d(k-1)(1-d+\zeta d)}{\beta(1-\zeta)}. \tag{D4}$$

Czas trwania pierwszej fazy procesów przejściowych t_1 oraz wartość prądu I_{ust1} pobieranego przez obwód wejściowy przerzutnika wynoszą odpowiednio:

$$t_1 = \tau_s \ln\frac{m}{m-1}, \tag{D5}$$

$$I_{ust1} = \frac{I_{cn}}{\beta} m(k-1). \tag{D6}$$

Czas trwania drugiej fazy wyznaczono dla dwóch skrajnych przypadków, a mianowicie:

1) przy dominującym wpływie pojemności złączowych ($C_{Tc}R_k \gg \tau_c$),

2) gdy wpływ tych pojemności na przebiegi przejściowe jest pomijalnie mały ($\tau_c \gg C_{Tc}R_k$).

W pierwszym przypadku czas trwania drugiej fazy t_2 oraz stała czasowa τ_z, z jaką zachodzą zmiany prądu pobieranego przez obwód wejściowy przerzutnika wynoszą odpowiednio:

$$t_2 = \frac{R_1 R_2}{R_1+R_2}(C_{Tc}+C_{Te})\ln\left[1+\frac{\psi d[(1-\zeta)\beta+1]-\dfrac{C_{Tc}}{C_{Tc}+C_{Te}}\zeta d(m-1)(k-1)(1+\psi)}{(m-m_{kr})(k-1)(1-d+\zeta d)}\right], \tag{D7}$$

$$\tau_z = \frac{\tau_c + C_{Tc}R_k\left(1+\dfrac{\zeta d}{1-d}\right)}{\dfrac{1}{\beta}+1-\zeta}, \tag{D8}$$

zaś dla drugiego przypadku uzyskano następujące zależności:

$$t_2' = \tau_z' \ln\left[1+\frac{\psi d[1+\beta(1-\zeta)]}{(m-m_{kr})(k-1)(1-d+\zeta d)}\right], \tag{D9}$$

$$\tau_z' = \frac{\tau_c}{\dfrac{1}{\beta}+1-\zeta}. \tag{D10}$$

Dla przebiegów prądów kolektorowych w trzeciej fazie procesów przejściowych uzyskano zależności następujące:

$$i_{k1} = A_1 - Be^{\frac{t}{\tau_1}} + C_1 e^{-\frac{t}{\tau_2}}, \tag{D11}$$

$$i_{k2} = A_2 + Be^{\frac{t}{\tau_1}} + C_2 e^{-\frac{t}{\tau_2}}, \tag{D12}$$

gdzie:

$$\tau_1 = \frac{\beta(\tau_c+C_{Tc}R_k)}{k-1}; \quad \tau_2 = \frac{\beta(\tau_c+C_{Tc}R_k)}{1+k\lambda}; \quad \lambda = 1+2\zeta^2\frac{d}{1-d}+(1-\zeta)\left(\frac{\beta}{k}-1\right), \tag{D13}$$

zaś współczynniki A_1, A_2, B, C_1, C_2 są funkcjami parametrów: ζ, d, β, k, m, ψ.

W przypadku, gdy impuls oddziaływujący na przerzutnik zanika w trzeciej fazie procesów przejściowych przy wartościach prądów kolektorowych równych odpowiednio: $i_{k1}^{w} = \frac{i_{k1}}{I_{cn}} = \delta_1$, $i_{k2}^{w} = \frac{i_{k2}}{I_{cn}} = \delta_2$, wówczas po zaniku tego impulsu przebiegi prądów i_{k1}, i_{k2} opisane są następującymi zależnościami:

$$i_{k1}^{w} = \frac{k}{(1-d)\left(1+k\frac{1+d}{1-d}\right)} - \frac{\delta_2-\delta_1}{2} e^{\frac{(k-1)t}{\beta(\tau_c+C_{Tc}R_k)}} + \\ + \frac{1}{2}\left(\delta_1+\delta_2-1-\frac{k-1}{1+k\frac{1+d}{1-d}}\right) e^{-\left(1+k\frac{1+d}{1-d}\right)\frac{t}{\beta(\tau_c+C_{Tc}R_k)}}, \tag{D14}$$

$$i_{w}^{k2} = \frac{k}{(1-d)\left(1+k\frac{1+d}{1-d}\right)} + \frac{\delta_2-\delta_1}{2} e^{\frac{(k-1)t}{\beta(\tau_c+C_{Tc}R_k)}} + \\ + \frac{1}{2}\left(\delta_1+\delta_2-1-\frac{k-1}{1+k\frac{1+d}{1-d}}\right) e^{-\left(1+k\frac{1+d}{1-d}\right)\frac{t}{\beta(\tau_c+C_{Tc}R_k)}}. \tag{D15}$$

Aby po zaniku impulsu wejściowego przerzutnik zmienił swój stan na przeciwny, prąd i_{k1} powinien maleć, i_{k2} — rosnąć. Człon z wykładnikiem dodatnim w wyrażeniu (D14) powinien być ujemny, zaś w wyrażeniu (D15) — dodatni. A zatem powinien być spełniony warunek:

$$\delta_2 > \delta_1. \tag{D16}$$

W celu wyznaczenia czasu t_r, po upływie którego zostanie spełniony warunek (D16), skorzystano z wyrażeń (D11), (D12) pomijając człony z wykładnikami ujemnymi (dla $t = t_r$ wartości ich są pomijalnie małe), uzyskując następujące wyrażenie:

$$t_r = \frac{\beta(\tau_c+C_{Tc}R_k)}{k-1} \ln \frac{(1+\lambda)(1-d)}{2(1+k\lambda)(1-d+\zeta d)(m-m_{kr})} \left\{(1-\zeta)\left[\beta+1-d+(m-1)(1-d+kd)+\right.\right. \\ - \frac{2kd\zeta}{1-d} - \frac{\psi kd\beta}{(1-d)(k-1)} - \frac{\psi kd(1-d-2kd\zeta)}{(k-1)(1-d)^2}\left.\right] + \zeta\frac{(1-d+kd)m+k}{1-d} - \zeta\frac{\psi k^2 d}{(k-1)(1-d)^2} + \\ \left. + k\left[\frac{1-d+\zeta d}{1-d}(m-1) - (1-\zeta)\frac{\psi d}{(k-1)(1-d)}\left(\beta - \frac{2kd\zeta}{1-d}\right) - \psi d\frac{1-d+kd\zeta}{(k-1)(1-d)^2}\right]\right\}. \tag{D17}$$

Ustalona wartość prądu pobieranego przez obwód wejściowy przerzutnika w fazie trzeciej określona jest wzorem:

$$I_{ust3} = I_3 + (m-m_{kr})\frac{I_{cn}}{\beta}(k-1)\frac{1-d+\zeta d}{1-d}\cdot\frac{k+1-d+kd}{1+k\lambda}. \tag{D18}$$

LITERATURA CYTOWANA W TEKŚCIE

1. I. Z. Cypkin, Teoria układów impulsowych, PWN, Warszawa 1965.
2. W. Kiełek, Tranzystorowy przerzutnik bistabilny — analiza dynamiczna, Rozdz. 9 w pracy zbiorowej pod kier. S. Ryżki — Elektroniczne mierniki zliczające, WKŁ, Warszawa 1965.
3. R. A. Lipman, Przekaźniki tranzystorowe, WNT, Warszawa 1965.
4. H. Mroczek, Tranzystorowe elementy opóźniające w bezstykowych układach sterowania, Przegląd Elektrotechniczny, 1966, nr 12, s. 509—513.

5. H. Mroczek, Analiza charakterystyki wejściowej przerzutnika tranzystorowego o szeregowo-równoległym sprzężeniu zwrotnym. Materiały Sesji Nauk.: 25-lecie Wydziału Elektrycznego Politechniki Łódzkiej, z. 1, Automatyka i Elektronika, s. 77—92.
6. I. Pinto, Semiconductor Time-Delay Circuits, Electronic Engineering, March 1964, s. 166—169.
7. Praca zbiorowa, Halbleiter — Schaltbeispiele, Siemens & Halske A. G., April 1966.

H. MROCZEK

THE ANALYSIS OF LONG TIME DELAY TRANSISTOR CIRCUIT

Summary

In this paper the analysis of a long time delay circuit is made. The circuit analysed consists of *RC*-elements, an auxiliary pulse source and a bipolar transistor bistable multivibrator with series-parallel feedback. This circuit enables the use of a large value of resistance R (tens MΩ), despite of current mode control of the bipolar transistors.

As a result of this analysis, formulas were obtained, which determine the value of the time delay, as a function of the pulse series parameters and static and dynamic parameters of the multivibrator.

Four stages of the time delay process are isolated in which the physical phenomena are qualitatively different and it is stated that in the individual stages the value of time delay depends on the impulse series parameters, *RC*-value, and on the input multivibrator characteristic parameters, but the time delay of the fourth stage depends additionally on dynamic parameters of the multivibrator.

H. MROCZEK

ANALYSE DE CONDITIONS DE TRAVAIL D'UN SYSTEME TRANSISTORISÉ DE DÈCALAGE À LONG TEMPS DU RETARD

Résumé

L'analyse d'un système de décalage conçu pour la réalisation de longs temps du retard, constitué d'un circuit — *RC*, d'une source auxiliaire des impulsions et d'un bistable transistorisé. Le systeme, malgré la commande des transistors bipolaires effectuée par courants, permet l'application des grandes resistances R (même quelques dizaines de MΩ).

Les valeurs du temps du retard en fonction des valeurs caractéristiques du train des impulsions et des paramètres statiques et dynamiques de bistable sont données.

H. MROCZEK

ANALYSE VON DIMENSIONIERUNGSBEDINGUNGEN EINER TRANSISTORSCHALTUNG BEI LANGEN VERZÖGERUNGSZEITEN

Zusammenfassung

In der Arbeit wurde die Analyse einer aus dem *RC*-Glied, der Impulsquelle und dem bistabilen Multivibrator zusammengesetzten Transistorschaltung bei langen Verzögerungszeiten durchgeführt. Diese Schaltung ermöglicht die Anwendung großwertiger R-Widerstände (Größenreihe von Zehnern von MΩ) trotz Stromsteuerung mittels Bipolartransistore, die Bauteile der Multivibratoren sind. Es wurden auch die Zusammenhänge zwischen dem Wert der Verzögerungszeit und den statischen und dynamischen Parametern des Multivibrators und auch der Impulsquelle bestimmt.

Г. МРОЧЕК

АНАЛИЗ УСЛОВИЙ РАБОТЫ ТРАНЗИСТОРНОЙ СХЕМЫ ЗАПАЗДЫВАНИЯ РЕАЛИЗУЮЩЕЙ БОЛЬШИЕ ВЫДЕРЖКИ ВРЕМЕНИ

Р е з ю м е

В работе проведен анализ схемы запаздывания способной реализовать большие выдержки времени собранной из контура *RC*, дополнительного источника импульсов и транзисторного бистабильного триггера с коллекторно-эмиттерной связью.

Эта схема предоставляет возможность применения больших значений резистанса *R* (порядка нескольких десятков MΩ), помимо токового характера управления биполярных транзисторов, из которых собран триггер.

Определены зависимости связывающие время запаздывания с параметрами ряда импульсов и статическими и динамическими параметрами триггера. В анализе выделены четыре этапа, в которых ход явлений качественно различен. Констатировано, что времена выдержки отдельных этапов зависят от параметров ряда импульсов, контура *RC* и параметров определяющих статическую входную характеристику триггера, причем время выдержки четвертого этапа зависят дополнительно от динамических параметров триггера.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 2, ss. 295—312

621.372.54

Zarys teorii syntezy filtrów elektromechanicznych i mikrofalowych o strukturze regularnej

FRANCISZEK KAMIŃSKI (WARSZAWA)
Instytut Fizyki PAN

Otrzymano 2.4.1973

Przedstawiono metodę analizy właściwości transmisyjnych czwórników o strukturze regularnej w zastosowaniu do filtrów elektromechanicznych i mikrofalowych. Podano określenie struktury regularnej I i II rodzaju oraz rozpatrzono konkretne przykłady w postaci filtrów elektromechanicznych ze sprzęgaczami prostymi, złożonymi i kombinowanymi oraz o strukturze mieszanej regularnej. Analizę prowadzi się w oparciu o unormowaną macierz transmisyjną czwórnika. Wyprowadzono ścisłe związki analityczne, określające funkcję transmisji i funkcję filtracji badanych układów. Uzyskane wyniki służą do syntezy filtrów o strukturze jednorodnej i mieszanej regularnej, umożliwiając kontrolę zarówno charakterystyki tłumieniowej, jak i fazowej czwórnika.

1. WSTĘP

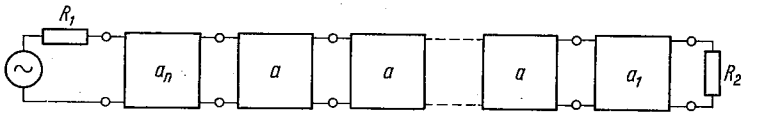
Potrzeba analizy właściwości transmisyjnych czwórnika o strukturze regularnej wynika z dwojakich przesłanek. Z jednej strony czwórniki tego typu znajdują zastosowanie praktyczne, jak np. filtry elektromechaniczne o strukturze regularnej [12], co związane jest z łatwością ich realizacji. Z drugiej strony wykorzystuje się je jako układy odniesienia (inaczej nazywając układy wzorcowe) przy syntezie filtrów elektromechanicznych i mikrofalowych metodą algebraiczną [6—9]. Z tych powodów właściwościom układów regularnych poświęcono wiele opracowań, a w szczególności [1—4, 11—14].

W niniejszej pracy omówiono w zarysie teorię syntezy filtrów elektromechanicznych i mikrofalowych o strukturze regularnej. W trakcie rozważań wykorzystuje się rezultaty uzyskane w pracach [1—4]. Przy opracowywaniu niniejszej publikacji starano się w pierwszym rzędzie uwypuklić zagadnienia nowe bądź niedostatecznie wszechstronnie omówione w literaturze.

2. UNORMOWANA MACIERZ TRANSMISYJNA CZWÓRNIKA O STRUKTURZE REGULARNEJ

Czwórnik o strukturze regularnej (rys. 1) składa się z łańcucha jednakowych czwórników z tym, że czwórniki skrajne mogą posiadać właściwości odmienne od pozostałych. Z tego względu wprowadza się podział układów regularnych na dwa rodzaje. Czwórnik o strukturze regularnej I rodzaju obejmuje wyłącznie czwórniki identyczne. Jeżeli nato-

miast czwórniki skrajne są odmienne od pozostałych w łańcuchu, to wtedy układ określa się mianem czwórnika o strukturze regularnej II rodzaju.



Rys. 1. Czwórnik o strukturze regularnej

Zgodnie z [1], str. 84, macierz łańcuchowa $\mathbf{a}'_n$ układu regularnego I rodzaju, złożonego z czwórników odwracalnych o macierzy łańcuchowej $\mathbf{a}$, wynosi

$$(\mathbf{a}'_n)_{\mathrm{I}} = \mathbf{a}^n = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^n = \begin{bmatrix} P_n(\varphi)+\phi U_{n-1}(\varphi) & a_{12}U_{n-1}(\varphi) \\ a_{21}U_{n-1}(\varphi) & P_n(\varphi)-\phi U_{n-1}(\varphi) \end{bmatrix}, \tag{1}$$

gdzie:

$$P_n(\varphi) = \mathrm{ch}(n\,\mathrm{Arch}\,\varphi), \qquad U_{n-1}(\varphi) = \frac{\mathrm{sh}(n\,\mathrm{Arch}\,\varphi)}{\mathrm{sh}(\mathrm{Arch}\,\varphi)},$$
$$\varphi = \frac{1}{2}(a_{11}+a_{22}), \qquad \phi = \frac{1}{2}(a_{11}-a_{22}). \tag{2}$$

Funkcja $P_n(x)$ jest wielomianem Czebyszewa I rodzaju n-tego stopnia względem zmiennej x; $U_{n-1}(x)$ jest wielomianem Czebyszewa II rodzaju $(n-1)$-ego stopnia względem zmiennej x.

Identyczny wzór obowiązuje w przypadku obliczania macierzy $\mathbf{T}_n$, będącej iloczynem n jednakowych unormowanych macierzy transmisyjnych $\mathbf{T}$ [1, 10], to znaczy

$$(\mathbf{T}_n^z)_{\mathrm{I}} = \mathbf{T}^n = \begin{bmatrix} P_n(\varphi)+\phi U_{n-1}(\varphi) & T_{12}U_{n-1}(\varphi) \\ T_{21}U_{n-1}(\varphi) & P_n(\varphi)-\phi U_{n-1}(\varphi) \end{bmatrix},$$
$$\varphi = \frac{1}{2}(T_{11}+T_{22}), \qquad \phi = \frac{1}{2}(T_{11}-T_{22}). \tag{3}$$

W przypadku układu regularnego II rodzaju korzysta się ze wzoru (1) celem określenia właściwości łańcucha czwórników z pominięciem macierzy a_1 i a_n, czyli że

$$(\mathbf{a}'_n)_{\mathrm{II}} = \mathbf{a}_n\mathbf{a}^{n-2}\mathbf{a}_1 = (\mathbf{a}_n\mathbf{a}^{-1})\mathbf{a}^n(\mathbf{a}^{-1}\mathbf{a}_1). \tag{4}$$

Przytoczony sposób zapisu wskazuje na możliwość uproszczenia macierzy $(\mathbf{a}'_n)_{\mathrm{II}}$, jeżeli $\mathbf{a}_n\mathbf{a}^{-1} \cong \mathbf{E}+\Delta\mathbf{a}_n$, $\mathbf{a}^{-1}\mathbf{a}_1 \cong \mathbf{E}+\Delta\mathbf{a}_1$ w roboczym pasmie częstotliwości, gdzie $\mathbf{E}$ — macierz jednostkowa, $\Delta\mathbf{a}_1$, $\Delta\mathbf{a}_n$ — macierze o elementach do pominięcia w pierwszym przybliżeniu.

Celem wyznaczenia macierzy transmisyjnej $\mathbf{T}_n$ czwórnika o strukturze regularnej przy obciążeniu rezystywnym R_1, R_2 można skorzystać ze wzoru [1, 10]

$$\mathbf{T}_n = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} \sqrt{\frac{R_2}{R_1}}\,(a'_{11})_n & \frac{(a'_{12})_n}{\sqrt{R_1R_2}} \\ \sqrt{R_1R_2}\,(a'_{21})_n & \sqrt{\frac{R_1}{R_2}}\,(a'_{22})_n \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \tag{5}$$

w którym uwzględnia się elementy macierzy $\mathbf{a}'_n$ według (1) lub (4).

Można też postąpić inaczej, a mianowicie wprowadzić w miejsce macierzy łańcuchowej **a** macierz transmisyjną **T** ogniwa podstawowego i skorzystać z twierdzenia [1, 10], że macierz transmisyjna łańcucha czwórników równa się iloczynowi macierzy transmisyjnych poszczególnych czwórników. Zatem wzór (3) jest ważny dla unormowanej macierzy transmisyjnej $\mathbf{T}_n$ czwórnika o strukturze regularnej I rodzaju. Stwierdzenie to jest istotne z tego względu, iż elementy macierzy transmisyjnej są ściśle związane z syntezą czwórnika w oparciu o jego parametry skuteczne. W szczególności element T_{21} unormowanej macierzy transmisyjnej jest funkcją filtracji czwórnika, a moduł elementu T_{11}, określanego mianem funkcji transmisji, stanowi charakterystykę amplitudową czwórnika.

Analiza właściwości układów regularnych za pomocą unormowanej macierzy transmisyjnej stanie się bardziej przejrzysta na przykładzie konkretnych struktur filtrów elektromechanicznych i mikrofalowych, spośród których wybrano filtry o strukturze jednorodnej regularnej i o strukturze mieszanej regularnej. Przez pojęcie *filtr elektromechaniczny o strukturze jednorodnej* rozumie się w danym przypadku to, iż w linii filtrującej nie występują dwójniki o stałych skupionych lub o selekcji skupionej (z wyjątkiem być może układu na wejściu i wyjściu filtru). Jest to równoznaczne z tym, że filtr elektromechaniczny o tej strukturze można rozpatrywać jako tor niejednorodny złożony z odcinków jednorodnych [5, 10]. Natomiast pojęcie *filtr elektromechaniczny* (*mikrofalowy*) *o strukturze mieszanej* odnosi się do układu złożonego z łańcucha odcinków jednorodnej linii długiej, współpracujących z dwójnikami o selekcji skupionej [1, 6].

Należy jeszcze zaznaczyć, że w dalszym ciągu pracy określenie *macierz transmisyjna* dotyczy wyłącznie unormowanej macierzy transmisyjnej czwórnika.

3. UNORMOWANA MACIERZ TRANSMISYJNA FILTRU ELEKTROMECHANICZNEGO O STRUKTURZE JEDNORODNEJ REGULARNEJ I RODZAJU

3.1. Wprowadzenie

Filtr elektromechaniczny o strukturze jednorodnej regularnej I rodzaju daje się schematycznie przedstawić jak na rys. 2. Czwórnik ten można rozbić na ogniwa podstawowe, złożone z dwu połówek rezonatora przedzielonych układem sprzęgającym (rys. 3). Na podstawie danych w [2—5] macierz transmisyjna tego filtru daje się napisać następująco:

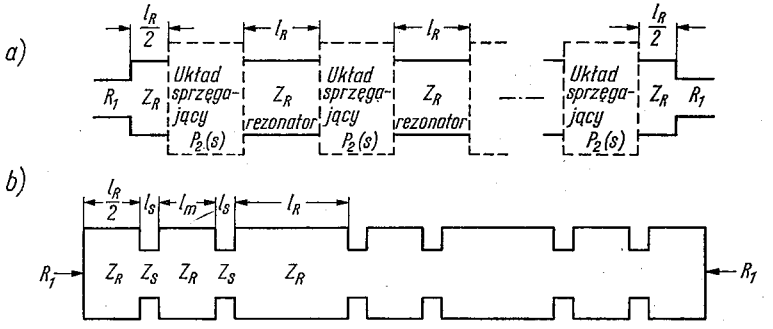
$$\mathbf{T}_n(s) = \mathbf{M}_1\mathbf{J}\left(\frac{s}{2}\right)\mathbf{P}_2(s)\mathbf{J}(s)\mathbf{P}_2(s)\ldots\mathbf{J}(s)\mathbf{P}_2(s)\mathbf{J}\left(\frac{s}{2}\right)\mathbf{M}_1^{-1} = \mathbf{T}_0^n,$$

$$\mathbf{T}_0 = \mathbf{M}_1\mathbf{J}\left(\frac{s}{2}\right)\mathbf{P}_2(s)\mathbf{J}\left(\frac{s}{2}\right)\mathbf{M}_1^{-1}, \tag{6}$$

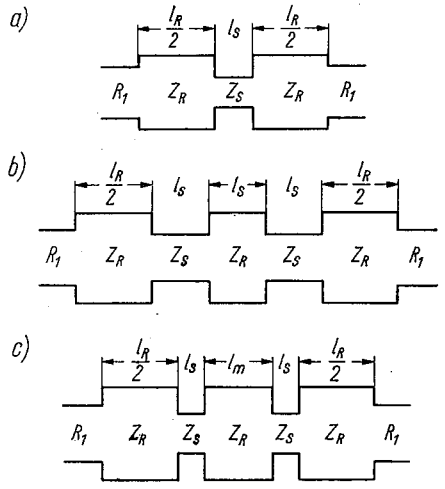
gdzie n jest liczbą ogniw podstawowych w łańcuchu.

Macierz $\mathbf{P}_2(s)$ jest macierzą transmisyjną układu sprzęgającego o postaci

$$\mathbf{P}_2(s) = \mathbf{M}_2\mathbf{J}(rs)\mathbf{M}_2^{-1}\mathbf{J}(qs)\mathbf{M}_2\mathbf{J}(rs)\ldots\mathbf{M}_2\mathbf{J}(rs)\mathbf{M}_2^{-1}\mathbf{J}(qs)\mathbf{M}_2\mathbf{J}(rs)\mathbf{M}_2^{-1}. \tag{7}$$



Rys. 2. Filtr elektromechaniczny o strukturze jednorodnej regularnej I rodzaju z dowolnym układem sprzęgającym o macierzy transmisyjnej $\mathbf{P}_2(s)$ (*a*) oraz ze sprzęgaczem kombinowanym (b))



Rys. 3. Przykłady ogniw podstawowych w filtrach elektromechanicznych o strukturze jednorodnej regularnej

a) ogniwo podstawowe ze sprzęgaczem prostym, b) ogniwo podstawowe ze sprzęgaczem złożonym dla $p = 3$, c) ogniwo podstawowe ze sprzęgaczem kombinowanym dla $u = 1$

Macierze typu **M** i **J** określa się za pomocą następujących zależności:

$$\mathbf{M}_i = \begin{bmatrix} \mathrm{ch}\nu_i & \mathrm{sh}\nu_i \\ \mathrm{sh}\nu_i & \mathrm{ch}\nu_i \end{bmatrix}, \quad i = 1, 2; \quad \nu_1 = \frac{1}{2}\ln\frac{Z_R}{R_1}, \quad \nu_2 = \frac{1}{2}\ln\frac{Z_S}{Z_R}, \tag{8}$$

$$\mathbf{J}(s) = \begin{bmatrix} e^{s} & 0 \\ 0 & e^{-s} \end{bmatrix}, \quad \mathbf{J}(rs) = \begin{bmatrix} e^{rs} & 0 \\ 0 & e^{-rs} \end{bmatrix}, \tag{9}$$

przy czym:

$$s = \sigma + j\Theta, \quad \Theta = 2\pi\frac{l_R}{\lambda}, \quad r = \frac{l_s}{l_R}, \quad q = \frac{l_m}{l_R}, \tag{10}$$

gdzie:

Z_R — impedancja falowa rezonatora,

Z_S — impedancja falowa skrajnych odcinków jednorodnych w układzie sprzęgającym,

R_1 — opór obciążenia,
l_R — długość rezonatora,
l_s — długość skrajnych odcinków jednorodnych w sprzęgaczu,
l_m — długość odcinka jednorodnego, występującego między dwoma odcinkami o długości l_s,
Θ — przesuwność falowa rezonatora,
λ — długość fali w linii.

Macierz $\mathbf{J}(qs)$ uzyskuje się z macierzy $\mathbf{J}(rs)$ zastępując współczynnik r przez q.

Macierz $\mathbf{P}_2(s)$ opisuje właściwości układu sprzęgającego, który jest sprzęgaczem kombinowanym, jeżeli $r \neq q$, lub sprzęgaczem złożonym, jeżeli $r = q$. W przypadku szczególnym, kiedy we wzorze (7) występuje tylko jeden czynnik typu $\mathbf{J}$, macierz $\mathbf{P}_2(s)$ opisuje właściwości transmisyjne sprzęgacza prostego. W tym miejscu warto zwrócić uwagę na fakt, iż sprzęgacze złożone oraz rozbudowane sprzęgacze kombinowane również należą do układów o strukturze regularnej.

Jak wynika ze związków (6), (7), zadanie wyznaczenia unormowanej macierzy transmisyjnej filtru o strukturze regularnej I rodzaju sprowadza się do obliczenia macierzy $\mathbf{T}_0$ dla zadanej konstrukcji sprzęgacza i następnie do podstawienia uzyskanego wyniku do wzoru (3).

3.2. Unormowana macierz transmisyjna filtru elektromechanicznego o strukturze jednorodnej regularnej I rodzaju ze sprzęgaczami prostymi

W przypadku sprzęgacza prostego macierz $\mathbf{P}_2(s)$ wynosi

$$\mathbf{P}_2(s) = \mathbf{M}_2\mathbf{J}(rs)\mathbf{M}_2^{-1},$$

a zatem

$$\mathbf{T}_0 = \begin{bmatrix} T_{11}^0 & T_{12}^0 \\ T_{21}^0 & T_{22}^0 \end{bmatrix} = \mathbf{M}_1\mathbf{J}\left(\frac{s}{2}\right)\mathbf{M}_2\mathbf{J}(rs)\mathbf{M}_2^{-1}\mathbf{J}\left(\frac{s}{2}\right)\mathbf{M}_1^{-1}, \tag{11}$$

skąd

$$\begin{aligned} T_{11}^0 &= e^s\mathrm{ch}^2\nu_1(e^{rs}\mathrm{ch}^2\nu_2 - e^{-rs}\mathrm{sh}^2\nu_2) + \mathrm{sh}2\nu_1\,\mathrm{sh}2\nu_2\,\mathrm{sh}rs - \\ &\quad - e^{-s}(e^{-rs}\mathrm{ch}^2\nu_2 - e^{rs}\mathrm{sh}^2\nu_2)\mathrm{sh}^2\nu_1, \qquad T_{22}^0(s) = T_{11}^0(-s), \end{aligned} \tag{12}$$

$$\begin{aligned} T_{21}^0 &= \mathrm{sh}2\nu_1\mathrm{ch}^2\nu_2\,\mathrm{sh}(1+r)s + \mathrm{ch}2\nu_1\,\mathrm{sh}2\nu_2\,\mathrm{sh}rs - \mathrm{sh}2\nu_1\mathrm{sh}^2\nu_2\,\mathrm{sh}(1-r)s, \\ T_{12}^0(s) &= T_{21}^0(-s) = -T_{21}^0(s). \end{aligned} \tag{13}$$

Po podstawieniu rezultatu (12) do (3) uzyskuje się

$$\begin{aligned} \varphi &= \mathrm{ch}^2\nu_2\,\mathrm{ch}(1+r)s - \mathrm{sh}^2\nu_2\,\mathrm{ch}(1-r)s, \\ \phi &= \mathrm{ch}2\nu_1\,\mathrm{ch}^2\nu_2\,\mathrm{sh}(1+r)s + \mathrm{sh}2\nu_1\,\mathrm{sh}2\nu_2\,\mathrm{sh}rs - \mathrm{ch}2\nu_1\,\mathrm{sh}^2\nu_2\,\mathrm{sh}(1-r)s \end{aligned} \tag{14}$$

oraz

$$(T_{11})_n = P_n(\varphi) + \phi U_{n-1}(\varphi), \qquad (T_{21})_n = T_{21}^0\,U_{n-1}(\varphi). \tag{15}$$

Wzory (14), (15) służą do określania charakterystyki tłumieniowej i fazowej filtru elektromechanicznego o strukturze jednorodnej regularnej I rodzaju ze sprzęgaczami prostymi przy obciążeniu rezystywnym. Są to zależności ścisłe, ważne dla dowolnej liczby ogniw podstawowych n oraz dla dowolnego stosunku długości elementu sprzęgającego do długości rezonatora.

3.3. Unormowana macierz transmisyjna filtru elektromechanicznego o strukturze jednorodnej regularnej I rodzaju ze sprzęgaczami złożonymi

Unormowana macierz transmisyjna sprzęgacza złożonego, obejmującego p ($p = 1, 3, 5, \ldots$) elementów o jednakowej długości, ma postać

$$\mathbf{P}_2(s) = \mathbf{M}_2\mathbf{J}(rs)\mathbf{M}_2^{-1}\mathbf{J}(rs)\mathbf{M}_2 \ldots \mathbf{J}(rs)\mathbf{M}_2\mathbf{J}(rs)\mathbf{M}_2^{-1}, \tag{16}$$

przy czym liczba macierzy typu $\mathbf{J}$ wynosi p, a macierzy typu $\mathbf{M}$ — $p+1$.

W iloczynie (16) dają się wyodrębnić następujące czynniki podstawowe:

$$\mathbf{P}' = \mathbf{M}_2'\mathbf{J}(rs)\mathbf{M}_2'^{-1} = \mathbf{P}\bar{\mathbf{E}}, \quad \mathbf{P}_t' = \mathbf{M}_2'^{-1}\mathbf{J}(rs)\mathbf{M}_2' = \bar{\mathbf{E}}\mathbf{P}, \tag{17}$$

które występują na przemian. Macierze we wzorze (17) wynoszą

$$\bar{\mathbf{E}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{M}_2'\mathbf{M}_2' = \mathbf{M}_2,$$
$$\mathbf{P} = \begin{bmatrix} \mathrm{ch}\nu_2\,\mathrm{sh}rs + \mathrm{ch}rs & \mathrm{sh}\nu_2\,\mathrm{sh}rs \\ \mathrm{sh}\nu_2\,\mathrm{sh}rs & \mathrm{ch}\nu_2\,\mathrm{sh}rs - \mathrm{ch}rs \end{bmatrix}. \tag{18}$$

Podstawienie zależności (17) do (16) przynosi

$$\mathbf{P}_2(s) = \mathbf{M}_2'\mathbf{P}^p\bar{\mathbf{E}}\mathbf{M}_2'^{-1} = j^p\mathbf{M}_2'(-j\mathbf{P})\mathbf{M}_2'\bar{\mathbf{E}} = j^p\mathbf{P}_z\bar{\mathbf{E}}. \tag{19}$$

Wyznacznik macierzy $(-j\mathbf{P})$ jest równy jedności, przeto wyrażenie $(-j\mathbf{P})^p$ daje się obliczyć na podstawie zależności podstawowej (1), skąd

$$(-j\mathbf{P})^p = \begin{bmatrix} P_p(\varphi_z) + \phi_z U_{p-1}(\varphi_z) & -j\,\mathrm{sh}\nu_2\,\mathrm{sh}rs\,U_{p-1}(\varphi_z) \\ -j\,\mathrm{sh}\nu_2\,\mathrm{sh}rs\,U_{p-1}(\varphi_z) & P_p(\varphi_z) - \phi_z U_{p-1}(\varphi_z) \end{bmatrix}, \tag{20}$$

$$\varphi_z = -j\,\mathrm{ch}\nu_2\,\mathrm{sh}rs = \mathrm{ch}\nu_2 \sin r\frac{s}{j}, \quad \phi_z = -j\,\mathrm{ch}rs = -j\cos r\frac{s}{j}. \tag{21}$$

Zatem elementy P_{ik}^z macierzy $\mathbf{P}_z$ według (19) wynoszą

$$P_{11}^z = \mathrm{ch}\nu_2\,U_p(\varphi_z) - je^{-rs}U_{p-1}(\varphi_z), \quad P_{12}^z = P_{21}^z = \mathrm{sh}\nu_2\,U_p(\varphi_z),$$
$$P_{22}^z = \mathrm{ch}\nu_2\,U_p(\varphi_z) + je^{rs}U_{p-1}(\varphi_z). \tag{22}$$

Uwzględniając z kolei uzyskane rezultaty w wyrażeniu (6), otrzymuje się następujące zależności, określające elementy macierzy transmisyjnej $\mathbf{T}_0$ ogniwa podstawowego linii filtrującej o strukturze regularnej I rodzaju ze sprzęgaczami złożonymi:

$$T_{11}^0 = j^p(e^s\mathrm{ch}^2\nu_1 P_{11}^z + e^{-s}\mathrm{sh}^2\nu_1 P_{22}^z + \mathrm{sh}2\nu_1 P_{21}^z) = j^p\{U_p(\varphi_z)[(\mathrm{ch}2\nu_1\,\mathrm{ch}s + \mathrm{sh}s)\mathrm{ch}\nu_2 +$$
$$+ \mathrm{sh}2\nu_1\,\mathrm{sh}\nu_2] - jU_{p-1}(\varphi_z)[\mathrm{ch}2\nu_1\,\mathrm{sh}(1-r)s + \mathrm{ch}(1-r)s]\}, \quad T_{22}^0(s) = T_{11}^0(-s). \tag{23}$$

$$T_{21}^{0} = j^{p}\left[\frac{1}{2}\operatorname{sh}2v_1(P_{11}^{z}e^{s}+P_{22}^{z}e^{-s})+\operatorname{ch}2v_1 P_{21}^{z}\right] = j^{p}[U_p(\varphi_z)(\operatorname{sh}2v_1\operatorname{ch}v_2\operatorname{ch}s+$$

$$+\operatorname{ch}2v_1\operatorname{sh}v_2)-j\operatorname{sh}2v_1\operatorname{sh}(1-r)sU_{p-1}(\varphi_z)], \quad T_{12}^{0}(s) = -T_{21}^{0}(s). \tag{24}$$

Na tej podstawie poszukiwane elementy macierzy transmisyjnej $\mathbf{T}_n$ badanego układu można wyznaczyć z następujących związków:

$$(T_{11})_n = P_n(\varphi)+\phi U_{n-1}(\varphi), \quad (T_{21})_n = T_{21}^{0}\,U_{n-1}(\varphi),$$
$$\varphi = j^{p}[U_p(\varphi_z)\operatorname{ch}v_2\operatorname{sh}s-jU_{p-1}(\varphi_z)\operatorname{ch}(1-r)s], \tag{25}$$
$$\phi = j^{p}[U_p(\varphi_z)(\operatorname{ch}2v_1\operatorname{ch}v_2\operatorname{ch}s+\operatorname{sh}2v_1\operatorname{sh}v_2)-jU_{p-1}(\varphi_z)\operatorname{ch}2v_1\operatorname{sh}(1-r)s].$$

Wyprowadzone związki (25) są ścisłe i ważne dla dowolnej liczby nieparzystej p odcinków sprzęgających w sprzęgaczu złożonym oraz dla dowolnej liczby n ogniw podstawowych w linii filtrującej o strukturze regularnej I rodzaju ze sprzęgaczami złożonymi przy obciążeniu rezystywnym. Na podstawie tych zależności realizuje się syntezę filtru elektromechanicznego o badanej strukturze oraz wyznacza się parametry charakteryzujące zmienną pomocniczą dla potrzeb algebraicznej metody syntezy.

3.4. Unormowana macierz transmisyjna filtru elektromechanicznego o strukturze jednorodnej regularnej I rodzaju ze sprzęgaczami kombinowanymi

Przechodząc do analizy przypadku linii filtrującej o strukturze regularnej I rodzaju ze sprzęgaczami kombinowanymi rozpatruje się w pierwszym rzędzie macierz $\mathbf{P}_2(s)$ o postaci (7). Macierz ta obejmuje u macierzy $\mathbf{J}(qs)$ oraz $u+1$ macierzy $\mathbf{J}(rs)$; liczba macierzy typu $\mathbf{M}$ wynosi $2(u+1)$. W iloczynie (7) można wyodrębnić czynnik podstawowy

$$\mathbf{P} = \mathbf{J}\left(\frac{q}{2}s\right)\mathbf{M}_2\mathbf{J}(rs)\mathbf{M}_2^{-1}\mathbf{J}\left(\frac{q}{2}s\right) = \begin{bmatrix} e^{qs}(\operatorname{ch}rs+\operatorname{ch}2v_2\operatorname{sh}rs) & -\operatorname{sh}2v_2\operatorname{sh}rs \\ \operatorname{sh}2v_2\operatorname{sh}rs & e^{-qs}(\operatorname{ch}rs-\operatorname{ch}2v_2\operatorname{sh}rs)\end{bmatrix} \tag{26}$$

i tym samym wyrazić macierz $\mathbf{P}_2(s)$ sprzęgacza kombinowanego za pomocą zależności

$$\mathbf{P}_2(s) = \mathbf{J}\left(-\frac{q}{2}s\right)\mathbf{P}^{u+1}\mathbf{J}\left(-\frac{q}{2}s\right) = \mathbf{J}\left(-\frac{q}{2}s\right)\mathbf{P}_k\mathbf{J}\left(-\frac{q}{2}s\right). \tag{27}$$

Elementy macierzy $\mathbf{P}_k$ wyznacza się na podstawie wyrażenia (3), skąd otrzymuje się

$$\mathbf{P}_k = \begin{bmatrix} P_{u+1}(\varphi_k)+\phi_k U_u(\varphi_k) & -\operatorname{sh}2v_2\operatorname{sh}rs\,U_u(\varphi_k) \\ \operatorname{sh}2v_2\operatorname{sh}rsU_u(\varphi_k) & P_{u+1}(\varphi_k)-\phi_k U_u(\varphi_k)\end{bmatrix}, \tag{28}$$

gdzie:

$$\varphi_k = \operatorname{ch}rs\operatorname{ch}qs+\operatorname{ch}2v_2\operatorname{sh}rs\operatorname{sh}qs,$$
$$\phi_k = \operatorname{ch}rs\operatorname{sh}qs+\operatorname{ch}2v_2\operatorname{sh}rs\operatorname{ch}qs. \tag{29}$$

Zatem macierz $\mathbf{T}_0$ ma postać

$$\mathbf{T}_0 = \mathbf{M}_1\mathbf{J}\left[(1-q)\frac{s}{2}\right]\mathbf{P}_k\mathbf{J}\left[(1-q)\frac{s}{2}\right]\mathbf{M}_1^{-1}, \tag{30}$$

której uwzględnienie we wzorze (6) przynosi w rezultacie poszukiwane elementy macierzy $\mathbf{T}_n$

$$(T_{11})_n = P_n(\varphi)+\phi U_{n-1}(\varphi), \quad (T_{21})_n = T_{21}^0 U_{n-1}(\varphi),$$

$$T_{21}^0 = \operatorname{ch}2\nu_1 P_{21}^k + \frac{1}{2}\operatorname{sh}2\nu_1 (P_{11}^k e^{(1-q)s} - P_{22}^k e^{-(1-q)s}) = \tag{31}$$

$$= \operatorname{sh}2\nu_1 [\operatorname{sh}(1-q)sP_{u+1}(\varphi_k)+\operatorname{ch}(1-q)s\phi_k U_u(\varphi_k)]+\operatorname{ch}2\nu_1 \operatorname{sh}2\nu_2 \operatorname{sh}rsU_u(\varphi_k),$$

przy czym:

$$\varphi = \frac{1}{2}(P_{11}^k e^{(1-q)s}+P_{22}^k e^{-(1-q)s}) = \operatorname{ch}(1-q)sP_{u+1}(\varphi_k)+\operatorname{sh}(1-q)s\phi_k U_u(\varphi_k),$$

$$\phi = \frac{1}{2}\operatorname{ch}2\nu_1 (P_{11}^k e^{(1-q)s} - P_{22}^k e^{-(1-q)s})+\operatorname{sh}2\nu_1 P_{21}^k = \tag{32}$$

$$= \operatorname{ch}2\nu_1 [\operatorname{sh}(1-q)sP_{u+1}(\varphi_k)+\operatorname{ch}(1-q)s\phi_k U_u(\varphi_k)]+\operatorname{sh}2\nu_1 \operatorname{sh}2\nu_2 \operatorname{sh}rsU_u(\varphi_k).$$

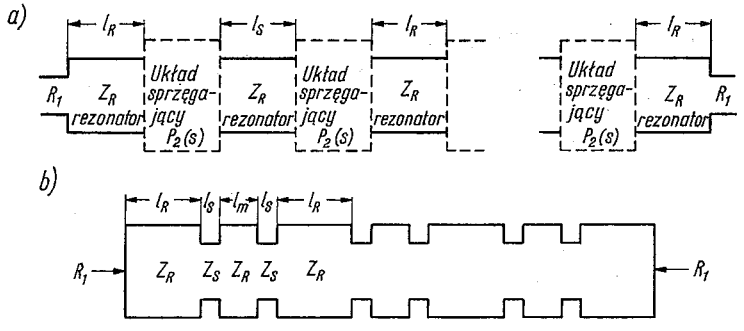
Zestaw wzorów (31)÷(32) daje informację o właściwościach transmisyjnych linii filtrującej o strukturze regularnej I rodzaju ze sprzęgaczami kombinowanymi w zależności od liczby ogniw podstawowych n, od liczby u określającej stopień komplikacji sprzęgacza, od parametru ν_1 uzależnionego od skoku impedancji na wejściu i wyjściu linii oraz od parametru ν_2 charakteryzującego skoki impedancji falowej w linii. Wyprowadzone zależności są ścisłe i nadają się do realizacji syntezy filtru elektromechanicznego o strukturze jednorodnej regularnej I rodzaju ze sprzęgaczami kombinowanymi.

4. UNORMOWANA MACIERZ TRANSMISYJNA FILTRU ELEKTROMECHANICZNEGO O STRUKTURZE JEDNORODNEJ REGULARNEJ II RODZAJU

Linia filtrująca o strukturze regularnej II rodzaju składa się z n jednakowych rezonatorów współpracujących z $n-1$ jednakowymi sprzęgaczami (rys. 4). Stąd unormowana macierz transmisyjna $\mathbf{T}_n$ tego czwórnika ma postać

$$\mathbf{T}_n(s) = \mathbf{M}_1 \mathbf{J}(s)\mathbf{P}_2(s)\mathbf{J}(s)\mathbf{P}_2(s) \ldots \mathbf{P}_2(s)\mathbf{J}(s)\mathbf{M}_1^{-1} =$$

$$= \mathbf{M}_1 \mathbf{J}\left(\frac{s}{2}\right)\mathbf{T}_0^{n-1}\mathbf{J}\left(\frac{s}{2}\right)\mathbf{M}_1^{-1}, \quad \mathbf{T}_0 = \mathbf{J}\left(\frac{s}{2}\right)\mathbf{P}_2(s)\mathbf{J}\left(\frac{s}{2}\right). \tag{33}$$



Rys. 4. Filtrująca linia mechaniczna o strukturze regularnej II rodzaju
a) dowolny układ sprzęgający o macierzy transmisyjnej $P_2(s)$, b) sprzęgacz kombinowany

Z porównania wzoru (33) z (6) wynika następująca równość:

$$(\mathbf{T}_n)_{II} = \mathbf{M}_1 \mathbf{J}\left(\frac{s}{2}\right)\mathbf{M}_1^{-1}(\mathbf{T}_{n-1})_I \mathbf{M}_1 \mathbf{J}\left(\frac{s}{2}\right)\mathbf{M}_1^{-1}, \quad (\mathbf{T}_0)_{II} = \mathbf{M}_1^{-1}(\mathbf{T}_0)_I \mathbf{M}_1, \tag{34}$$

gdzie indeks I wskazuje na strukturę regularną I rodzaju, a indeks II — na analogiczną strukturę II rodzaju.

Przytoczone zależności ukazują praktyczną drogę postępowania przy wyznaczaniu elementów macierzy transmisyjnej czwórnika o omawianej strukturze. Mianowicie celem określenia macierzy $(\mathbf{T}_n)_{II}$ wystarczy podstawić do wzoru (33) w miejsce macierzy $\mathbf{T}_0^{n-1}$ macierz $(\mathbf{T}_{n-1})_I$ o elementach według (15), (25) lub (31) w zależności od typu sprzęgacza z tym, że $(v_1)_I = 0$. Korzystając z tej wskazówki dochodzi się do następujących wzorów określających elementy macierzy $\mathbf{T}_n$ według (33):

$$\begin{aligned}(T_{11})_n &= \operatorname{ch} s P_{n-1}(\varphi) + \operatorname{sh} s\phi\, U_{n-2}(\varphi) + \operatorname{sh} 2v_1 T_{21}^0 U_{n-2}(\varphi) + \\ &\quad + \operatorname{ch} 2v_1 [\operatorname{sh} s P_{n-1}(\varphi) + \operatorname{ch} s\phi\, U_{n-2}(\varphi)], \\ (T_{21})_n &= \operatorname{sh} 2v_1 [\operatorname{sh} s P_{n-1}(\varphi) + \operatorname{ch} s\varphi U_{n-2}(\varphi)] + \operatorname{ch} 2v_1 T_{21}^0 U_{n-2}(\varphi),\end{aligned} \tag{35}$$

gdzie n jest liczbą rezonatorów w filtrze.

Funkcje φ, ϕ oraz T_{21}^0 zależą od typu sprzęgacza i zgodnie z danymi rozdz. 3 wynoszą: filtr elektromechaniczny ze sprzęgaczami prostymi

$$\begin{aligned}\varphi &= \operatorname{ch} 2v_2 \operatorname{sh} s \operatorname{sh} rs + \operatorname{ch} s \operatorname{ch} rs, \\ \phi &= \operatorname{ch} 2v_2 \operatorname{ch} s \operatorname{sh} rs + \operatorname{sh} s \operatorname{ch} rs, \\ T_{21}^0 &= \operatorname{sh} 2v_2 \operatorname{sh} rs;\end{aligned} \tag{36}$$

filtr elektromechaniczny ze sprzęgaczami złożonymi

$$\begin{aligned}\varphi &= j^p [\operatorname{ch} v_2 \operatorname{sh} s U_p(\varphi_z) - j \operatorname{ch}(1-r) s U_{p-1}(\varphi_z)], \\ \phi &= j^p [\operatorname{ch} v_2 \operatorname{ch} s U_p(\varphi_z) - j \operatorname{sh}(1-r) s U_{p-1}(\varphi_z)]; \\ T_{21}^0 &= j^p \operatorname{sh} v_2\, U_p(\varphi_z),\end{aligned} \tag{37}$$

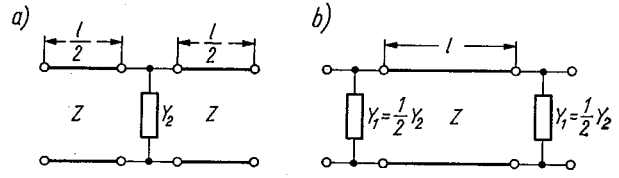
filtr elektromechaniczny ze sprzęgaczami kombinowanymi

$$\begin{aligned}\varphi &= \operatorname{ch}(1-q) s P_{u+1}(\varphi_k) + \operatorname{sh}(1-q) s \phi_k U_u(\varphi_k), \\ \phi &= \operatorname{sh}(1-q) s P_{u+1}(\varphi_k) + \operatorname{ch}(1-q) s \phi_k U_u(\varphi_k), \\ T_{21}^0 &= \operatorname{sh} 2v_2 \operatorname{sh} rs\, U_u(\varphi_k).\end{aligned} \tag{38}$$

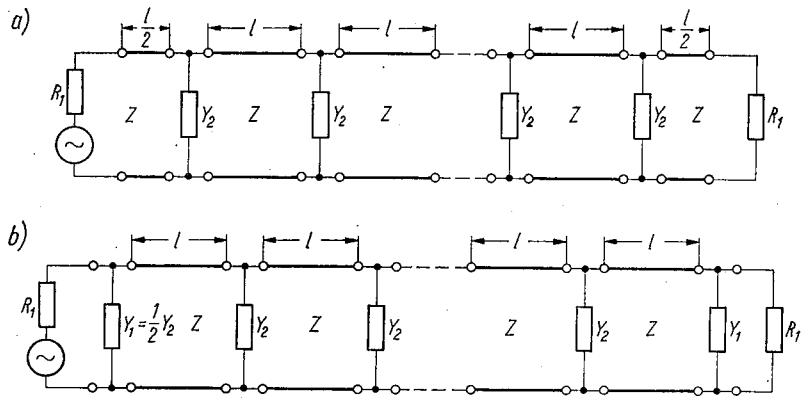
Przytoczone zależności są ścisłe i ważne dla dowolnych wartości parametrów decydujących o właściwościach filtru. Na ich podstawie przeprowadza się syntezę linii filtrującej o strukturze regularnej II rodzaju.

5. UNORMOWANA MACIERZ TRANSMISYJNA FILTRU O STRUKTURZE MIESZANEJ REGULARNEJ

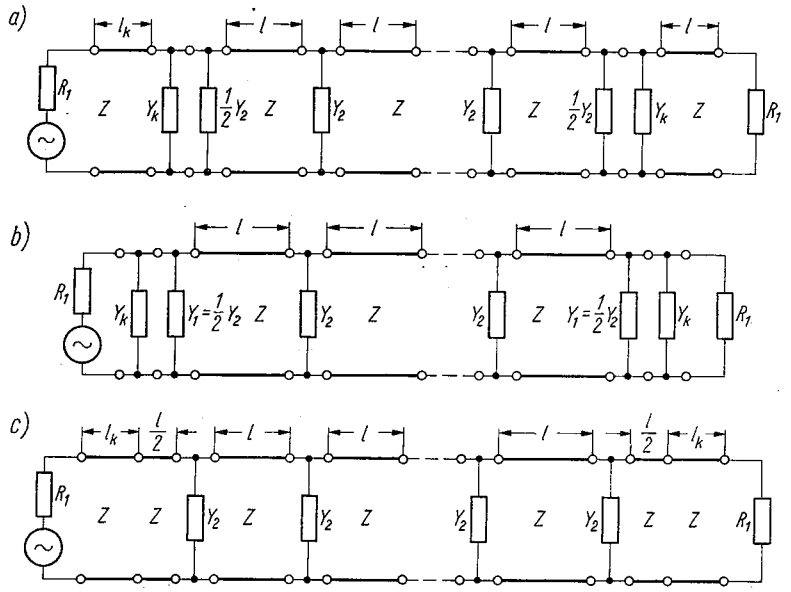
Filtry o strukturze mieszanej regularnej dzielą się na dwie grupy. Do jednej z nich należą filtry o strukturze mieszanej regularnej I rodzaju, a do drugiej — II rodzaju. Z kolei w każdej grupie wyodrębnia się podgrupy, przy czym za kryterium podziału służy typ



Rys. 5. Typy ogniw podstawowych w filtrze o strukturze mieszanej regularnej
a) ogniwo podstawowe typu *A*, b) ogniwo podstawowe typu *B*



Rys. 6. Schemat czwórnika o strukturze mieszanej regularnej I rodzaju
a) struktura regularna I*A*, b) struktura regularna I*B*



Rys. 7. Filtry o strukturze mieszanej regularnej II rodzaju
a) przypadek ogólny, ogniwo podstawowe typu *B*, b) przypadek szczególny, struktura II*B*, c) przypadek szczególny, struktura II*A*

ogniwa podstawowego, tworzącego strukturę regularną. Rozróżnia się ogniwa podstawowe *A* i *B* o schemacie jak na rys. 5 i odpowiednio struktury regularne I*A*, II*A*, I*B*, II*B*.

Dodanie do symbolu struktury regularnej litery *A* bądź *B* wskazuje na typ ogniwa podstawowego, z którego utworzono badany czwórnik łańcuchowy. Na rys. 6 podano schematy czwórnika o strukturze mieszanej regularnej I*A* i I*B*, a na rys. 7 — schematy filtru o strukturze mieszanej regularnej II rodzaju. W tym miejscu należy zaznaczyć, iż z reguły przez określenie *filtr o strukturze regularnej* II*A* będzie się rozumiało układ jak na rys. 7c, zaś przez określenie *filtr o strukturze regularnej* II*B* — układ jak na rys. 7b.

5.1. Unormowana macierz transmisyjna filtru o strukturze mieszanej regularnej I rodzaju

Zgodnie z rys. 6 macierz transmisyjna filtru o strukturze mieszanej regularnej I rodzaju daje się wyrazić za pomocą następującego wzoru:

$$\mathbf{T}_n = \mathbf{T}_0^n, \tag{39}$$

gdzie:

$\mathbf{T}_0$ — macierz transmisyjna ogniwa podstawowego przy obiążeniu R_1,

n — liczba ogniw podstawowych w łańcuchu.

Macierz $\mathbf{T}_0$ najlepiej wyznaczyć za pomocą wzoru (5), który wiąże elementy macierzy transmisyjnej czwórnika z elementami jego macierzy łańcuchowej. Macierz $\mathbf{a}_0$ ogniw podstawowych jak na rys. 5 wynosi

$$(\mathbf{a}_0)_A = \begin{bmatrix} \operatorname{ch} s + \frac{1}{2} Y_2' \operatorname{sh} s & Z\left[\operatorname{sh} s + \frac{1}{2} Y_2'(\operatorname{ch} s - 1)\right] \\ \frac{1}{Z}\left[\operatorname{sh} s + \frac{1}{2} Y_2'(\operatorname{ch} s + 1)\right] & \operatorname{ch} s + \frac{1}{2} Y_2' \operatorname{sh} s \end{bmatrix}, \tag{40}$$

$$(\mathbf{a}_0)_B = \begin{bmatrix} \operatorname{ch} s + Y_1' \operatorname{sh} s & Z \operatorname{sh} s \\ \frac{1}{Z \operatorname{sh} s}\left[(\operatorname{ch} s + Y_1' \operatorname{sh} s)^2 - 1\right] & \operatorname{ch} s + Y_1' \operatorname{sh} s \end{bmatrix}, \tag{41}$$

gdzie:

$$Y_i' = \frac{Z}{\omega_0 L_i} \cdot \frac{\Theta_0}{s}\left[1 - \left(\frac{s}{j\Theta_0} \cdot \frac{f_0}{f_i}\right)^2\right], \quad i = 1, 2, \quad Y_i' = ZY_i, \tag{42}$$

$$s = \sigma + j\Theta, \quad \Theta = \Theta_0 \frac{f}{f_0}, \quad f_0 = \frac{1}{2}(f_p + f_{-p}),$$
$$\left(\frac{1}{4} + k\right)\pi \leqslant \Theta_0 \leqslant \left(\frac{3}{4} + k\right)\pi, \quad k = 0, 1, 2, \ldots \tag{43}$$

Sens poszczególnych oznaczeń:

Z — impedancja falowa jednorodnego odcinka sprzęgającego,

Θ — przesuwność falowa (długość elektryczna) jednorodnego odcinka sprzęgającego,

Θ_0 — przesuwność falowa odcinka sprzęgającego dla częstotliwości f_0,
Y_i — admitancja dwójnika o parametrach L_i, C_i,
f_i — częstotliwość rezonansowa dwójnika Y_i,
f_0, f_{-p}, f_p — częstotliwości środkowa, dolna i górna pasma przepustowego użytkowego.

Podstawienie wzoru (40) do (5) daje w wyniku unormowaną macierz transmisyjną $(\mathbf{T}_0)_A$ ogniwa podstawowego A przy obciążeniu R_1; jej elementy są następujące:

$$(T_{11}^0)_A = \operatorname{ch} s + \frac{1}{2} Y_2' \operatorname{sh} s - \frac{1}{2} Y_2' \operatorname{sh} 2\nu + \operatorname{ch} 2\nu \left(\operatorname{sh} s + \frac{1}{2} Y_2' \operatorname{ch} s \right),$$

$$(T_{21}^0)_A = \operatorname{sh} 2\nu \left(\operatorname{sh} s + \frac{1}{2} Y_2' \operatorname{ch} s \right) - \frac{1}{2} Y_2' \operatorname{ch} 2\nu, \quad \nu = \frac{1}{2} \ln \frac{Z}{R_1}, \tag{44}$$

$$(T_{11}^0)_B = (\operatorname{ch} s + Y_1' \operatorname{sh} s) \left[1 + \frac{R_1}{2Z \operatorname{sh} s} (\operatorname{ch} s + Y_1' \operatorname{sh} s) \right] + \frac{1}{2} \left(\frac{Z \operatorname{sh} s}{R_1} - \frac{R_1}{Z \operatorname{sh} s} \right),$$

$$(T_{21}^0)_B = \frac{1}{2} \left(\frac{Z \operatorname{sh} s}{R_1} + \frac{R_1}{Z \operatorname{sh} s} \right) - \frac{R_1}{2Z \operatorname{sh} s} (\operatorname{ch} s + Y_1' \operatorname{sh} s)^2. \tag{45}$$

Pozostałe elementy macierzy $\mathbf{T}_0$ wyznacza się na podstawie znanych wzajemnych związków między jej elementami.

Uzyskane rezultaty stanowią punkt wyjścia do obliczenia elementów macierzy $\mathbf{T}_n$, wzór (39). W tym celu wystarczy skorzystać ze związku (3), skąd otrzymuje się

$$(T_{11})_n = P_n(\varphi) + \phi U_{n-1}(\varphi), \quad (T_{21})_n = T_{21}^0 U_{n-1}(\varphi), \tag{46}$$

gdzie:

$$\varphi = \begin{cases} \operatorname{ch} s + \frac{1}{2} Y_2' \operatorname{sh} s, & \text{ogniwo A}, \\ \operatorname{ch} s + Y_1' \operatorname{sh} s, & \text{ogniwo B}, \end{cases} \tag{47}$$

$$\phi = \begin{cases} -\frac{1}{2} Y_2' \operatorname{sh} 2\nu + \operatorname{ch} 2\nu \left(\operatorname{sh} s + \frac{1}{2} Y_2' \operatorname{ch} s \right), & \text{ogniwo A}, \\ \frac{R_1}{2Z \operatorname{sh} s} (\operatorname{ch} s + Y_1' \operatorname{sh} s)^2 + \frac{1}{2} \left(\frac{Z \operatorname{sh} s}{R_1} - \frac{R_1}{Z \operatorname{sh} s} \right), & \text{ogniwo B}. \end{cases} \tag{48}$$

Wyprowadzone zależności (46) ÷ (48) są ścisłe w ramach przyjętych założeń i nadają się do przeprowadzenia syntezy filtrów o strukturze mieszanej regularnej I rodzaju w oparciu o parametry skuteczne czwórnika.

5.2. Unormowana macierz transmisyjna filtru o strukturze mieszanej regularnej II rodzaju

Wyniki uzyskane w rozdz. 5.1 są przydatne także przy poszukiwaniu macierzy transmisyjnej $\mathbf{T}_n$ filtru o strukturze mieszanej regularnej II rodzaju, gdyż macierz $\mathbf{T}_n$ czwórnika jak na rys. 7 daje się napisać następująco:

$$\mathbf{T}_n = \mathbf{T}_k \mathbf{T}_0^n \mathbf{T}_k', \tag{49}$$

gdzie:

$\mathbf{T}_k$ — macierz transmisyjna czwórnika dołączonego do ogniwa podstawowego na wejściu filtru,

$\mathbf{T}'_k$ — macierz transmisyjna czwórnika dołączonego do ogniwa podstawowego na wyjściu filtru,

$\mathbf{T}_0$ — macierz transmisyjna ogniwa podstawowego,

n — liczba ogniw podstawowych w łańcuchu.

W tym miejscu należy zaznaczyć, że elementy czwórników o macierzy $\mathbf{T}_k$ oraz $\mathbf{T}'_k$ mogą być ujemne pod warunkiem, iż układ wypadkowy będzie miał elementy fizycznie realizowalne.

Elementy macierzy $\mathbf{T}_0^n$ wyznacza się na podstawie zależności (46)÷(48) z uwzględnieniem (44), (45). W ten sposób zadanie sprowadza się do ustalenia macierzy $\mathbf{T}_k$, $\mathbf{T}'_k$ i wykonania odpowiednich działań zgodnie ze wzorem (49).

Dla przypadku ogólnego, przedstawionego na rys. 7a, elementy macierzy $\mathbf{T}_k$ i $\mathbf{T}'_k$ wynoszą

$$
\begin{aligned}
[T_{11}(s)]_k &= \operatorname{ch}\frac{s_k}{2} + \frac{1}{2} Y'_k \operatorname{sh}\frac{s_k}{2} + \operatorname{ch} 2\nu \operatorname{sh}\frac{s_k}{2} + \frac{1}{2} Y'_k \mathrm{e}^{-2\nu} \operatorname{ch}\frac{s_k}{2}, \\
[T_{21}(s)]_k &= \frac{1}{2} Y'_k \operatorname{sh}\frac{s_k}{2} + \operatorname{sh} 2\nu \operatorname{sh}\frac{s_k}{2} - \frac{1}{2} Y'_k e^{-2\nu} \operatorname{ch}\frac{s_k}{2}, \\
s_k &= \left(\frac{2l_k}{l}\right) s
\end{aligned}
\tag{50}
$$

oraz

$$
\begin{aligned}
[T_{11}(-s)]_k &= [T_{22}(s)]_k = [T'_{11}(-s)]_k = [T'_{22}(s)]_k, \\
[T_{21}(-s)]_k &= [T_{12}(s)]_k = -[T'_{21}(s)]_k = -[T'_{12}(-s)]_k.
\end{aligned}
\tag{51}
$$

Postać funkcji Y'_k jest taka sama jak we wzorze (42); parametr ν jest zdefiniowany przez zależność (44); l jest długością jednorodnego odcinka sprzęgającego między dwoma sąsiednimi dwójnikami Y_2; l_k jest długością skrajnych odcinków jednorodnych w czwórniku jak na rys. 7a.

Po uwzględnieniu związków (51) we wzorze (49) i po wykonaniu mnożenia otrzymuje się

$$
\begin{aligned}
[(T_{11})_n]_{II} &= [(T_{11})_n]_{IB}(T_{11}^2)_k + 2[(T_{21})_n]_{IB}(T_{11}T_{12})_k - [(T_{22})_n]_{IB}(T_{12}^2)_k, \\
[(T_{21})_n]_{II} &= [(T_{11})_n]_{IB}(T_{11}T_{21})_k + [(T_{21})_n]_{IB}[1 + 2(T_{12}T_{21})_k] - [(T_{22})_n]_{IB}(T_{12}T_{22})_k.
\end{aligned}
\tag{52}
$$

Wzór (52) określa elementy macierzy transmisyjnej filtru o strukturze mieszanej regularnej II rodzaju w funkcji elementów macierzy transmisyjnej filtru o strukturze mieszanej regularnej IB oraz elementów macierzy transmisyjnej skrajnych czwórników współpracujących z ogniwami podstawowymi typu B.

Jeżeli z kolei podstawić do wzoru (52) zależności (45)÷(48), (50,) to w wyniku uzyska się poszukiwane wyrażenie na elementy macierzy $\mathbf{T}_n$ w funkcji elementów układu o strukturze mieszanej regularnej II rodzaju. W dalszych rozważaniach ograniczono się do przed-

stawienia rezultatu takiego podstawienia w odniesieniu do funkcji $(T_{21})_n$ dla różnych odmian struktury regularnej II rodzaju.

Jak można sprawdzić, uwzględnienie związków (45)÷(48), (50) we wzorze (52) prowadzi do następującej postaci funkcji filtracji układu łańcuchowego jak na rys. 7a (przypadek ogólny analizowanego filtru o strukturze mieszanej regularnej II rodzaju):

$$(T_{21})_n = P_n(\varphi_A)[T^0_{21}(s_k, 2Y'_k)]_A + \frac{1}{2}U_{n-1}(\varphi_A)\Big\{\varphi_B(Y'_k)[T^0_{21}(s_k, 2Y'_k)]_A + \varphi_A[T^0_{21}(s_k)]_A +$$

$$+2\operatorname{sh}2\nu\operatorname{sh}(s-s_k) + \left(\frac{1}{2}Y'_2 - Y'_k\right)[\operatorname{sh}2\nu\operatorname{ch}(s-s_k) - \operatorname{ch}2\nu\operatorname{ch}s]\Big\}. \tag{53}$$

W tym wzorze funkcję φ bierze się zgodnie ze wzorem (47) z uwzględnieniem typu ogniwa podstawowego zaznaczonego za pomocą wskaźnika A lub B. Funkcja $(T^0_{21})_A$ jest określona przez wzór (44). Wymienienie argumentów obok symbolu funkcji oznacza, że należy zastąpić nimi odpowiednie parametry we wzorze określającym funkcję.

Z kolei korzystając z zależności (53) można podać postać funkcji filtracji czwórnika o strukturze mieszanej regularnej IIA (rys. 7c) i IIB (rys. 7b). Dla układu jak na rys. 7c obowiązuje równość $Y'_k = \frac{1}{2}Y'_2$, przeto

$$[(T_{21})_n]_{IIA} = [T^0_{21}(s_k)]_A U_n(\varphi_A) + \operatorname{sh}2\nu\operatorname{sh}(s-s_k)\,U_{n-1}(\varphi_A). \tag{54}$$

Natomiast w przypadku filtru o strukturze mieszanej regularnej IIB zachodzi $s_k \equiv 0$ i na tej podstawie jego funkcja filtracji wynosi

$$[(T_{21})_n]_{IIB} = -\frac{R_1}{Z}\left\{U_{n-1}(\varphi_B)\left[\frac{1}{2}(Y'^2_1 + Y'^2_k)\operatorname{sh}s + Y'_1\operatorname{ch}s - \frac{1}{2}\left(\frac{Z^2}{R_1^2} - 1\right)\operatorname{sh}s\right] + P_n(\varphi_B)\,Y'_k\right\}. \tag{55}$$

Należy zaznaczyć, że we wszystkich wzorach, dotyczących właściwości transmisyjnych filtru o strukturze mieszanej regularnej II rodzaju, parametr n określa liczbę ogniw podstawowych typu B, które dają się wyodrębnić w łańcuchu; zatem n jest o jedność mniejsze od liczby dwójników typu Y w układzie.

Wyprowadzone zależności służą do syntezy filtru o badanej strukturze, a także do wyznaczenia parametrów zmiennej pomocniczej x przy posługiwaniu się algebraiczną metodą syntezy [6]. Są one także pomocne przy szacowaniu stosunku $\frac{Z}{R_1}$ celem dokładnego spełnienia wymagań dotyczących tłumienności skutecznej dla jednej konkretnej częstotliwości.

6. UWAGI O SYNTEZIE UKŁADÓW FILTRUJĄCYCH O STRUKTURZE REGULARNEJ

Właściwości transmisyjne układów filtrujących o strukturze regularnej zależą jedynie od stosunkowo niewielkiej liczby zmiennych parametrów, które muszą być dobrane w procesie syntezy. W przypadku linii filtrujących są to: liczba ogniw podstawowych

lub rezonatorów n, współczynnik skoku impedancji falowej na wejściu linii i w sprzęgaczach $\left(\text{tzn. } m_1 = \frac{R_1}{Z_R},\ m_2 = \frac{Z_R}{Z_S}\right)$, długość odcinka linii długiej przyjętego za odcinek odniesienia, a także ew. liczba p lub u określająca stopień komplikacji sprzęgaczy złożonych lub kombinowanych.

Jeżeli synteza dotyczy układu o strukturze mieszanej regularnej, to należy wyznaczyć liczbę ogniw podstawowych n, elementy dwójników Y_2 oraz Y_k, długość i impedancję falową odcinków linii długiej między sąsiednimi dwójnikami oraz na wejściu i wyjściu układu. Oczywiście całość obliczeń zależy także od swobody wyboru rodzaju struktury regularnej.

Zależności wyprowadzone w ramach niniejszej pracy określają wszystkie elementy macierzy transmisyjnej filtru. Dzięki temu istnieje możliwość ścisłego obliczenia charakterystyki tłumieniowej i fazowej czwórnika, gdyż jak wiadomo [1, 2, 10] dla $s = j\Theta$

$$\begin{aligned} \ln T_{11} &= \ln|T_{11}| + \text{jarg}\, T_{11} = A_s + jB_s, \\ A_s &= \frac{1}{2}\ln(1+|T_{21}|^2), \end{aligned} \tag{56}$$

gdzie:

A_s — tłumienność skuteczna,

B_s — przesuwność skuteczna czwórnika.

Synteza filtru o strukturze regularnej na podstawie wymagań odnośnie do charakterystyki tłumieniowej jest realizowana w oparciu o funkcję filtracji $(T_{21})_n$ z uwzględnieniem związku (56) w trakcie dwu etapów podstawowych. Zadanie etapu I polega na ustaleniu niezbędnej liczby ogniw podstawowych n. Uzyskuje się to za pomocą metody kolejnych przybliżeń. W tym celu dokonuje się cyklicznego powtarzania pewnego zestawu obliczeń, uwarunkowanego przez wyprowadzone uprzednio zależności, aż osiągnie się zadowalającą zgodność wyników z postawionymi wymaganiami. W trakcie etapu II wyznacza się pozostałe parametry układu.

Zależności opisujące właściwości transmisyjne filtru o strukturze regularnej są ścisłe, co pozwala przewidzieć wszelkie niuanse w przebiegu charakterystyki tłumieniowej, a w szczególności jej asymetrię w pasmie przepustowym, i sprawdzać zgodność wyników z postawionymi wymaganiami dla częstotliwości krytycznych. Po wykonaniu obliczeń istnieje możliwość porównania założonej charakterystyki filtru z obliczoną na podstawie podanych w tej pracy zależności.

Pewne pomocnicze wskazówki o sposobie realizacji syntezy filtru elektromechanicznego o strukturze regularnej wraz z konkretnym przykładem obliczeń są podane w pracy [3].

Reasumując należy podkreślić, że synteza układów filtrujących o strukturze regularnej korzysta ze związków ścisłych, umożliwiających kontrolę zarówno charakterystyki tłumieniowej, jak i fazowej czwórnika.

LITERATURA CYTOWANA W TEKŚCIE

1. A. L. Feldsztejn, L. R. Jawicz, Sintiez czetyriochpolusnikow i wosmipolusnikow na SWCz, Wyd. II, Izdat. „Swiaz", Moskwa 1971.
2. F. Kamiński, Sintiez szirokopołosnych elektromiechaniczeskich cepoczecznych filtrow s riegularnoj strukturoj, Tocznoje rieszenije, Bull. Acad. Polon. Sci., Sér. sci. techn., **15**, 2, 1(165)—5(169), 1967.
3. F. Kamiński, Uwagi o syntezie elektromechanicznych filtrów łańcuchowych o strukturze regularnej, Prace ITR, **11**, 4, 33—46, 1967.
4. F. Kamiński, O sintiezie swierchuzkopołosnych elektromiechaniczeskich cepoczecznych filtrow s riegularnoj strukturoj, Bull. Acad. Polon. Sci., Sér. sci. techn. **17**, 1, 1(101)—5(105), 1969.
5. F. Kamiński, Synteza łańcuchowych filtrów elektromechanicznych za pomocą układów zastępczych o stałych rozłożonych, Rozpr. Elektrotechn., **15**, 4, 717—749, 1969.
6. F Kamiński. O syntezie układów filtrujących o strukturze mieszanej II rodzaju, Arch. Elektrotechn., **22**, 3, 539—548, 1973.
7. F. Kamiński, Synteza pewnej grupy torów niejednorodnych metodą algebraiczną z uwzględnieniem układów zastępczych o elementach skupionych, Arch. Elektrotechn., **21**, 2 374—380, 1972.
8. F. Kamiński, Synteza filtru elektromechanicznego ze sprzęgaczami prostymi, Arch. Elektrotechn., **22**, 3, 525—537, 1973.
9. F. Kamiński, O syntezie pewnej grupy filtrów elektromechanicznych za pomocą układów zastępczych o stałych rozłożonych lub skupionych, Arch. Elektrotechn., **22**, 2, 517—521, 1973.
10. F. Kamiński, Cechy charakterystyczne torów niejednorodnych złożonych z odcinków jednorodnych, Przegl. Telekom., **38**, 8, 225—229, 1966.
11. C. Kurth, Eine Wellenparametertheorie für mechanische Vierpole in Kompressions- oder Torsionsschwingungen, Nachrichtentechnik, **9**, 11, 490—502, 1959.
12. A. K. Łosiew, Tieorija i rasczot elektromiechaniczeskich filtrow, Izdat. „Swiaz", Moskwa 1965.
13. A. N. Pietrow, W. F. Szmatczenko, Radiowe filtry elektromechaniczne, WNT, Warszawa 1964.
14. M. Skowrońska, Teoria filtru elektromechanicznego o drganiach skręcających lub podłużnych, Zeszyty Naukowe PW, 86 (Elektryka 33), 39—69, 1964.

F. KAMIŃSKI

OUTLINE OF THE THEORY OF THE REGULAR-STRUCTURE MECHANICAL AND MICROWAVE FILTER SYNTHESIS

Summary

In the paper is presented a method of the regular-structure two-port network analysis for designing mechanical and microwave bandpass filters. Basic relations and a definition of regular-structure fourpoles are contained in chapter 2. On the basis of these data is derived a formula describing the normalized wave transmission matrix of regular-homogeneous-structure mechanical filters with a diverse coupler realization (chapters 3, 4) as well as a formula for the same matrix of mechanical and microwave bandpass filters with regular mixed structure (chapter 5). The expressions derived are strict taking into account the idealization of physical phenomena in devices under consideration; they describe transmission properties of the filter considered within a operating frequency range satisfactorily.

On the basis of the paper results, effective network parameter synthesis of mechanical and microwave filters with regular structure is performed. It is possible to design them with due regard for prescribed insertion loss or phase characteristic conditions (see eq. (56)). Side by side with this the results obtained are of use for synthesizing mechanical and microwave filters with homogeneous or mixed irregular structure by means of an algebraic procedure.

F. KAMIŃSKI

LE PRÉCIS DE LA THÉORIE DE LA SYNTHÈSE DE FILTRES ÉLECTROMÉCANIQUES ET À HYPERFREQUENCE À STRUCTURE RÉGULIÈRE

Résumé

On présente dans l'article une méthode d'analyse de quadripôles à structure régulière afin de faire la synthèse de filtres électromécaniques et à hyperfréquence. Les rapports de base initiales et la définition d'une circuit à structure régulière se trouvent dans le chapitre 2. On les applique à obtenir une matrice de transmission normée de filtres électromécaniques à structure régulière homogène étant construits avec des coupleurs différents (le chapitre 3 et 4) et aussi une matrice analogique de filtres électromécaniques et à hyperfréquence à structure régulière mixte (le chapitre 5). Compte tenu de l'idéalisation de phénomènes physiques les expressions obtenues sont strictes. Elles décrivent bien des propriétés de transmission de circuits en question dans la bande de fonctionnement.

Au moyen des résultats obtenus on effectue la synthèse de filtres en question. On peut le faire en prenant pour base des conditions d'affaiblissement effectif ou de caractéristique phase-fréquence (f. (56)). En outre les résultats en question conviennent à la synthèse de filtres électromécaniques et à hyperfréquence à structure irrégulière homogène et mixte à l'aide d'un procédé algébrique.

F. KAMIŃSKI

ENTWURF EINER SYNTHESENTHEORIE MECHANISCHER UND MIKROWELLENFILTER MIT REGULÄRER STRUKTUR

Zusammenfassung

Es wird ein Verfahren zur Analyse des Übertragungsverhaltens strukturregulärer Vierpole dargestellt, die als mechanische und mikrowelle Filter angewandt werden. Grundlegende Beziehungen für diese Vierpole und ihre Definition sind im Kapitel 2 angeführt worden. Diese Daten dienen zur Ableitung einer Betriebsmatrix für mechanische Filter mit regulärer homogener Struktur und mit verschiedenartigen Köpplern (Kapitel 3, 4) und einer ähnlichen Matrix für mechanische und mikrowelle Filter (Kapitel 5) mit einer regulären gemischten Struktur. Die abgeleiteten Formeln treffen hinsichtlich angenommener Voraussetzungen genau zu und beschreiben das Vierpolverhalten im Betriebsfrequenzgebiet richtig.

Auf Grund theoretischer Resultate dieser Arbeit wird die Synthese erwähnter Filtergruppen realisiert, wobei die Berücksichtigung von Betriebsdämpfungs - oder Winkelmaßbedingungen (siehe Gl. 56) möglich ist. Darüber hinaus sind die Ergebnisse bei Anwendung von algebraischer Synthesenmethode zur Dimensionierung mechanischer und mikroweller Filter mit irregulären, homogenen oder gemischten Strukturen brauchbar.

Ф. КАМИНЬСКИ

ОЧЕРК ТЕОРИИ СИНТЕЗА ЭЛЕКТРОМЕХАНИЧЕСКИХ И МИКРОВОЛНОВЫХ ФИЛЬТРОВ С РЕГУЛЯРНОЙ СТРУКТУРОЙ

Резюме

Представлен метод синтеза трансмиссионных свойств четырехполюсников с регулярной структурой, используемых в качестве электромеханических и микроволновых фильтров. Основные исходные соотношения и определение понятия „регулярная структура" даны в главе 2. На основании этих данных выведено выражение для нормированной волновой матрицы передачи электромеханических фильтров с однородной регулярной структурой, изготовляемых с применением простых, сложных и комбинированных связок (глава 3, 4), а также выражение для аналогичной

матрицы электромеханических и микроволновых фильтров со смешанной регулярной структурой (глава 5). Полученные выражения являются точными в пределах принятой идеализации физических явлений в реальных устройствах и удовлетворительно описывают свойства четырехполюсника в рабочей полосе частот.

Результаты настоящей работы используются в синтезе данной группы фильтров, причем имеется возможность учета требований относительно амплитудной или фазовой характеристики фильтра (см. ф-лу (56)). Наряду с этим полученные результаты полезны при синтезе электромеханических и микроволновых фильтров с однородной и смешанной нерегулярной структурой и применением алгебраического метода.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 2, ss. 313—330

621.382

Pomiar właściwości szumowych warstw rezystywnych Pd/PdO/Ag za pomocą miernika typu Quan Tech

ANDRZEJ KUSY (RZESZÓW)

Wyższa Szkoła Inżynierska w Rzeszowie

Otrzymano 1.2.1973

Przedstawiono podstawowe wielkości charakteryzujące szumy, które występują w warstwach rezystywnych, ze szczególnym uwzględnieniem szumów typu $\frac{1}{f}$. Podano dokładną analizę metodyki pomiaru szumów prądowych miernikiem typu Quan Tech oraz sposobu wyznaczania wskaźnika k_{sz}, stanowiącego miarę *jakości szumowej* kompozycyjnych warstw rezystywnych. Definicja wskaźnika szumów prądowych k_{sz} standaryzuje pomiar, sprowadzając napięcie skuteczne szumów do określonego napięcia zasilania równego 1 V oraz dekadowego pasma częstotliwości.

Przeprowadzono dyskusję wpływu postaci funkcji gęstości widmowej mocy szumów elementów rezystywnych na wyniki pomiarów. Podano wyniki badań, dotyczące wpływu niektórych parametrów technologicznych, a mianowicie ułamka atomowego palladu w składnikach metalicznych pasty x_1 oraz czasów i temperatury szczytowej wypalania na wielkości charakteryzujące właściwości szumowe warstw rezystywnych Pd/PdO/Ag. Analiza uzyskanych wyników wykazała, że największe znaczenie z punktu widzenia poziomu szumów prądowych, a także i rezystancji warstw mają dwa czynniki: ułamek Pd x_1 oraz temperatura wypalania ϑ. Przedstawiono koncepcję wyjaśnienia uzyskanych zależności opierając się na dokonanej analizie fazowej warstw.

1. WSTĘP

Fluktuacje elektryczne, jakimi są szumy występujące w kompozycyjnych warstwach rezystywnych, np. typu Pd/PdO/Ag, należą do procesów przypadkowych, których gęstość widmowa mocy jest malejącą funkcją częstotliwości, najczęściej postaci $\frac{1}{f}$. Stąd też pochodzi nazwa tych wielkości przypadkowych określanych jako *szumy typu* $\frac{1}{f}$. Często stosuje się także nazewnictwo *szumy nadmiarowe* (excees noise) oraz *szumy prądowe* (current noise). Ostatnia nazwa pochodzi stąd, że szumy te występują podczas przepływu prądu przez rezystor i znikają przy jego braku, stanowiąc nadwyżkę lub nadmiar w stosunku do szumów termicznych występujących także wtedy, gdy rezystor nie jest spolaryzowany. Szumy termiczne mogą być określone teoretycznie na podstawie znanej reguły Nyquista, jeśli znane są: wartość rezystancji oraz pasmo częstotliwości. Wielkości charakteryzują-

cych szumy prądowe nie da się w taki sposób teoretycznie przewidzieć i dlatego muszą być one wyznaczane doświadczalnie w odpowiednich układach pomiarowych.

W niniejszej pracy przedstawiono metodę określania właściwości szumowych kompozycyjnych warstw rezystywnych wykorzystując miernik typu Quan Tech Model 315. Stosowana w pracy definicja wskaźnika szumów prądowych posiada tę zaletę, że ułatwia porównywanie i ocenę jakości rezystorów wytwarzanych różnymi technologiami przez różnych producentów, z punktu widzenia ich właściwości szumowych.

Przedmiot badań stanowiły rezystywne warstwy grube wytwarzane metodami sitodruku i wypalania odpowiednich kompozycji (past) na bazie palladu i srebra [4] [6] [7]. Warstwy te, jako elementy o strukturze ziarnistej, wykazują szumy prądowe, których wielkość jest uwarunkowana technologią wytwarzania. Przedstawiono wyniki pomiarów i obliczeń, dotyczące wpływu niektórych parametrów technologicznych i strukturalnych (skład wyjściowy, parametry wypalania, skład fazowy warstwy) na charakterystyki szumowe badanych warstw. Wyniki te stanowią fragment ogólnych badań w zakresie określania relacji pomiędzy procesem technologicznym — strukturą i parametrami elektrycznymi warstw, które nie znalazły dotąd należytego wyjaśnienia w literaturze.

2. WIELKOŚCI CHARAKTERYZUJĄCE SZUMY PRĄDOWE

Moc średnia zawarta w sygnale szumów jest ogólnie dana wzorem [1]:

$$P = \int_0^\infty G(f)\,df, \tag{1}$$

w którym $G(f)$ oznacza gęstość widmową mocy szumów.

Jeśli szumy określa się w paśmie częstotliwości od f_1 do f_2, to związana z nimi moc średnia dana jest zależnością:

$$P(f_1 \div f_2) = \int_{f_1}^{f_2} G(f)\,df. \tag{2}$$

Jak już wspomniano we wstępie, gęstość widmowa mocy szumów prądowych jest malejącą funkcją częstotliwości. Zależność tę można w ogólnym przypadku zapisać następująco [8] [10]:

$$G_p(f) = C' U_S^{\alpha} f^{-\gamma}, \tag{3}$$

gdzie:

G_p — gęstość widmowa mocy szumów prądowych,

C', α i γ — stałe zależne od rodzaju rezystora,

U_z — napięcie stałe przełożone do rezystora.

Zgodnie z danymi literaturowymi [2] [9] [10] można przyjąć, że dla rezystorów kompozycyjnych $\alpha = 2$ oraz $\gamma = 1$ i stosować wzór w postaci:

$$G_p = C' U_z^2 f^{-1}. \tag{4}$$

Wstawiając (4) do (2) otrzymujemy wyrażenie na moc średnią szumów prądowych dla pasma częstotliwości określonego częstotliwościami f_1 i f_2:

$$P_p(f_1 \div f_2) = \int_{f_1}^{f_2} C' U_z^2 f^{-1} df = C' U_z^2 \ln \frac{f_2}{f_1}. \tag{5}$$

Jak widać ze wzoru (5) moc średnia szumów prądowych jest proporcjonalna do kwadratu napięcia zasilania rezystora oraz do logarytmu naturalnego ze stosunku częstotliwości granicznych. W myśl definicji mocy średniej szumów, P_p jest mocą szumów wydzieloną na jednoomowej rezystancji. Ponieważ w dalszym ciągu niniejszej pracy wielkością używaną do charakteryzowania szumów będzie napięcie, więc zależności (4) i (5) zostaną teraz napisane w nieco innej postaci:

$$G_u(f) = C U_z^2 f^{-1}, \tag{4a}$$

$$\overline{u_p^2} = C U_z^2 \ln \frac{f_2}{f_1}, \tag{5a}$$

gdzie $G_u(f)$ jest gęstością widmową kwadratu napięcia, a $\overline{u_p^2}$ — średnim kwadratem napięcia szumów prądowych. Jeżeli chodzi o stałą materiałową C we wzorach (4a) i (5a), to jest ona wielkością bezwymiarową, podczas gdy stała C' (wzory (4) i (5)) posiada wymiar oma.

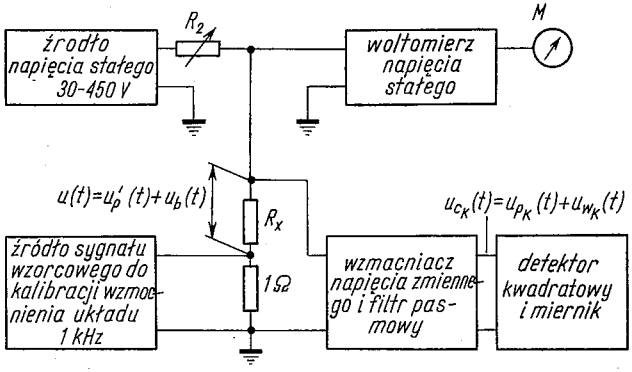
3. METODYKA POMIARU SZUMÓW PRĄDOWYCH

3.1. Schemat blokowy miernika

W zastosowanym w badaniach przyrządzie typu Quan Tech określenie wskaźnika szumów prądowych wymaga przeprowadzenia pomiaru wartości skutecznej napięcia szumów prądowych, danej wzorem:

$$U_p \underset{(\text{def})}{=} \sqrt{\overline{u_p^2}}, \tag{6}$$

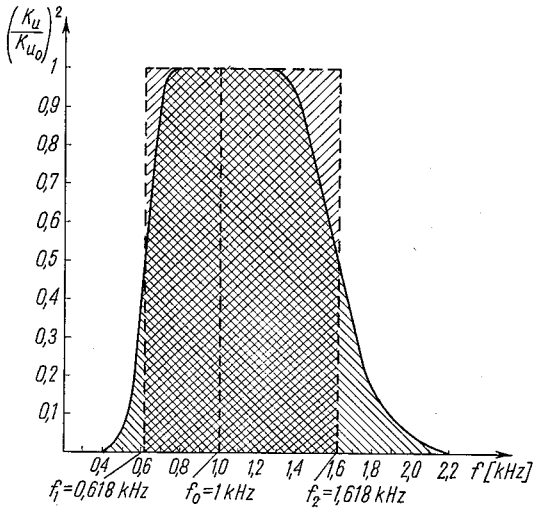
gdzie U_p jest wartością skuteczną napięcia szumów prądowych, zaś $\overline{u_p^2}$ jest dane wzorem (5a). Dokonuje się tego zgodnie ze schematem przedstawionym na rys. 1.



Rys. 1. Schemat blokowy przyrządu typu Quan Tech do pomiaru szumów prądowych rezystorów

$u(t)$, $u_p'(t)$ i $u_b(t)$ oznaczają napięcia szumów odpowiednio całkowitych, prądowych i cieplnych występujących na badanym rezystorze, natomiast $u_{c_K}(t)$, $u_{p_K}(t)$ i $u_{w_K}(t)$ — napięcia szumów odpowiednio całkowitych, prądowych oraz wzmacniacza i cieplnych rezystora po przejściu przez wzmacniacz i filtr pasmowy

Napięcie U_z polaryzujące badany rezystor o rezystancji R_x pochodzi ze źródła napięcia stałego regulowanego w zakresie od 30 do 400 V. Wartość napięcia U_z występującego na badanym rezystorze mierzy się za pomocą woltomierza lampowego. Rezystor izolacyjny o rezystancji R_z ma za zadanie zabezpieczyć przed znacznym tłumieniem sygnału szumów, występującego na badanym rezystorze, przez równoległą impedancję wyjściową źródła zasilania. Nie może on być źródłem szumów prądowych i dlatego jest odpowiednio wykonanym rezystorem drutowym. Wartość jego rezystancji jest regulowana i może wynosić 1 kΩ, 10 kΩ, 100 kΩ lub 1 MΩ w zależności od rezystancji R_x. Konkretna wartość rezystancji R_z nie jest krytyczna, jednakże w celu uniknięcia przeciążenia zarówno



Rys. 2. Charakterystyka częstotliwościowa filtru zastosowanego w przyrządzie do pomiaru szumów prądowych

rezystora badanego jak i izolacyjnego, wartość rezystancji tego ostatniego oraz napięcie zasilania muszą być odpowiednio dobrane.

Napięcie szumów występujące na badanym rezystorze jest wzmacniane we wzmacniaczu liniowym, którego poziom szumów własnych odniesiony do wejścia nie przekracza 10^{-2} $\mu V/\sqrt{Hz}$ na częstotliwości 1000 Hz, co odpowiada zastępczej rezystancji szumów cieplnych równej 6 kΩ oraz przechodzi przez filtr pasmowy o częstotliwości środkowej 1000 Hz i charakterystyce przenoszenia pokazanej na rys. 2. Otrzymane z wyjścia wzmacniacza i filtru pasmowego napięcie szumów oddziaływuje na detektor kwadratowy (przyrząd cieplny) oraz miernik.

Generator kalibracyjny (rys. 1) wytwarza sygnał sinusoidalny o częstotliwości 1000 Hz na rezystancji wynoszącej 1 Ω. Tak mała rezystancja wyjściowa źródła napięcia kalibracyjnego umożliwia jej pominięcie w porównaniu z rezystancją badanego rezystora (ta ostatnia zawiera się w zakresie od 100 Ω do 22 MΩ). Kalibracja jest konieczna między innymi ze względu na to, że wzmocnienie wzmacniacza zależy od wartości rezystancji

R_x; zastosowanie kalibracji zapewnia uzyskanie takich samych wartości wzmocnienia dla sygnałów szumów pochodzących od elementów o różnych rezystancjach.

Drugie, istotne zadanie jakie spełnia kalibracja zostanie przedstawione w rozdz. 3.3.

3.2. Definicja i wyznaczanie wskaźnika szumów prądowych

Na detektor kwadratowy, zgodnie z rys. 1, oddziaływa sygnał:

$$u_{c_k}(t) = u_{p_k}(t) + u_{w_k}(t), \tag{7}$$

gdzie:

$u_{p_k}(t)$ — napięcie pochodzące od szumów prądowych badanego rezystora,

$u_{w_k}(t)$ — napięcie pochodzące od szumów wzmacniacza oraz od szumów cieplnych rezystora.

Ponieważ wielkości przypadkowe $u_{p_k}(t)$ i $u_{w_k}(t)$ są wielkościami niezależnymi, więc całkowity kwadrat średni napięcia szumów, który oddziaływa na przyrząd cieplny, jest sumą średnich kwadratów napięć, pochodzących od poszczególnych źródeł szumów. Możemy zatem napisać:

$$\overline{u_c^2} = \overline{u_p^2} + \overline{u_w^2}, \tag{8}$$

gdzie $\overline{u_c^2}$, $\overline{u_p^2}$ i $\overline{u_w^2}$ są średnimi kwadratami napięć szumów kolejno, całkowitych, prądowych oraz wzmacniacza i cieplnych rezystora odniesionych do wejścia, czyli do zacisków rezystora o rezystancji 1 Ω (rys. 1).

Wskaźnik szumów prądowych k_{sz} jest zdefiniowany następująco:

$$k_{sz} = 20\log\frac{\sqrt{\overline{u_p^2}}}{U_z} = 20\log\frac{U_p}{U_z} \qquad \text{w dekadzie częstotliwości,} \tag{9}$$

gdzie $U_p = \sqrt{\overline{u_p^2}}$ jest napięciem skutecznym szumów prądowych w μV w zakresie jednej dekady częstotliwości, a U_z napięciem stałym przyłożonym do rezystora w V.

Pisząc wzór (9) w nieco innej postaci, a mianowicie:

$$k_{sz} = 20\log\sqrt{\overline{u_p^2}} - 20\log U_z, \tag{9a}$$

widać, że wskaźnik szumów prądowych jest różnicą napięcia skutecznego szumów oraz napięcia zasilania wyrażonych w dB, przy czym dla szumów poziomowi odniesienia, tj. 0 dB, odpowiada napięcie szumów $U_p = 1$ μV, zaś dla napięć polaryzacji poziomowi odniesienia 0 dB odpowiada napięcie $U_z = 1$ V. Tak więc stwierdzenia, że poziom szumów prądowych wynosi np. −20 dB lub 0,1 μV/V są sobie równoważne.

Jak wynika z podanej definicji, wskaźnik szumów k_{sz} jest odniesiony do idealnego, dekadowego pasma przepuszczania, tzn. takiego, które posiada kształt prostokątny oraz częstotliwość górną dziesięciokrotnie większą niż częstotliwość dolną. Należy tu podkreślić fakt, że moc lub też średni kwadrat napięcia szumów prądowych, po przejściu przez pasmo dekadowe, nie zależy od umiejscowienia tej dekady na osi częstotliwości, jako że widmo tych szumów jest typu $\frac{1}{f}$.

Ogólnie, średni kwadrat napięcia szumów po przejściu przez człon, którego kwadrat wzmocnienia napięciowego wynosi $K_u^2(f)$, jest dany wzorem:

$$\overline{u^2} = \int_0^\infty G_u(f) K_u^2(f) df, \tag{10}$$

gdzie $G_u(f)$ jest gęstością widmową kwadratu napięcia szumów.

W przypadku członu o idealnym pasmie przenoszenia wzór (10) przyjmuje postać:

$$\overline{u^2} = K_u^2 \int_{f_1}^{f_2} G_u(f) df, \tag{10a}$$

przy czym kwadrat wzmocnienia napięciowego K_u^2 posiada tu wartość stałą. Jeśli rozważane szumy są typu $\frac{1}{f}$, to zgodnie ze wzorem (4a) możemy napisać, że:

$$f G_u(f) = C U_z^2 = D, \tag{11}$$

czyli, że parametr D będący iloczynem częstotliwości f i gęstości widmowej $G_u(f)$, charakteryzujący konkretne źródło szumów prądowych, jest wielkością niezależną od częstotliwości i zależną tylko od materiału rezystora i występującego na tym rezystorze napięcia stałego U_z. Zgodnie z (10a), wyrażenie na średni kwadrat napięcia szumów prądowych po przejściu przez idealny filtr pasmowy ma postać:

$$\overline{u_{p_k}^2} = K_u^2 G_u(f) f \int_{f_1}^{f_2} f^{-1} df = K_u^2 \overline{u_p^2}. \tag{12}$$

Obliczając całkę i kładąc $K_u^2 = 1$ oraz uwzględniając (11) otrzymujemy:

$$\overline{u_p^2} = G_u(f) f \ln \frac{f_2}{f_1} = C U_z^2 \ln \frac{f_2}{f_1},$$

czyli zależność identyczną jak (5a). Przyjęcie $K_u^2 = 1$ upraszcza interpretację sygnału wyjściowego, odnosząc go do wejścia. Jeśli we wzorze (5a) wstawimy w myśl definicji wskaźnika szumów $\frac{f_2}{f_1} = 10$, to otrzymamy:

$$\overline{u_p^2} = C U_z^2 \ln 10 = D \ln 10. \tag{13}$$

Podstawiając z kolei (13) do (9), mamy dalej:

$$k_{sz} = 20 \log \frac{\sqrt{D \ln 10}}{U_z}. \tag{14}$$

Często konieczna jest także znajomość wskaźnika szumów dla dowolnego pasma częstotliwości. Można go łatwo wyznaczyć, gdy znana jest wartość k_{sz}. Obliczmy w tym celu ze wzoru (14) wielkość $\sqrt{D}$:

$$\sqrt{D} = \frac{U_z 10^{k_{sz}/20}}{\ln 10}. \tag{15}$$

Jeżeli teraz uwzględnimy, że zgodnie z zależnością (11) $D = CU_z^2$ i wstawimy (15) do (5a), to po dokonaniu prostych przekształceń otrzymamy:

$$\sqrt{\overline{u_p^2}} = U_z 10^{k_{sz}/20} \sqrt{\log \frac{f_2}{f_1}}. \tag{16}$$

Stąd wskaźnik szumów dla pasma o częstotliwościach granicznych f_2 i f_1 wyniesie:

$$\left(20\log \frac{\sqrt{\overline{u_p^2}}}{U_z}\right)_{f_1 \div f_2} = k_{sz} + 10\log\log\frac{f_2}{f_1}. \tag{17}$$

Zgodnie ze wzorem (8) średni kwadrat napięcia szumów prądowych odniesionych do wejścia dany jest wzorem:

$$\overline{u_p^2} = \overline{u_c^2} - \overline{u_w^2}.$$

Podstawiając otrzymane wyrażenie na $\overline{u_p^2}$ do (9a), mamy:

$$k_{sz} = 10\log(\overline{u_c^2} - \overline{u_w^2}) - 20\log U_z. \tag{18}$$

Jak już wspomniano przyrząd wychyłowy detektora kwadratowego jest wyskalowany w wartościach skutecznych napięcia szumów wyrażonych w decybelach. Ponieważ wychylenie tego przyrządu, przy włączonym zasilaniu rezystora, jest proporcjonalne do logarytmu z $\overline{u_c^2}$, więc jako wynik pomiaru otrzyma się:

$$k'_{sz} = 10\log\overline{u_c^2} - 20\log U_z, \tag{19}$$

czyli wartość różną od k_{sz}. Aby wyznaczyć wskaźnik szumów prądowych należy tu uwzględnić dodatkowo funkcję poprawkową zależną od różnicy $\overline{u_c^2} - \overline{u_w^2}$, a mianowicie:

$$k_{sz} = k'_{sz} - f(\overline{u_c^2} - \overline{u_w^2}), \tag{20}$$

lub wykorzystując (19),

$$k_{sz} = 10\log\overline{u_c^2} - f(\overline{u_c^2} - \overline{u_w^2}) - 20\log U_z. \tag{21}$$

Przyrównując stronami (18) i (21), otrzymujemy po przekształceniu:

$$f(\overline{u_c^2} - \overline{u_w^2}) = -10\log\left(1 - \frac{\overline{u_w^2}}{\overline{u_c^2}}\right). \tag{22}$$

Jeśli dla przejrzystości oznaczymy:

$$A = 10\log\overline{u_c^2}, \quad B = 10\log\overline{u_w^2} \quad \text{i} \quad Z = 20\log Z_z, \tag{23}$$

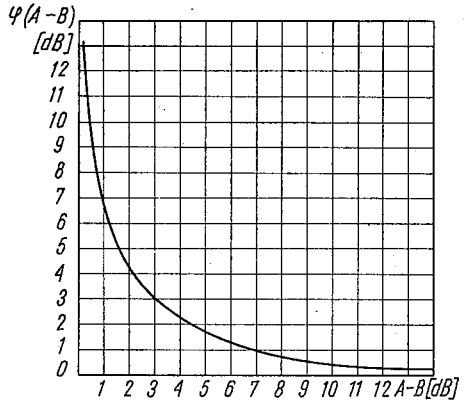
to zamiast (22) możemy napisać:

$$f(\overline{u_c^2} - \overline{u_w^2}) = -10\log\left(1 - 10^{-\frac{A-B}{10}}\right) = \varphi(A-B). \tag{24}$$

Zgodnie z (24) i (23) mamy teraz:

$$k_{sz} = A - \varphi(A-B) - Z. \tag{21a}$$

Zależność funkcji poprawkowej od różnicy $A-B$ przedstawiono na rys. 3.



Rys. 3. Wpływ różnicy pomiędzy poziomami szumów całkowitych i prądowych A—B na wartość funkcji poprawkowej

W przypadkach, gdy poziom szumów prądowych jest znaczny, tzn. gdy $A-B > 15$ dB $\left(\text{tj. } \frac{\overline{u_c^2}}{\overline{u_w^2}} > 31{,}6\right)$, to $\varphi(A-B) \cong 0$ i funkcji poprawkowej można nie uwzględniać. Jeżeli natomiast $A-B < 15$ dB, to konieczne jest uwzględnienie poprawki stosując wzór (24) lub rys. 3 oraz zależność (21a).

Tak więc w celu wyznaczenia wskaźnika szumów prądowych należy najpierw dokonać pomiaru szumów wzmacniacza, czyli określić przy wyłączonej polaryzacji rezystora wielkość B, a następnie zmierzyć szumy całkowite A po włączeniu zasilania rezystora. W zależności od uzyskanej różnicy $A-B$, gdy $A-B < 15$ dB, należy uwzględnić poprawkę oraz można ją pominąć w przypadku przeciwnym.

3.3. Pasmo rzeczywiste i pasmo odniesienia

Amplituda sygnału kalibracyjnego jest tak dobrana, że wskazania przyrządu na wyjściu detektora kwadratowego są równoważne wskazaniom, które by uzyskano, gdyby zamiast napięcia kalibracyjnego (sygnał sinusoidalny 1000 Hz) użyć sygnału szumów prądowych o napięciu skutecznym równym 1 mV, przepuszczonego przez idealny filtr dekadowy o wzmocnieniu równym jedności. W takim przypadku

$$\overline{u_{p_k}^2} = \overline{u_p^2} = D \ln 10.$$

Stąd, jeśli uwzględnimy, że $\overline{u_p^2} = 1\ \text{mV}^2$, otrzymamy:

$$D = \frac{1\ \text{mV}^2}{\ln 10}. \tag{25}$$

Taką wartość parametru D, tj. iloczynu gęstości widmowej i częstotliwości, musi posiadać źródło szumów prądowych, aby po przepuszczeniu pochodzącego od niego sygnału przez idealny filtr dekadowy o wzmocnieniu równym jedności uzyskać poziom 60 dB.

Taki sam skutek uzyska się używając jako źródła kalibracyjnego sygnału sinusoidalnego o napięciu skutecznym 1 mV i częstotliwości leżącej w środku rozważanego pasma dekadowego. Jednakże w przyrządzie znajduje się w rzeczywistości nie filtr dekadowy, lecz filtr o charakterystyce pokazanej na rys. 2, którego częstotliwość dolna i górna wynoszą odpowiednio $f_1 = 618$ Hz i $f_2 = 1618$ Hz. Dlatego też, aby wskazania miernika były zgodne z definicją wskaźnika szumów prądowych, potrzeba, aby sygnał kalibracyjny o częstotliwości 1000 Hz posiadał określoną wartość skuteczną $U = U_p = \sqrt{\overline{u_p^2}}$ różną od 1 mV. Wartość tę można wyznaczyć wstawiając (25) do (5a):

$$\overline{u_p^2} = \frac{1\ \text{mV}^2}{\ln 10} \ln \frac{f_2}{f_1} = 1\ \text{mV}^2 \log \frac{f_2}{f_1} \tag{26}$$

i uwzględniając, że $\frac{f_2}{f_1} = \frac{1618}{618} = 2{,}618$. Po obliczeniu uzyskuje się wartość $U_p = \overline{u_p^2} =$ $= 0{,}646$ mV. Wychylenie miernika od takiego sygnału szumów prądowych można traktować jako poziom 60 dB w zakresie jednej dekady sygnału szumów pochodzącego z tego samego źródła. Tak więc, równoważna wartość skuteczna sinusoidalnego napięcia kalibracyjnego wynosi $U = 0{,}646$ mV.

Analogiczny wynik można uzyskać, rozważając zamiast filtru idealnego filtr rzeczywisty. Zgodnie z zależnościami (10) i (11) średni kwadrat napięcia szumów prądowych po przejściu przez filtr rzeczywisty jest dany wzorem:

$$\overline{u_{p_k}^2} = G_u(f) f \int_0^\infty \frac{K_u^2(f)}{f} df = D \int_0^\infty \frac{K_u^2(f)}{f} df. \tag{27}$$

W celu odniesienia sygnału szumów do wejścia należy dokonać normalizacji współczynnika wzmocnienia. Otrzymamy wtedy:

$$\overline{u_p^2} = D \int_0^\infty \frac{K_u^2(f)}{K_{u_0}^2} \frac{1}{f} df, \tag{28}$$

gdzie $K_u(f)$ jest wzmocnieniem napięciowym, a K_{u_0} wzmocnieniem napięciowym dla częstotliwości środkowej f_0. Parametr D, tj. iloczyn gęstości widmowej i częstotliwości, powinien reprezentować szumy prądowe, których średni kwadrat napięcia w zakresie 1 dekady wynosi 1 mV^2 (tj. 60 dB na dekadę). Takie źródło szumów prądowych posiada wartość parametru D określoną wzorem (25). W związku z tym:

$$\overline{u_p^2} = \frac{1\ \text{mV}^2}{\ln 10} \int_0^\infty \frac{K_u^2(f)}{K_{u_0}^2} \frac{1}{f} df. \tag{28a}$$

Tak więc, aby wyznaczyć wartość skuteczną napięcia sygnału kalibracyjnego, należy dla konkretnej charakterystyki filtru obliczyć całkę we wzorze (28a), np. metodą graficzną, a następnie wyznaczyć według tego samego wzoru wartość skuteczną $U_p = \sqrt{\overline{u_p^2}}$. Wartość ta winna być równa obliczonej uprzednio wartości $U_p = 0{,}646$ mV dla filtru idealnego.

Jednakże z uwagi na fakt, że poszczególne wykonania filtru mogą wykazywać nieznaczne różnice, więc również i uzyskane wartości napięcia skutecznego U_p mogą nieco odbiegać od podanej wartości dla filtru idealnego.

Częstotliwości dolna, środkowa i górna filtru są związane zależnościami:

$$\ln\frac{f_2}{f_1} = \int_0^\infty \frac{K_u^2(f)}{K_{u0}^2}\frac{1}{f}df \tag{29}$$

oraz

$$\ln\frac{f_0}{f_1} = \ln\frac{f_2}{f_0}. \tag{30}$$

Wzór (29) stanowi warunek, jaki musi być spełniony, aby moc szumów typu $\frac{1}{f}$ przenoszona przez filtr idealny była równa mocy szumów przenoszonej przez filtr rzeczywisty. Natomiast wzór (30) oznacza, że w paśmie of f_1 do f_0 jest przenoszona taka sama moc szumów $\frac{1}{f}$, co w paśmie od f_0 do f_2. Z ostatniej zależności otrzymuje się bezpośrednio, że

$$f_0^2 = f_1 f_2, \tag{31}$$

czyli że częstotliwość środkowa filtru jest średnią geometryczną jego częstotliwości granicznych.

3.4. Dyskusja wpływu postaci funkcji gęstości widmowej szumów na wyniki pomiarów

W rozdziałach od 3.1. do 3.3. zakładano, że gęstość widmowa kwadratu napięcia szumów prądowych jest dana wzorem (4a). W ogólnym przypadku mogą jednak występować odchylenia od tej zależności i dlatego celowe jest tu rozważenie ich wpływu. Weźmy pod uwagę postać funkcji gęstości widmowej analogiczną do określonej wzorem (3), a mianowicie:

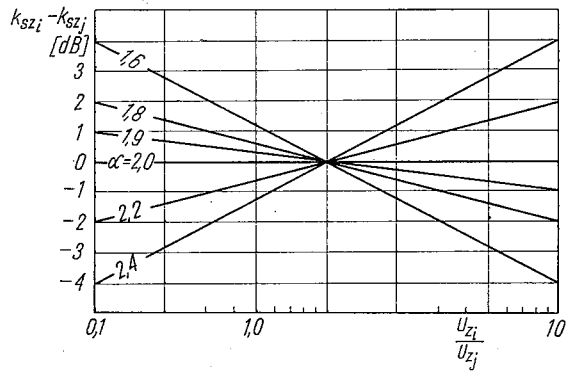
$$G_u(f) = CU_z^a f^{-\gamma}.$$

Rozważmy najpierw odchylenie γ od jedności. Jak wynika z danych literaturowych, wartości γ dla większości typów rezystorów wykazujących szumy prądowe zawierają się w zakresie od 0,8 do 1,2, przy czym najczęściej γ jest bardzo bliskie jedności [8] [9] [10]. Zgodnie z wynikami obliczeń przeprowadzonych przez Conrada i innych [2], zmiany wartości skutecznej napięcia sygnału kalibracyjnego, równoważnego sygnałowi szumów o poziomie 60 dB/dekadę, wynoszą mniej niż 1%, gdy w procesie wyznaczania $U = U_p$, opisanym w rozdz. 3.3., zamiast $\gamma = 1$ uwzględnimy $\gamma = 0{,}8$ lub 1,2. Wynika stąd, że opisana metoda zapewnia pomiar szumów prądowych praktycznie niezależnie od wartości γ w granicach zmian tego współczynnika przyjmowanych zwykle za maksymalne

Napięcie skuteczne szumów prądowych w zakresie jednej dekady częstotliwości jest niezależne od umiejscowienia tej dekady na osi częstotliwości tylko w przypadku, gdy $\gamma = 1$. Gdy $\gamma \neq 1$, to umiejscowienie dekady na osi częstotliwości jest czynnikiem, który

powoduje powstanie pewnych błędów. Jeżeli za dekadę odniesienia przyjmiemy taką dekadę, której częstotliwość środkowa f_0 wynosi 1000 Hz (jak w definicji k_{sz}), to w przypadku $\gamma = 0{,}8$ wartość skuteczna napięcia szumów prądowych jest dla $f_0 = 632$ Hz o 5% mniejsza, a dla $f_0 = 10$ kHz aż o 60% większa niż ta, którą uzyska się przy założeniu, że $\gamma = 1$ [2]. Jak widać z tego przykładu, powstałe błędy są do pominięcia, gdy f_0 mało różni się od 1000 Hz oraz osiągają bardzo duże wartości dla f_0 znacznie odległych od 1000 Hz.

Jednakże rezystory wykazujące szumy prądowe posiadają bardzo rzadko γ różne od 1 więcej niż o pojedyncze procenty [8] [9] [10] i dlatego wpływ zmian γ można uznać w większości zastosowań za pomijalnie mały.



Rys. 4. Zależność wskaźnika szumów prądowych od napięcia zasilania dla różnych wartości wykładnika potęgowego α

Jak wynika ze wzoru (9), wskaźnik szumów prądowych jest w ogólnym przypadku proporcjonalny do napięcia zasilania U_z w potędze $\frac{\alpha}{2}$. Ilustrację tego faktu stanowi rys. 4 [2]. W poprzednich rozdziałach przyjmowano, że $\alpha = 2$. Jeśli takie założenie jest słuszne, to wskaźnik szumów jest niezależny od napięcia zasilania. Wartości graniczne, jakie może zwykle przyjmować α, wynoszą 1,6 i 2,4, przy czym są obserwowane bardzo rzadko; dla większości rezystorów α jest bardzo bliskie 2 [9] [10]. Z rys. 4 widać, że przy $\alpha = 1{,}9$ dziesięciokrotna zmiana napięcia zasilania powoduje zmianę wskaźnika szumów tylko o 1 dB. Wynika stąd, że poza rzadko stosowanymi typami rezystorów i warunkami pracy, wskaźnik szumów prądowych zdefiniowany w rozdz. 3.2 może być z powodzeniem traktowany jako niezależny od napięcia zasilania w szerokim zakresie zastosowań.

4. WYNIKI BADAŃ

Zgodnie z definicją wskaźnika szumów prądowych oraz zasadą jego pomiaru, wartość skuteczną napięcia szumów prądowych we wzorze (9) określa się w zakresie jednej dekady częstotliwości, czyli dla $\frac{f_2}{f_1} = 10$. Na podstawie wzoru (13) mamy zatem:

$$\sqrt{\overline{u_p^2}} = U_z \sqrt{C \ln 10}\,.$$

Tablica 1

Wpływ ułamka x_1 w paście na wielkości charakteryzujące właściwości szumowe badanych warstw

x_1	R [Ω]	δ [μm]	U_z [V]	I_z [mA]	$\sqrt{\overline{u_p^2}}$ [μV]	$\sqrt{\overline{i_p^2}}$ [nA]	$\sqrt{\overline{u_p^2}}\,U_z$ [μV/V]	k_{sz} [dB]	C [10^{-14}]
0,333	132	14,23	3,80	28,80	1,192	9,030	0,3140	—10,05	4,28
0,400	1125	11,43	17,37	15,43	3,059	2,720	0,1760	—15,08	1,344
0,500	1415	11,83	19,49	13,78	1,793	1,268	0,0920	—20,70	0,367
0,600	1140	12,24	17,37	15,25	0,920	0,807	0,0530	—25,50	0,122
0,667	895	15,77	14,28	15,96	0,416	0,465	0,0291	—30,48	0,0388
0,750	680	15,91	13,02	19,15	0,306	0,450	0,0231	—31,70	0,0232

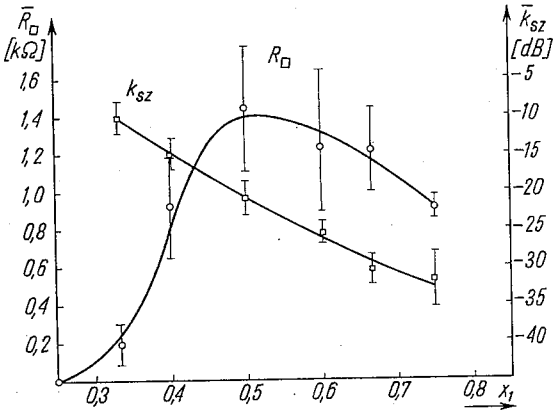
Stosując tę zależność można wyznaczyć stałą materiałową C. Jeśli chodzi o wartości $\overline{u_p^2}$ i U_z, to są one znane bezpośrednio na podstawie pomiarów.

W tablicy 1 przedstawiono wyniki pomiarów i obliczeń, dotyczące wielkości charakteryzujących warunki pomiaru wskaźnika szumów prądowych oraz wartości stałej materiałowej C, obliczone w opisany wyżej sposób, dla typowych struktur rezystywnych, wykonanych z kompozycji o różnym składzie, przy ustalonych pozostałych parametrach technologicznych. W składzie wyjściowym zmieniano ułamek atomowy palladu w składnikach metalicznych pasty określony wzorem:

$$x_1 = \frac{N_{Pd}}{N_{Pd}+N_{Ag}}, \tag{32}$$

gdzie N_{Pd} i N_{Ag} oznaczają liczby atomów odpowiednio palladu i srebra w kompozycji, przy stałej zawartości szkliwa wynoszącej 55% wagowo. Porównując wyniki podane w ostatnich trzech kolumnach tablicy 1 dla poszczególnych wartości ułamka x_1 można zauważyć, że wielkości charakteryzujące właściwości szumowe warstw, tzn. wartość stosunku napięć $\sqrt{\overline{u_p^2}}/U_z$ wyrażona w μV/V, lub wartość równoważnego wskaźnika k_{sz} wyrażona w dB, czy wreszcie wartość stałej materiałowej C, maleją ze wzrostem ułamka atomowego Pd w paście. Natomiast charakter zmian $\sqrt{\overline{u_p^2}}$ i $\sqrt{\overline{i_p^2}}$ jest niekiedy zupełnie inny, aniżeli wymienionych uprzednio trzech wielkości, ponieważ różnym strukturom odpowiadają różne rezystancje oraz napięcia i prądy zasilania oznaczone w tablicy odpowiednio przez R, U_z i I_z. Tak więc, w celu dokonania oceny elementów rezystywnych z punktu widzenia ich właściwości szumowych, należy porównywać napięcia szumów odniesione do 1 V napięcia zasilania, czyli wielkości $\sqrt{\overline{u_p^2}}/U_z$, wyrażone w μV/V lub też w dB. Dzieląc licznik i mianownik wyrażenia pod logarytmem we wzorze (9) przez rezystancję rezystora, otrzyma się równoważną wielkość $\sqrt{\overline{i_p^2}}/I_z$ w μA/A, równą co do wartości wskaźnikowi szumów prądowych. Wielkość taką można by uzyskać przeprowadzając pomiar średniego kwadratu natężenia prądu *zwarciowego*, nie zaś jak w opisanym mierniku średniego kwadratu napięcia w stanie rozwarcia (bez obciążenia).

Opierając się na podstawowych rozważaniach można wykazać, że przy ustalonych pozostałych parametrach strukturalnych, wartość stałej materiałowej C jest odwrotnie proporcjonalna do grubości warstwy [9]. Ponieważ dla różnych wartości ułamka atomowego palladu x_1 grubości warstw różniły się nieco (tablica 1), więc dla wyeliminowania wpływu grubości warstwy rezystywnej przeprowadzono korekcję wartości stałej materiałowej C oraz wskaźnika szumów, uwzględniając, że $C \approx \frac{1}{\delta}$ i odnosząc wszystkie wyniki do jednej grubości równej $\delta = 12{,}24$ μm.



Rys. 5. Wpływ ułamka atomowego palladu w kompozycji wyjściowej na wskaźnik szumów i rezystancję warstw Pd/PdO/Ag

Na rys. 5 przedstawiono zależności wskaźnika szumów i rezystancji na kwadrat od ułamka x_1. Wartości $\bar{k}_{sz}$ i $\bar{R}_\square$ podano na rysunku jako średnie dla 24 elementów w serii wraz z odpowiadającymi im przedziałami $\pm 3s$ (s — odchylenie standardowe), uwzględniając korekcję na grubość warstwy.

W tablicy 2 przedstawiono wyniki pomiarów i obliczeń, dotyczące wpływu parametrów wypalania, a mianowicie: czasów — narastania (t_n), wygrzewania w temperaturze szczytowej (t_w) i opadania (t_o) oraz temperatury szczytowej wypalania ϑ przy $x_1 = 0{,}5$, na wartość wskaźnika szumów i rezystancję badanych warstw. Jak widać z rys. 5 i tablicy 2, największe znaczenie z punktu widzenia poziomu szumów, a także i rezystancji warstw mają dwa czynniki, a mianowicie ułamek atomowy palladu w składnikach metalicznych pasty oraz temperatura szczytowa wypalania. Obserwuje się przy tym, że ze wzrostem x_1 szumy wyraźnie maleją oraz rosną ze wzrostem ϑ. Pozostałe parametry wypalania mają stosunkowo niewielki wpływ na poziom szumów.

Jak wynika z przeprowadzonych ostatnio badań, dotyczących relacji struktura — parametry elektryczne warstw Pd/PdO/Ag, przyczyna charakteru uzyskanych zależności tkwi w ilościowym składzie fazowym badanych struktur. Zasadnicze znaczenie ma tu udział objętościowy szkliwa, które w porównaniu z pozostałymi fazami, tzn. tlenkiem palladawym i roztworem stałym pallad-srebro, charakteryzuje się znacznie większą rezystywnością i jako faza izolująca determinuje w głównej mierze liczbę nieciągłości, lub — co się z tym wiąże — *szumiących kontaktów* w strukturze warstwy.

Tablica 2

Zależności wskaźnika szumów i rezystancji na kwadrat od parametrów wypalania warstw Pd/Pd O/Ag

Zmieniany parametr wypalania		$\overline{R_\square}$ [Ω]	$\frac{3s}{\overline{R_\square}}$ [%]	$\overline{k_{sz}}$ [dB]	$\frac{3s}{\overline{k_{sz}}}$ [%]
t_n [min]	5	1380	25,66	−20,00	8,32
	12,5	1472	24,69	−20,90	10,20
	25	1455	20,28	−21,30	7,07
t_w [min]	6	1058	14,86	−23,00	5,96
	10	1199	24,92	−21,40	8,03
	20	1472	24,69	−20,90	10,20
t_o [min]	7	1358	32,75	−21,20	9,96
	17,5	1472	24,69	−20,90	10,20
	35	1447	27,85	−21,50	13,34
ϑ [°C]	700	1129	18,93	−22,01	9,34
	750	1472	24,69	−20,90	10,20
	800	1682	39,33	−17,89	15,27

Jeżeli chodzi o rezystancję na kwadrat, to należy tu wziąć pod uwagę superpozycję wpływu zawartości PdO w PdO+Pd-Ag oraz udziału objętościowego szkliwa w materiale warstwy, przy czym dla $x_1 < 0{,}55$ (rys. 5) przeważa pierwszy czynnik, zaś dla $x_1 > 0{,}55$ drugi. Natomiast w przypadku temperatury szczytowej wypalania, dominujący wpływ na $R_\square$ wywierą prawdopodobnie drugi czynnik, czyli wzrost udziału objętościowego szkliwa, wynikły ze zmniejszenia się zawartości PdO w warstwie.

5. PODSUMOWANIE

Opisany sposób określania i oceny właściwości szumowych rezystorów, wynikający z samej koncepcji pomiaru szumów przyrządem typu Quan Tech, jest szczególnie przydatny w warunkach produkcyjnych oraz gdy chodzi o porównywanie jakości różnych struktur rezystywnych wykonywanych różnymi technikami. Definicja wskaźnika szumów prądowych dana wzorem (9) standaryzuje pomiar, sprowadzając wielkości charakteryzujące szumy do określonego napięcia zasilania rezystora równego 1 V oraz do konkretnego pasma, a mianowicie pasma dekadowego o częstotliwości środkowej równej 1000 Hz. Należy tu zwrócić uwagę na fakt, że porównywanie wskaźników szumów różnych elementów jest formalnie słuszne wtedy, gdy ich gęstość widmowa szumów podlega prawu danemu wzorami (4) lub (4a). Jednakże jak wynika z danych literaturowych [2], [8] [9] [10], większość rezystorów wykazujących szumy prądowe stosuje się do tego prawa bardzo

dokładnie, a występujące ewentualnie odchyłki są pomijalne w rozważanym zakresie zastosowań.

Poza tym, ponieważ między niezawodnością rezystorów kompozycyjnych a wskaźnikiem szumów, podobnie jak i między niezawodnością i nieliniowością napięciową wyznaczoną przez tzw. wskaźnik zawartości trzeciej harmonicznej, istnieje określona korelacja [3] [5], więc pomiary wskaźnika szumów serii rezystorów stwarzają możliwości selekcji i odrzucania elementów potencjalnie zawodnych w przypadkach, gdy wskaźnik szumów przekroczy pewne, z góry określone granice. Pozwala to często uniknąć dokonywania żmudnych badań niezawodnościowych i przeprowadzać wspomnianą selekcję bezpośrednio w procesie produkcyjnym.

Uzyskane wyniki badań wpływu wybranych parametrów procesu technologicznego na właściwości szumowe warstw rezystywnych Pd/PdO/Ag wykazały, że najbardziej krytycznymi spośród uwzględnionych parametrów są: zawartość palladu w składnikach metalicznych pasty określona ułamkiem x_1 oraz temperatura szczytowa wypalania ϑ. Jak wynika z przeprowadzonych badań strukturalnych, wspólną przyczyną uzyskanych zależności jest wpływ udziału objętościowego szkliwa w strukturze warstwy. Wynika stąd, że zasadniczą przyczyną występowania szumów prądowych w warstwach rezystywnych Pd/PdO/Ag jest nieciągłość ich struktury, związana z obecnością fazy izolującej szkliwa. Bowiem nawet, gdy zawartość szkliwa w paście jest ustalona (w przypadku wszystkich badanych kompozycji wynosiła ona 55% wagowo), to występujące w związku ze zmianami x_1 i ϑ stosunkowo niewielkie zmiany udziału objętościowego szkliwa, spowodowane powstaniem PdO w warstwie, wydają się mieć generalne znaczenie. Dlatego też uzyskiwanie struktur rezystywnych o dużych rezystancjach na kwadrat, co się wiąże z koniecznością stosowania dużych zawartości szkliwa w samej paście, pociąga za sobą pogorszenie właściwości szumowych warstw.

Aby osiągnąć zmniejszenie poziomu szumów badanych elementów rezystywnych, co jest szczególnie pożądane w przypadku dużych wartości rezystancji na kwadrat, konieczne jest dokonanie odpowiednich zmian w procesie wytwarzania warstw, a przede wszystkim w kompozycji wyjściowej. W tym celu należy dokładnie poznać mechanizm i przyczyny występowania szumów prądowych w rezystorach Pd/PdO/Ag. Wymaga to przeprowadzenia oddzielnych badań, dotyczących określania wpływu struktury warstw na ich właściwości szumowe, Wydaje się, że szczególnie przydatne będzie w tym względzie zbadanie zależności widmowej szumów od częstotliwości i napięcia zasilania dla różnych składów fazowych warstw.

LITERATURA CYTOWANA W TEKŚCIE

1. J. S. Bendat, Osnowy tieorii słuczajnych szumow i jejo primienienije, Izdat. „Nauka", Moskwa 1965 (tłum. z ang.).
2. G. T. Conrad, N. Newmann, A. P. Stainsbury, A Recommended Standard Resistor Noise Test System, IRE Transactions of the Professional Group on Component Parts, Vol. CP-7, Nr 3, September 1960.
3. J. G. Curtis, Current Noise Test Indicate Resistor Quality, International Electronics, May 1962.
4. W. Kalita, A. Kusy, Technologia grubowarstwowych mikroukładów biernych, Prace Techniczne RTPN, nr 1, 1969.

5. P. L. Kirby, Non-linearity Measurements Improve Resistor Reliability, Electronics Components, May 1967.
6. A. Kusy, Porównanie charakterystyk wypalania i rezystywności rezystorów grubowarstwowych wykonanych z kompozycji: Pd+szkliwo, Pd0+szkliwo, Pd+Ag+szkliwo, PdO+Ag+szkliwo, Prace ITE Politechniki Wrocławskiej, nr 1, 1969.
7. A. Kusy, Struktura i własności rezystorów grubowarstwowych na bazie palladu i srebra, Rozprawy Elektrotechniczne, nr 3, 1971.
8. J. Kotecki, Rezystory, WKŁ, Warszawa 1970.
9. Van der Ziel, Fluktuacji w radiotiechnikie i fizikie, Goseniergoizdat, Moskwa 1958 (tłum. z ang.).
10. A. Zaklikiewicz, Pomiary szumów w opornikach, Przegląd Elektroniki, nr 11, 1967.

A. KUSY

MEASUREMENT OF NOISE PROPERTIES OF Pd/PdO/Ag FILMS BY QUAN TECH TYPE METER

Summary

Fundamental quantities characterizing noise, which occurs in resistive films, particularly taking into consideration the $1/f$ type noise, have been presented. A detailed analysis of measuring methods of current noise, applying the Quan Tech type meter and the way of evaluating the k_{sz} index, which is the „noise quality" measure of composition resistive films, has been given. A definition of the current noise index k_{sz}, standarizes the measurements, referring the root mean square voltage to the defined, constant supply voltage equal 1 V and to the decade pass band of frequency. A discussion of the influence of the form of the noise power spectral density function on the results of measurements has been carried out. The results of the investigations dealing with the influence of some technological parameters, namely palladium atomic fraction x_1 in the metallic constituents of the paste and also the times and peak temperature of firing on the quantities characterizing the noise properties of Pd/PdO/Ag resistive films have been given. An analysis of the obtained results proved that the greatest significance from the noise level viewpoint and also the resistivity of the films have two agents: Pd atomic fraction x_1 and the temperature of firing ϑ. A conception of the explanation of the obtained dependences on the basis of the performed phase analysis of the films has been given.

A. KUSY

BEMESSUNG VON RAUSCHEIGENSCHAFTEN DER Pd/PdO/Ag-WIDERSTANDSSCHICHTEN MITTELS MEßGERÄT TYP QUAN TECH

Zusammenfassung

Es wurden die das Rauschen charakterisierenden Grundgrößen für das in Widerstandsschichten auftretende Rauschen unter besonderer Berücksichtigung der $1/f$-Rauschart behandelt. Es wurde auch eine genaue Analyse der Meßmethodik mittels Meßgerät Typ Quan Tech für Stromrauschen sowie die Bestimmungsmethoden für den k_{sz}-Index, der ein Maß für die „Rauschqualität" ist, gegeben. Die Definition des Stromrauschindex k_{sz} standarisiert die Messungen, indem sie die effektive Rauschspannung der bekannten Versorgungsspannung von 1 V sowie der Banddekadenfrequenz beschreibt.

Der Einfluß der Funktionsform der Rauschleistungsspektraldichte von Widerstandselementen auf die Meßergebnisse wurde diskutiert. Untersuchsergebnisse wurden angegeben, die den Einfluß mancher technischer Parameter erklären, und zwar eines Atombruchs vom Palladium in metallischen Bestandteilen der x_1-Paste sowie der Verbrennungszeiten und -höchstemperatur auf die Größen, die die Rauscheigenschaften der Pd/PdO/Ag-Resistanzschichten charakterisieren.

Eine Analyse der erhaltenen Ergebnisse zeigte, daß die größte Bedeutung vom Gesichtspunkt des Stromrauschpegels, aber auch der Schichtresistanz zwei Faktoren haben: der Bruchteil Pd x_1 sowie die Verbrennungstemperatur. Es wurde eine Konzeption für die Erklärung der Beziehungen in Anlehnung an die Phasenschichtanalyse dargestellt.

А. КУСЫ

ИЗМЕРЕНИЕ ШУМОВЫХ СВОЙСТВ РЕЗИСТИВНЫХ ПЛЕНОК Pd/PdO/Ag ПРИБОРОМ ТИПА Quan Tech

Р е з ю м е

Представлены основные величины характеризующие шумы, которые выступают в резистивных пленках, с особым учетом шумов типа $1/f$. Приведен подробный анализ метода измерения токовых шумов прибором типа Quan Tech и способа определения индекса k_{sz}, являющегося мерой „шумового качества" композиционных резистивных пленок. Дефиниция индекса токовых шумов k_{sz} стандартизирует измерение и приводит эффективное напряжение шумов к определенному напряжению питания равному 1 V b декадовой полосе частоты. Проведена дискуссия влияния вида функции спектральной плотности токовых шумов резистивных элементов на результаты измерений. Поданы результаты исследования относительно влияния некоторых технологических параметров, а в частности атомной дроби палладия x_1 в металлических слагаемых пасты, а также времени и температуры отжига на величины характеризующие шумовые свойства резистивных пленок Pd/PdO/Ag. Анализ полученных результатов показал, что самое большое значение с точки зрения уровня шумов, а также и сопротивления пленок, имеют два фактора: дробь Pd x_1 и температура отжига ϑ. Представлена концепция объяснения полученных зависимостей на основании проведенного фазового анализа.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 2, ss. 331—358

621.382

Elektromagnetyczna teoria promieniowania laserów złączowych

MAREK OSIŃSKI (WARSZAWA)

Instytut Fizyki Polskiej Akademii Nauk, OD-2

Otrzymano 2.4.1973

Celem artykułu jest przedstawienie polskiemu czytelnikowi rezultatów, jakie do tej pory udało się uzyskać traktując laser półprzewodnikowy na złączu *p-n* jako falowód. W najprostszym przypadku falowód ten składa się z warstw różniących się wartością zespolonej stałej dielektrycznej. Warstwa czynna, charakteryzująca się ujemnym przewodnictwem (co jest równoważne ujemnemu współczynnikowi absorpcji, a więc wzmacnianiu promieniowania elektromagnetycznego), jest otoczona warstwami biernymi o dodatnim przewodnictwie. Jednocześnie warstwa czynna ma wyższy współczynnik załamania niż warstwy bierne, co powoduje efekt falowodowy w laserze złączowym. Pole elektromagnetyczne wewnątrz lasera otrzymuje się rozwiązując równanie falowe dla takiego falowodu. Pozwala to poprawnie ocenić skład widmowy promieniowania (strukturę modową) oraz rozkłady natężenia pola elektromagnetycznego w pobliżu powierzchni emitującej promieniowanie oraz w strefie dalekiej (rozkład kątowy).

Model ten można ulepszać zakładając bardziej zbliżony do faktycznego kształt zależności zespolonej przenikalności dielektrycznej od położenia. Otrzymuje się wtedy zgodne z eksperymentem wyniki dotyczące polaryzacji promieniowania i kolejności wzbudzania się poszczególnych typów drgań ze wzrostem prądu przepływającego przez złącze (selekcja modów). Uwzględnienie zależności zespolonej przenikalności dielektrycznej od położenia nie tylko w kierunku prostopadłym do złącza *p-n*, lecz także w płaszczyźnie złącza, daje ograniczenie obszaru generującego promieniowanie do wąskiego włókna. Teoria elektromagnetyczna wyjaśnia zatem w prosty sposób zasadnicze właściwości promieniowania półprzewodnikowego lasera złączowego.

1. WSTĘP

Zagadnienia, z jakimi mamy do czynienia w teorii półprzewodnikowego lasera złączowego można podzielić na trzy zasadnicze działy:

1) identyfikacja poziomów, pomiędzy którymi zachodzą przejścia laserowe oraz oszacowanie prawdopodobieństwa tych przejść,

2) rozkład przestrzenny i rekombinacja nośników w obszarze złącza,

3) mody promieniowania elektromagnetycznego.

Rozwiązanie trzeciego problemu w zasadzie zależy od rozwiązania pierwszych dwóch; można by więc oczekiwać, że niewiele można osiągnąć nie znając dokładnych opracowań zagadnień 1) i 2). Wbrew pozorom można jednak stosować proste modele lasera złączowego, pozwalające na uzyskiwanie niektórych rozwiązań problemu 3).

Celem niniejszej pracy jest zapoznanie Czytelnika z różnymi modelami lasera półprzewodnikowego, stosowanymi przy rozwiązywaniu problemu 3) i zasadniczymi wynikami, jakie osiągnięto dzięki ich przyjęciu, począwszy od pierwszych prac na ten temat, aż do opublikowanych w latach ostatnich. Niezbędne do zrozumienia tego artykułu wiadomości na temat budowy, działania i różnych nowych typów laserów złączowych można znaleźć w artykułach [1, 2]. Czytelnikowi, pragnącemu pogłębić swoje wiadomości z dziedziny laserów półprzewodnikowych, można polecić obszerniejsze opracowania [3] i [4], a także książkę C. H. Goocha [5]. Tam też można znaleźć wyczerpującą bibliografię na ten temat.

W artykule tym ograniczymy się do zagadnienia modów promieniowania elektromagnetycznego, powstającego w laserach złączowych z arsenku galu lub heterozłączowych GaAs-$Al_xGa_{1-x}As$, ponieważ obecnie te materiały są najczęściej stosowane. Pominiemy model uwzględniający własności piezoelektryczne arsenku galu [6], który jest niezbędny dopiero przy analizie modulacji światła laserowego przez wykorzystanie efektu Pockelsa [7].

Pierwsze lasery półprzewodnikowe były laserami homozłączowymi z GaAs [8÷11]. Wkrótce potem ukazały się prace analizujące teoretycznie własności promieniowania laserowego [12÷15]. W pracach tych przyjęto wyidealizowany model lasera narzucony przez strukturę diody laserowej, który będziemy nazywać modelem trójwarstwowym. Model ten jest w miarę szczegółowo opisany w drugiej części pracy. Pomimo swej prostoty, oddaje on dosyć dobrze rozkłady pola promieniowania w strefie bliskiej i rozkłady kątowe. Model trójwarstwowy stosowano też do opisu promieniowania laserów z pojedynczym i podwójnym heterozłączem [16, 17].

W trzeciej części artykułu zajmiemy się krótko zagadnieniem polaryzacji promieniowania laserowego dla laserów homo- i heterozłączowych.

W opisanym w części drugiej modelu trójwarstwowym zawsze mod podstawowy ma najniższy próg wzbudzenia, a następne mody powinny się wzbudzać kolejno ze wzrostem prądu [17]. Tymczasem w laserach z szeroką wnęką rezonansową (LOC — Large Optical Cavity Laser), a także w laserach heterozłączowych obserwuje się znacznie wcześniejsze wzbudzanie się modów wyższych rzędów, niż należałoby oczekiwać z modelu trójwarstwowego (lub czterowarstwowego dla lasera z szeroką wnęką rezonansową) [16, 18]. Poprawne wyniki otrzymuje się dzieląc obszar czynny w laserze złączowym na dwie warstwy różniące się wartościami współczynnika wzmocnienia — jedna z nich może być nawet pochłaniająca [19]. Ten model lasera opisany jest w części czwartej.

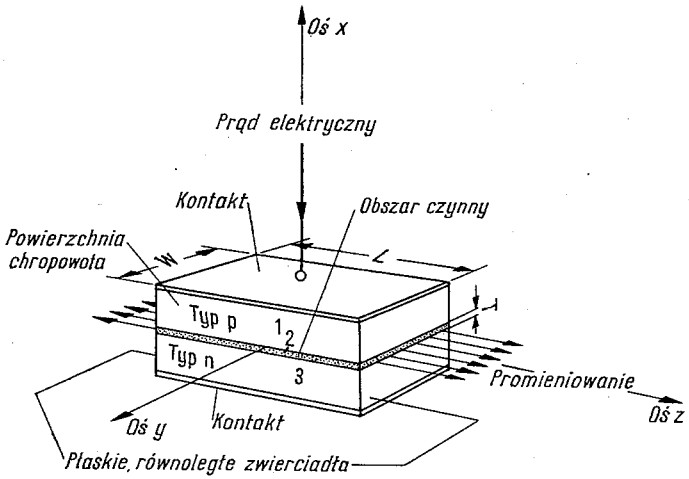
Wreszcie w ostatniej części podana jest teoria lasera z paskową geometrią kontaktu omowego (który dla zwięzłości będziemy nazywać dalej laserem paskowym), nie wykazującego struktury włóknistej promieniowania. Mody promieniowania takiego lasera posiadają symetrię hermitowsko-gaussowską, co wskazuje na efekt skupiania promieniowania w obszarze czynnym [20]. Podany w części piątej model lasera paskowego pozwala wyjaśnić dokładnie strukturę widmową promieniowania laserowego, oddaje też dobrze rozkłady kątowe i pole w strefie Fresnela (bliskiej) [21÷23].

Artykuł ten nie wyczerpuje tematu w całości. Pominięte zostały takie zagadnienia, jak modulacja promieniowania laserowego, szumy lasera, wydajność kwantowa, synchronizacja modów (mode-locking), koherencja. Czytelnika interesującego się tymi problema-

mi można odesłać do artykułów [24÷28]. Ze względu na konieczną zwięzłość, w wielu miejscach podane są tylko odnośniki do prac oryginalnych, które na ogół są łatwo dostępne. Autor starał się natomiast doprowadzić Czytelnika do najbardziej bieżących zagadnień, jednocześnie zapoznając Go z „historią" teorii modów w laserze złączowym.

2. MODEL TRÓJWARSTWOWY LASERA ZŁĄCZOWEGO

Najczęściej stosowaną strukturą półprzewodnikowego lasera złączowego jest struktura Fabry'ego-Pérota z dwoma zwierciadłami, utworzonymi przez przeciwległe ścianki prostopadłe do płaszczyzny złącza p—n. Laserami o innej geometrii w artykule tym nie będziemy się zajmować. Laser złączowy o rezonatorze typu Fabry'ego-Pérota przedstawiony jest schematycznie na rys. 1. Jeśli przez taką diodę laserową przepuścimy odpowiednio duży



Rys. 1. Schemat półprzewodnikowego lasera złączowego o rezonatorze typu Fabry'ego-Pérota
L — długość rezonatora, W — jego szerokość, T — grubość. Zaznaczone są osie układu współrzędnych, przyjętego do opisu rozchodzenia się powstających fal elektromagnetycznych

prąd w kierunku przewodzenia, w obszarze złącza p—n wytwarza się inwersja obsadzeń (elektrony są wstrzykiwane ze strony n na stronę p, a dziury w przeciwnym kierunku); wskutek rekombinacji promienistej elektronów i dziur w warstwie oznaczonej na rysunku cyfrą 2 wytwarzane jest promieniowanie elektromagnetyczne. Fotony, które padają na zwierciadła, ulegają częściowo odbiciu, co umożliwia wzmacnianie promieniowania wskutek emisji wymuszonej w warstwie czynnej.

Opisana struktura lasera złączowego narzuca prosty model, jaki należy przyjąć do opisu wytwarzanego promieniowania. Diodę laserową dzielimy na trzy warstwy, równoległe do płaszczyzny złącza p—n. Warstwa środkowa o grubości T jest warstwą czynną, którą, gdy prąd przepływający przez diodę przewyższa wartość progową, można charakteryzować przez podanie ujemnego współczynnika absorpcji (ponieważ promieniowanie jest w tej warstwie wytwarzane, a nie pochłaniane). Obszar czynny jest z obu stron ograniczony obszarami biernymi, w których współczynnik absorpcji jest dodatni. Długość

rezonatora Fabry'ego-Pérota oznaczać będziemy przez L, a szerokość przez W. Aby opisać rozchodzenie się promieniowania wewnątrz diody, wybierzemy układ współrzędnych kartezjańskich, którego oś z pokrywa się z osią rezonatora, zaś oś x jest prostopadła do płaszczyzny złącza. Początek układu współrzędnych jest tak dobrany, że zwierciadła mają współrzędne z odpowiednio $-L$, 0. Płaszczyznę yz równoległą do złącza będziemy nazywać płaszczyzną poziomą, natomiast prostopadłą do złącza płaszczyznę xz — płaszczyzną pionową.

2.1. Zależność przenikalności dielektrycznej od położenia

W ramach takiego modelu musimy jeszcze przyjąć postać zależności zespolonej przenikalności dielektrycznej od położenia we wszystkich trzech warstwach. Najprostszym założeniem jest przyjęcie, że w każdej z wymienionych wyżej warstw zespolona przenikalność dielektryczna ma inną, lecz stałą w danej warstwie, wartość (model taki będziemy nazywać trójwarstwowym). Możemy wtedy zapisać

$$\varepsilon_k = \varepsilon_k' + i\varepsilon_k'', \qquad k = 1, 2, 3, \tag{2.1}$$

gdzie ε_k', ε_k'' są rzeczywiste i nie zależą od położenia; wskaźnik k numeruje poszczególne warstwy, zgodnie z rysunkiem 1. Zachodzą znane związki między ε_k', ε_k'' a współczynnikami załamania n_k i absorpcji α_k (względnie przewodnictwem σ_k) w tych samych warstwach:

$$\varepsilon_k' = n_k^2 - \left(\frac{c\alpha_k}{2\omega}\right)^2, \tag{2.2}$$

$$\varepsilon_k'' = \frac{c\alpha_k n_k}{\omega} = \frac{\sigma_k}{\omega}; \tag{2.3}$$

ω oznacza tu częstość kołową generowanego promieniowania, c — prędkość światła w próżni.

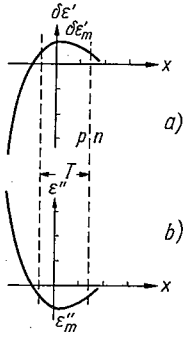
W pierwszych pracach dotyczących elektromagnetycznej teorii promieniowania lasera złączowego [12÷15] brano pod uwagę tylko nieciągłość ε_k'' przy przechodzeniu od jednej warstwy do drugiej, natomiast ε_k' pozostawało stałe. Dawało to jednak straty dyfrakcyjne (wynikające z przenikania promieniowania do pochłaniających obszarów 1, 3) o rząd wielkości większe niż wynikające z doświadczenia [12, 29].

Aby otrzymać efekt falowodowy, a przez to jednocześnie zmniejszenie strat dyfrakcyjnych, konieczne jest również uwzględnienie nieciągłości ε_k' przy przejściu od warstwy czynnej do warstw sąsiednich [30]; w obszarach biernych w laserze homozłączowym jest ona niższa o ok. 10^{-3} głównie ze względu na występowanie nośników swobodnych [31], w laserach heterozłączowych różnice te są kilkakrotnie większe [32]. Taki właśnie model opiszemy szczegółowo w tej części. Był on przyjmowany zarówno do opisu promieniowania lasera homozłączowego [15, 31, 33÷35], jak i laserów heterozłączowych [16, 17, 36÷42].

Model trójwarstwowy jest oczywiście uproszczeniem realnej sytuacji. F. Stern [30] zwrócił uwagę na fakt, że współczynnik załamania w warstwie czynnej zmienia się ze zmianą położenia w kierunku prostopadłym do złącza. Przyczyną tego jest zmienna koncen-

tracja domieszek w tym kierunku, a zatem i wiążąca się z tym zmienna szerokość przerwy energetycznej. Ścisłe potraktowanie tego problemu jest trudne ze względu na to, że rozkład domieszek w obszarze czynnym nie jest dokładnie znany i dla każdej diody może być inny. Stern [30] próbował ominąć te trudności, dzieląc obszar czynny na trzy warstwy różniące się wartościami współczynnika załamania, które można oszacować ze związków dyspersyjnych Kramersa-Kroniga [43].

W części trzeciej opiszemy bliżej model zastosowany przez J. Hatza i E. Mohna [44], w którym zależność przenikalności dielektrycznej od położenia wzdłuż osi x ma w warstwie



Rys. 2. Przebieg funkcji: a) $\delta\varepsilon'(x)$ i b) $\varepsilon''(x)$ w modelu R. G. Allachwerdiana i in. [45]
T — grubość obszaru czynnego

czynnej kształt trójkąta. R. G. Allachwerdian i in. [45] przyjęli jeszcze bardziej skomplikowany model, w którym zespolona przenikalność dielektryczna jest wyrażona wzorem

$$\varepsilon(x) = \varepsilon_0' + \delta\varepsilon'(x) + i(\varepsilon_0'' + \varepsilon''(x)), \tag{2.4}$$

gdzie ε_0' jest częścią rzeczywistą przenikalności dielektrycznej w obszarach biernych (przyjęto, że jest ona jednakowa w materiale typu n i p), ε_0'' opisuje straty na promieniowanie, zaś $\delta\varepsilon'(x)$ i $\varepsilon''(x)$ dane są wzorami

$$\delta\varepsilon'(x) = \delta\varepsilon_m'(2e^{-\beta x} - e^{-2\beta x}), \tag{2.5}$$

$$\varepsilon''(x) = \varepsilon_m''(e^{-2\beta x} - 2e^{-\beta x}) \tag{2.6}$$

i są przedstawione na rys. 2; $\delta\varepsilon_m'$ i ε_m'' są wziętymi z doświadczenia wartościami $\delta\varepsilon'(x)$ i $\varepsilon''(x)$ w punktach ekstremalnych, zaś współczynnik β charakteryzuje grubość warstwy czynnej. Kształt krzywych przedstawionych na rys. 2 jest zbliżony do danych eksperymentalnych [46÷48] dla GaAs typu p i n. Model ten staje się bardzo przydatny przy analizie promieniowania lasera pracującego impulsowo. Do opisu lasera pracującego w sposób ciągły w zupełności wystarcza prosty model trójwarstwowy.

2.2. Rozwiązywanie równań Maxwella

Mając już przyjęty model struktury lasera złączowego można przystąpić do szukania rozwiązań równań Maxwella, odpowiadających modom rezonansowym wzbudzającym się w rezonatorze Fabry'ego-Pérota. Zakłada się, że kontakty doprowadzające prąd są

dostatecznie odległe od złącza, by nie trzeba było uwzględniać niejednorodności rozpływu prądu w obszarze czynnym. Podobnie jak w przypadku płyty dielektrycznej [49] istnieją dwa rodzaje modów rezonansowych — mody TE (poprzeczne elektryczne) z wektorem **E** spolaryzowanym w płaszczyźnie poziomej oraz mody TM (poprzeczne magnetyczne) spolaryzowane w płaszczyźnie pionowej. Jedynymi niezerowymi składowymi pól **E** i **H** dla modów TE są E_y, H_z, H_x, zaś dla modów TM H_y, E_z, E_x. Kształtem poszczególnych składowych w różnych warstwach zajmiemy się poniżej.

Z równań Maxwella wynika, że pola **E**, **H** spełniają równanie falowe, na przykład dla pola **E**:

$$\Delta \mathbf{E}_k - \frac{\varepsilon_k}{c^2} \frac{\partial^2 \mathbf{E}_k}{\partial t^2} = 0, \qquad k = 1, 2, 3, \tag{2.7}$$

gdzie k podobnie jak we wzorach (2.1 ÷ 3) numeruje obszary.

Ponieważ szerokość diody W w kierunku y jest dużo większa od grubości obszaru czynnego T, na ogół pomija się zależność pól **E**, **H** od współrzędnej y. Jedynie K. Unger [50] rozważał efekt falowodowy związany z powolną zmianą w przestrzeni w kierunku osi y współczynników załamania i absorpcji. Przyczyną zmienności przenikalności dielektrycznej wzdłuż osi y mogą być np. niejednorodności domieszkowania złącza. Uwzględniając ściśle tę zależność powinniśmy otrzymać obserwowane doświadczalnie zjawisko „włóknistej" struktury promieniowania lasera złączowego [8, 51 ÷ 55], tzn. zawężenie obszaru aktywnego do wąskich włókien o przekroju rzędu 4 μ^2, z grubsza równoległych do płaszczyzny złącza. Jest to jednak bardzo trudne do wykonania ze względu na brak powtarzalności w położeniu i ilości włókien nawet w tej samej próbce [56]. Dlatego też istniejące jak dotąd modele opisują tylko promieniowanie laserów świecących całą powierzchnią złącza, a nie oderwanymi plamkami (porównaj część 3 i część 5). Ograniczamy się zatem do modelu, w którym pola **E**, **H** są wzdłuż osi y stałe.

Jeśli pomijamy zależność $\mathbf{E}(y)$ i $\mathbf{H}(y)$, równanie (2.7) redukuje się do

$$\frac{\partial^2 \mathbf{E}_k}{\partial x^2} + \frac{\partial^2 \mathbf{E}_k}{\partial z^2} - \frac{\varepsilon_k}{c^2} \frac{\partial^2 \mathbf{E}_k}{\partial t^2} = 0 \tag{2.8}$$

i analogicznie dla pola **H**.

Zakładamy, że fala zmienia się z czasem jak $e^{i\omega t}$, zaś wzdłuż osi z jak e^{iqz}, gdzie q — zespolona stała propagacji, tzn.

$$\mathbf{E}_k(x, y, z, t) = \mathbf{E}_k(x) e^{i(qz+\omega t)}, \tag{2.9}$$

$$\mathbf{H}_k(x, y, z, t) = \mathbf{H}_k(x) e^{i(qz+\omega t)}. \tag{2.10}$$

Jeśli przyjmiemy postać którejkolwiek składowej pola $\mathbf{E}_k(x)$ lub $\mathbf{H}_k(x)$ we wszystkich trzech warstwach, możemy z równań Maxwella wyznaczyć pozostałe. W tablicy 1 podane są wszystkie nieznikające składowe dla modów TE i TM; dla modów TE przyjęty został kształt składowej E_y, zaś dla modów TM — składowej H_y. Stałe A, B, Φ, p_k ($k = 1, 2, 3$) są zespolone ($p_k = p_k' + ip_k''$). Podstawiając te składowe pól do równania falowego dostajemy trzy równania

$$p_k^2 + q^2 - \frac{\varepsilon_k \omega^2}{c^2} = 0, \qquad (k = 1, 2, 3). \tag{2.11}$$

Tablica 1

Składowe pól E, H w trzech obszarach dla modów TE i TM

		1	2	3
TE	E_y	Ae^{ip_1x}	$\cos(p_2x+\Phi)$	Be^{-ip_3x}
	H_x	$\frac{qc}{\omega} Ae^{ip_1x}$	$\frac{qc}{\omega}\cos(p_2x+\Phi)$	$\frac{qc}{\omega} Be^{-ip_3x}$
	H_z	$\frac{-p_1c}{\omega} Ae^{ip_1x}$	$\frac{-ip_2c}{\omega}\sin(p_2x+\Phi)$	$\frac{p_3c}{\omega} Be^{-ip_3x}$
TM	H_y	Ae^{ip_1x}	$\cos(p_2x+\Phi)$	Be^{-ip_3x}
	E_x	$-\frac{qc}{\omega\varepsilon_1} Ae^{ip_1x}$	$-\frac{qc}{\omega\varepsilon_2}\cos(p_2x+\Phi)$	$-\frac{qc}{\omega\varepsilon_3} Be^{-ip_3x}$
	E_z	$\frac{p_1c}{\omega\varepsilon_1} Ae^{ip_1x}$	$\frac{ip_2c}{\omega\varepsilon_2}\sin(p_2x+\Phi)$	$\frac{-p_3c}{\omega\varepsilon_3} Be^{-ip_3x}$

2.2.1. *Warunki brzegowe*

Nakładamy warunki graniczne ciągłości składowych poprzecznych pola elektrycznego i magnetycznego na powierzchniach granicznych ($x = \pm T/2$). Dla fal TE oznacza to

$$E_{1y}\left(-\frac{T}{2}\right) = E_{2y}\left(-\frac{T}{2}\right), \tag{2.12a}$$

$$E_{2y}\left(\frac{T}{2}\right) = E_{3y}\left(\frac{T}{2}\right), \tag{2.12b}$$

$$H_{1z}\left(-\frac{T}{2}\right) = H_{2z}\left(-\frac{T}{2}\right), \tag{2.12c}$$

$$H_{2z}\left(\frac{T}{2}\right) = H_{3z}\left(\frac{T}{2}\right), \tag{2.12d}$$

co prowadzi do równania

$$\operatorname{tg}(p_2T) = \frac{ip_2(p_1+p_3)}{p_2^2+p_1p_3}. \tag{2.13}$$

W analogiczny sposób dla modów TM dostajemy związek

$$\operatorname{tg}(p_2T) = i\frac{p_2}{\varepsilon_2}\frac{p_1/\varepsilon_1+p_3/\varepsilon_3}{(p_2/\varepsilon_2)^2+p_1p_3/\varepsilon_1\varepsilon_3}. \tag{2.14}$$

Równania te określają dozwolone mody poprzeczne w rezonatorze. Możemy je charakteryzować podaniem liczby całkowitej N, określonej jako wielokrotność 2π, która musi być dodana do p_2T w rozwiązaniu powyższych równań. Liczba N związana jest z ilością maksimów w rozkładzie pola **E** w funkcji x. Tak np. mod z $N = 0$ ma jedno maksimum

w rozkładzie pola elektrycznego wzdłuż osi x, mod z $N = 1$ — dwa maksima itd. [17]. W praktyce wystarcza zwykle rozważać mody poprzeczne najniższych rzędów (tzn. o małych wartościach N), ponieważ w modelu trójwarstwowym mają one najniższe wartości prądu progowego.

Otrzymujemy zatem zarówno dla przypadku TE jak i TM cztery równania (2.11) i (2.13 lub 14) pozwalające wyznaczyć p_i ($i = 1, 2, 3$) oraz q w funkcji pozostałych parametrów. Początkowo [12÷15] rozwiązywano te równania rozwijając tg $(p_2 T)$ w szereg potęgowy i zakładając, że $|p_2 T| \ll 1$. Jednakże Anderson [35] pokazał rozwiązując te równania numerycznie, że odrzucenie dalszych wyrazów rozwinięcia jest w laserach z GaAs nieuzasadnione, szczególnie w przypadku, gdy zależność przenikalności dielektrycznej od położenia w obszarze czynnym nie jest symetryczna. Tak więc chcąc wyznaczyć stałe propagacji modów rezonansowych należy posługiwać się metodami numerycznymi.

2.3. Warunek wzbudzenia drgań

Musimy też nałożyć warunek graniczny na wystąpienie oscylacji w kierunku osi z. Energia przenoszona przez falę dana jest przez wektor Poyntinga, jest więc proporcjonalna do e^{2iqz}. Chcemy, aby fala po przejściu przez wnękę i odbiciu się od ścianki nie ulegała osłabieniu, zatem po wykonaniu pełnego cyklu.

$$R_1 e^{2iqL} R_2 e^{2iqL} \geqslant 1, \tag{2.15}$$

gdzie R_1, R_2 są współczynnikami odbicia dla obu ścianek tworzących rezonator Fabry'ego-Pérota; w przybliżeniu $R_1 \simeq R_2 \simeq \left(\frac{n_2 - 1}{n_2 + 1}\right)^2$. Gdy ograniczymy się do znaku równości, dostaniemy warunek progowy na wzmacnianie promieniowania

$$\sqrt{R_1 R_2}\, e^{2iq'L} e^{-2q''L} = 1, \tag{2.16}$$

gdzie

$$q = q' + iq''. \tag{2.17}$$

Biorąc część rzeczywistą i urojoną tego równania otrzymujemy

$$q' = \frac{M\pi}{L}, \tag{2.18}$$

$$q'' = -\frac{1}{2L} \ln\left(\frac{1}{\sqrt{R_1 R_2}}\right). \tag{2.19}$$

Pierwsze z tych równań podaje dopuszczalne wartości wektora falowego $\left(q' = \frac{2\pi}{\lambda} n_2\right)$. Gdy $|p_2| \ll |q|$, można to w przybliżeniu zapisać jako warunek na energię fotonu

$$E \simeq \frac{q' hc}{2\pi n_2} = \frac{Mhc}{2n_2 L}. \tag{2.20}$$

Natomiast drugie równanie daje warunek progowy na akcję laserową. Często warunek ten zapisuje się też w postaci

$$-\alpha_2 = \alpha + \frac{1}{L}\ln\left(\frac{1}{\sqrt{R_1 R_2}}\right), \tag{2.21}$$

bowiem $q'' = \frac{\alpha+\alpha_2}{2}$, gdzie α_2 — ujemny współczynnik absorpcji (tzn. współczynnik wzmocnienia) w obszarze 2, α — efektywny człon absorpcyjny opisujący straty dyfrakcyjne, absorpcję na wolnych nośnikach i inne możliwe mechanizmy powodujące straty [4].

2.4. Wzmocnienie promieniowania

Uśrednioną w czasie moc rozpraszanego promieniowania na jednostkę długości i szerokości w trzech obszarach można zapisać następująco [35]:

$$P_k = \frac{\omega\varepsilon_k''}{8\pi}\int_{a_k}^{b_k} |\mathbf{E}_k|^2\,dx, \qquad (k = 1, 2, 3), \tag{2.22}$$

gdzie $a_1 = -\infty$, $b_1 = a_2 = -T/2$, $b_2 = a_3 = T/2$, $b_3 = \infty$.

Podobnie znajdujemy uśrednioną w czasie energię nagromadzoną w trzech obszarach na jednostkę długości i szerokości:

$$U_k = \frac{\varepsilon_k'}{8\pi}\int_{a_k}^{b_k} |\mathbf{E}_k|^2\,dx, \qquad (k = 1, 2, 3), \tag{2.23}$$

a_k, b_k nadal oznaczają to samo co w (2.22).

Ze związku

$$\sqrt{\varepsilon_k' + i\varepsilon_k''} = n_k + i\frac{c}{2\omega}\alpha_k \tag{2.24}$$

otrzymujemy przybliżony wzór

$$\alpha_k \simeq \frac{\omega\varepsilon_k''}{c\sqrt{\varepsilon_k'}} \tag{2.25}$$

$\left(\text{po wzięciu pierwszego wyrazu rozwinięcia pierwiastka w szereg względem potęg } \left(\frac{\varepsilon_k''}{\varepsilon_k'}\right)^2\right)$.

Zatem stosując przybliżenie $\varepsilon_1' \simeq \varepsilon_2' \simeq \varepsilon_3' \simeq \varepsilon'$ dostajemy proste wyrażenie opisujące wzmocnienie lub osłabienie promieniowania rozchodzącego się w diodzie laserowej:

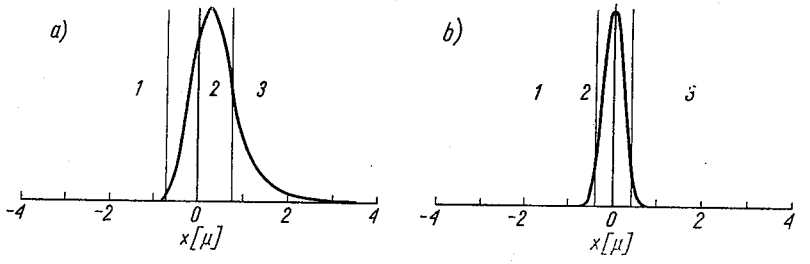
$$\alpha = \frac{P_1 + P_2 + P_3}{U_1 + U_2 + U_3}\frac{\sqrt{\varepsilon'}}{c}. \tag{2.26}$$

Możemy w ten sposób wyznaczyć próg ($\alpha = 0$), powyżej którego występuje akcja laserowa.

Uwzględniając znaną z doświadczenia zależność współczynnika załamania we wszystkich warstwach [48, 57] oraz współczynnika absorpcji w warstwach biernych [58 ÷ 62] od temperatury, koncentracji domieszek, składu stopu AlAs-GaAs (w przypadku laserów heterozłączowych) i energii fotonu można, mając numeryczne rozwiązania równań Maxwella, badać zależność właściwości lasera od takich parametrów, jak grubość obszaru czynnego T, długość rezonatora L, domieszkowanie we wszystkich warstwach czy skład stopu. Pozwala to ocenić, jak należy dobrać wartości tych parametrów, aby otrzymać jak najniższy prąd progowy.

2.5. Pole w strefie bliskiej

Po numerycznym obliczeniu stałych propagacji modów rezonansowych możemy bardzo łatwo otrzymać rozkład natężenia pola elektromagnetycznego w strefie bliskiej (Fresnela). W tym celu wystarczy położyć $z = 0$ i pole zależy już tylko od współrzędnej x,



Rys. 3. Pole w strefie bliskiej w płaszczyźnie pionowej dla: a) lasera monoheterozłączowego z warstwą $p^+-Al_{0.5}Ga_{0.5}As$, $T = 1.4\,\mu$, $\alpha' = 10\ cm^{-1}$; b) lasera biheterozłączowego z warstwami n- oraz $p^+-Al_{0.5}Ga_{0.5}As$, $T = 0.8\,\mu$, $\alpha' = 2\ cm^{-1}$ [17]

w sposób przedstawiony w tablicy 1. Na rysunku 3 można zobaczyć, jak wygląda typowy rozkład pola w strefie Fresnela dla modu podstawowego w laserze z pojedynczym (rys. 3a) i podwójnym heterozłączem (rys. 3b). Wyraźnie widać, jak heterozłącze poprawia efekt falowodowy, zmniejszając przenikanie pola do obszarów biernych.

2.6. Pole w strefie dalekiej

Znając rozwiązanie równań Maxwella wewnątrz diody laserowej możemy otrzymać pole promieniowania w strefie dalekiej, biorąc transformatę Fouriera pola E_y dla modów TE lub H_y dla modów TM na powierzchni zwierciadła ($z = 0$) [31]. Jeśli przez γ oznaczymy

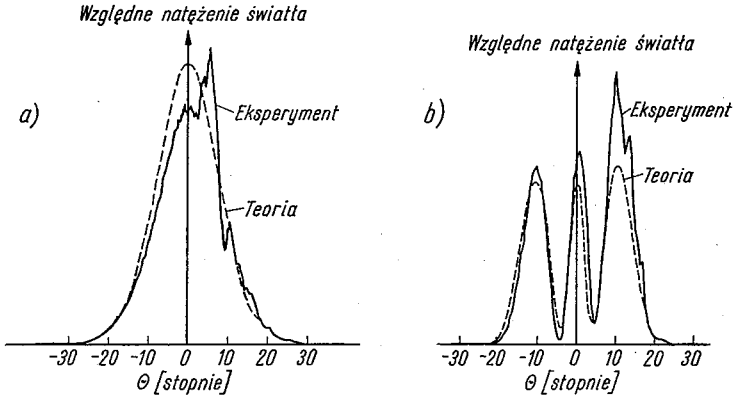
$$\gamma = \frac{\omega}{c}\sin\Theta, \tag{2.27}$$

gdzie Θ — kąt w płaszczyźnie xz pomiędzy kierunkiem obserwacji a normalną do powierzchni promieniującej, to obserwowany rozkład natężenia w strefie dalekiej dany jest przez $|f(\gamma)|^2$, gdzie np. dla modów TE

$$f(\gamma) = \int_{-\infty}^{\infty} e^{i\gamma x} E_y(x)\,dx = \frac{iBe^{i(\gamma-p_3)T/2}}{\gamma-p_3} - \frac{iAe^{-i(\gamma+p_1)T/2}}{\gamma+p_1} + \frac{e^{i\Phi}}{\gamma+p_2}\sin\left[\frac{(\gamma+p_2)T}{2}\right] + \frac{e^{-i\Phi}}{\gamma-p_2}\sin\left[\frac{(\gamma-p_2)T}{2}\right]. \quad (2.28)$$

Pole elektryczne modów wyższych rzędów ma coraz więcej maksimów wzdłuż kierunku osi x. Wynika stąd, że szerokość kątowa wiązki w płaszczyźnie prostopadłej do złącza rośnie ze wzrostem rzędu modu.

Ze względu na to, że grubość obszaru czynnego T jest dużo mniejsza od szerokości W diody, szerokość kątowa wiązki laserowej w płaszczyźnie prostopadłej do powierzchni



Rys. 4. Pole w strefie dalekiej w płaszczyźnie pionowej dla lasera monoheterozłączowego pracującego na: a) modzie podstawowym ($T = 2.5\ \mu$), b) modzie trzeciego rzędu ($T = 5\ \mu$) [16]

złącza jest większa, niż w płaszczyźnie równoległej do złącza [29]. Przykładowo, dla diod dyfuzyjnych w temperaturze 77°K wynoszą one odpowiednio $15 \div 25°$ i 3° [31, 54, 63, 64]. Rozkłady kątowe pola promieniowania lasera homozłączowego w płaszczyźnie prostopadłej do złącza sa ną ogół niesymetryczne — prócz silnego maksimum wokół osi rezonatora występuje zwykle dodatkowe maksimum po stronie n złącza [54]. Efekt ten otrzymuje się też w modelu trójwarstwowym; jest on rezultatem różnicy współczynników absorpcji po stronie n i p złącza. Występowanie dodatkowego maksimum po stronie n świadczy o niezbyt dobrym ograniczeniu pola promieniowania w obszarze czynnym, ponieważ przenikanie fali na stronę n jest dosyć znaczne [31].

Na rysunkach 4a, b przykładowo porównane są eksperymentalne i obliczone z modelu trójwarstwowego rozkłady kątowe w płaszczyźnie równoległej do złącza promieniowania lasera monoheterozłączowego (AlGa)As, pracującego odpowiednio na modzie podstawowym (rys. 4a) i modzie trzeciego rzędu (rys. 4b) [16]. Jak widać, kształt i położenie rozkładów teoretycznych jest w dosyć dobrej zgodności z wynikami pomiarów. Model trójwarstwowy daje zupełnie dobre wartości szerokości kątowych wiązek laserowych.

Jednakże obserwowane rozkłady pola wykazują bardziej asymetryczną i złożoną strukturę niż przewidziana w prostym modelu trójwarstwowym. Ta nakładająca się na zasadniczą dodatkowa struktura zmienia się przypadkowo od diody do diody, nie można więc nawet oczekiwać, by tak prosty model mógł oddać te szczegóły.

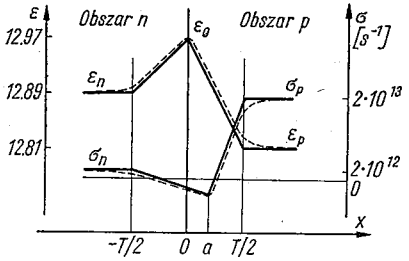
Model trójwarstwowy był też stosowany do opisu pola promieniowania laserów biheterozłączowych [39], nie oddaje on jednak poprawnie kolejności wzbudzania się poszczególnych modów. Do problemu tego powrócimy szerzej w części czwartej.

3. POLARYZACJA

Modele, w których zespolona przenikalność dielektryczna w obszarze czynnym (tzn. ε_2 ze wzoru (2.1)) jest stała w kierunku prostopadłym do złącza [13, 30, 34, 35, 65], dają jednakowy próg wzbudzenia dla modów TE i TM w laserze homozłączowym. Mogłoby się wydawać, że wynik ten jest potwierdzony przez pomiary polaryzacji diod laserowych [66÷70], w których przy progu stwierdzono występowanie obu polaryzacji TE i TM w mniej więcej jednakowym stopniu. Należy jednak pamiętać, że większość diod laserowych wysyła promieniowanie nie całą powierzchnią złącza, lecz pojedynczymi plamkami [8, 51÷55]. Oznacza to, że promieniowanie ma strukturę włóknistą, tzn. w płaszczyźnie prostopadłej do złącza wzbudzone mody mają strukturę dwuwymiarową (pola **E**, **H** są funkcjami nie tylko współrzędnej x, ale także y). Tymczasem w modelach ze stałą zespoloną przenikalnością dielektryczną w obszarze czynnym mamy do czynienia z modami jednowymiarowymi (pola **E**, **H** nie zależą od współrzędnej y). Dlatego też nie należy oczekiwać zgodności takich modeli z wynikami doświadczeń, w których diody laserowe promieniują w postaci plamek.

3.1. Lasery homozłączowe

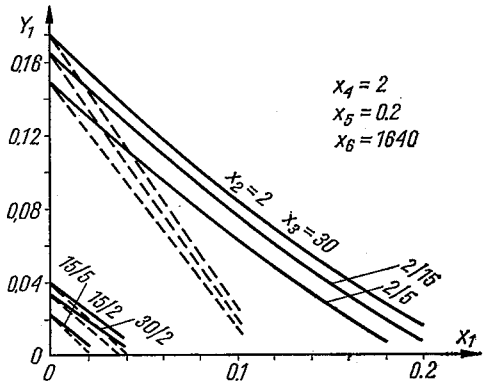
Okazuje się, że w laserach homozłączowych promieniujących jednorodnie na całej powierzchni złącza [55], np. wykonanych z materiału wolnego od dyslokacji, wzbudzane są tylko mody TM. Wynik taki można otrzymać w modelu zaproponowanym przez J. Hatza i E. Mohna, w którym zakłada się trójkątny rozkład zespolonej przenikalności dielektrycznej w obszarze czynnym, przedstawiony na rys. 5 [44]. Maksimum części rze-



Rys. 5. Linią przerywaną zaznaczony jest faktyczny przebieg części rzeczywistej przenikalności dielektrycznej $\varepsilon'(x)$ oraz przewodnictwa $\sigma(x)$ w kierunku prostopadłym do złącza dla lasera homozłączowego z GaAs. Linia ciągła przedstawia $\varepsilon'(x)$ i $\sigma(x)$ w modelu przyjętym przez J. Hatza i E. Mohna [44].

a — położenie minimum przewodnictwa (czyli maksimum inwersji)

czywistej przypada w środku złącza ($x = 0$), gdzie jest najmniejszy ujemny wkład od nośników swobodnych (obecność nośników swobodnych obniża wartość współczynnika załamania). Położenie minimum części urojonej przenikalności dielektrycznej ($x = a$) jest określone przez mechanizm wstrzykiwania nośników. Rekombinacja promienista zachodzi na ogół po stronie p złącza, ponieważ strona n jest silniej zdegenerowana oraz ruchliwość elektronów jest większa niż dziur, wobec tego minimum przewodnictwa przesunięte jest w stronę materiału typu p [51, 69].



Rys. 6. Zależność progu dla modów TE i TM od przesunięcia maksimum inwersji względem maksimum części rzeczywistej przenikalności dielektrycznej a. Krzywe ciągłe oznaczają mody TE, krzywe przerywane — mody TM;

$$X_1 = \frac{2a}{T}, \quad X_2 = \frac{\omega^2 T^2}{4c^2}(\varepsilon_0 - \varepsilon_p), \quad X_3 = \frac{\omega^2 T^2}{4c^2}(\varepsilon_0 - \varepsilon_n), \quad X_4 = \frac{\omega T^2}{4c^2}\sigma_p, \quad X_5 = \frac{\omega T^2}{4c^2}\sigma_n,$$

$X_6 = \frac{\omega^2 T^2}{4c^2}\varepsilon_0$; Y_1 opisuje inwersję przy progu $Y_1 = \frac{-\omega T^2}{4c^2}\sigma_0$;

$\varepsilon_0, \varepsilon_p, \varepsilon_n, \sigma_p, \sigma_n, a$ zaznaczone są na rysunku 6 [44]

Założenie, że maksimum inwersji obsadzeń pokrywa się przestrzennie z maksimum części rzeczywistej przenikalności dielektrycznej prowadzi do wyników bardzo podobnych do otrzymywanych w modelu, w którym przenikalność dielektryczna jest stała w obszarze złącza. Dopiero uwzględnienie przesunięcia tych maksimów ($a \neq 0$) daje obniżenie progu dla modów TM w porównaniu z TE i to nawet o rząd wielkości (rys. 6).

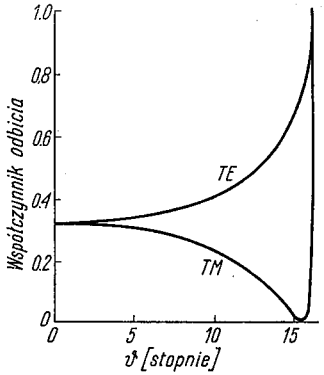
3.2. Lasery biheterozłączowe

Podobnie jak w laserze homozłączowym, również w laserze monoheterozłączowym zaobserwowano dużą przewagę polaryzacji TM [16]. Istotna różnica pojawia się dopiero w laserach biheterozłączowych (DH) [32, 70], z szeroką wnęką rezonansową (LOC) [71] i w laserach z podwójną heterostrukturą o geometrii paskowej [72]. We wszystkich tych typach laserów stwierdzono wyłącznie polaryzację TE.

Niższy próg dla modów TE niż dla TM otrzymali M. J. Adams i M. Cross [73] w modelu trójwarstwowym, zakładając, że efekt ten jest rezultatem nieprostopadłego padania

fali na zwierciadła. Jak już wiemy, w modelu tym pole w obszarze czynnym zmienia się jak $e^{iqz}\cos(p_2 x+\Phi)e^{i\omega t}$. Oznacza to, że fala rozchodząca się we wnęce jest superpozycją dwóch fal płaskich padających na zwierciadła pod kątami

$$\vartheta = \pm\operatorname{arctg}\left(\frac{\operatorname{Re} p_2}{\operatorname{Re} q}\right). \tag{3.1}$$



Rys. 7. Współczynnik odbicia od powierzchni granicznej GaAs — powietrze dla modów TE i TM w funkcji kąta padania [73]

Stosując prawo załamania można znaleźć kąt załamania ϑ', a potem współczynniki odbicia dla modów TE i TM [58]

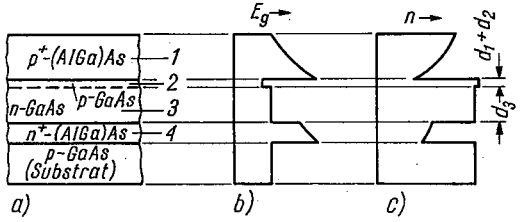
$$R_{\mathrm{TE}} = \frac{\sin^2(\vartheta'-\vartheta)}{\sin^2(\vartheta'+\vartheta)}, \tag{3.2}$$

$$R_{\mathrm{TM}} = \frac{\operatorname{tg}^2(\vartheta'-\vartheta)}{\operatorname{tg}^2(\vartheta'+\vartheta)}. \tag{3.3}$$

Wykresy R_{TE} i R_{TM} jako funkcji ϑ dla powierzchni granicznej GaAs — powietrze są podane na rys. 7. Fakt, że $R_{\mathrm{TE}} \geqslant R_{\mathrm{TM}}$ tłumaczy, dlaczego mody TE są uprzywilejowane

4. LASERY Z SZEROKĄ WNĘKĄ REZONANSOWĄ, SELEKCJA MODÓW

Cechą charakterystyczną w działaniu lasera złączowego jest to, że jeśli struktura lasera dopuszcza wzbudzanie się modów wyższych rzędów (np. jeśli wnęka optyczna jest dostatecznie szeroka, co można zrealizować w laserach biheterozłączowych (DH) [74—76] i w laserach z szeroką wnęką rezonansową (LOC) (rys. 8) [71]), wówczas na ogół faworyzowany jest mod najwyższego rzędu spośród dozwolonych [16, 18]. Wraz ze wzrostem rzędu modu pole elektromagnetyczne przenika coraz głębiej w sąsiadujące z wnęką warstwy pochłaniające, rosną więc straty absorpcyjne. Można by wobec tego oczekiwać, że zawsze mod najniższego rzędu (podstawowy) powinien mieć najwyższy współczynnik wzmocnienia, potem powinien się wzbudzać mod drugiego rzędu itd. Skoro jednak obserwuje się wcześniejsze wzbudzenie modów wyższych rzędów, oznacza to, że współczynnik wzmocnienia nie może być rozłożony jednorodnie we wnęce rezonansowej.

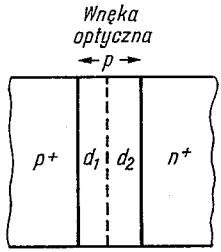


Rys. 8. Schemat struktury lasera LOC
a) przekrój diody w płaszczyźnie pionowej, b) zależność przerwy energetycznej E_g od współrzędnej x, c) zależność współczynnika załamania od współrzędnej x

Opisany wcześniej model trójwarstwowy lasera złączowego ze stałą przenikalnością dielektryczną w obszarze czynnym daje stosunkowo dobry opis pola promieniowania lasera w strefie bliskiej i dalekiej dla laserów homozłączowych i monoheterozłączowych. W modelu tym jednak zawsze mod podstawowy ma najniższy próg wzbudzenia [16]. Chcąc więc otrzymać zgodną z doświadczeniem selekcję modów należy przyjąć model bardziej skomplikowany.

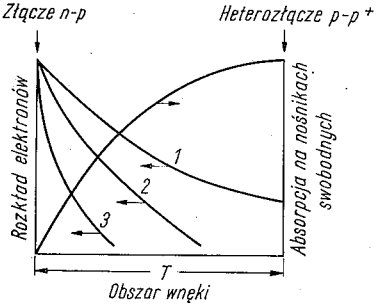
4.1. Model Butlera

Uprzywilejowanie modów wyższego rzędu otrzymał J. K. Butler [19] zakładając, że współczynnik wzmocnienia nie jest stały we wnęce rezonansowej. W modelu przyjętym przez Butlera obszar skompensowanego GaAs typu p dzieli się na dwie warstwy różniące się wartością współczynnika wzmocnienia. Warstwa sąsiadująca z p^+—$Al_xGa_{1-x}As$ ma niższy współczynnik wzmocnienia (może być nawet warstwą pochłaniającą) niż pozostała część wnęki, zawierająca materiał czynny optycznie (rys. 9). Grubość warstwy czynnej



Rys. 9. Przyjęty przez J. K. Butlera model lasera DH z wnęką optyczną podzieloną na część o dużym współczynniku wzmocnienia i grubości d_2 oraz część o małym (lub nawet ujemnym) współczynniku wzmocnienia i grubości d_1 [18]

jest bardzo mała (poniżej 1 μ). Można próbować uzasadniać takie postępowanie dwoma czynnikami: 1) inwersja obsadzeń pasm w rezonatorze może nie być przestrzennie jednorodna; 2) przy jednorodnej inwersji obsadzeń może być przestrzennie niejednorodny rozkład strat na nośnikach swobodnych. Obie te sytuacje przedstawione są na rys. 10. Jeśli inwersja obsadzeń w warstwie skompensowanego materiału typu p nie jest jednorodna,



Rys. 10. Z lewej — rozkład elektronów w obszarze wnęki optycznej o grubości T; z prawej — absorpcja na nośnikach swobodnych w obszarze wnęki dla złącz dyfuzyjnych GaAs: Zn [19]

rozkład przestrzenny elektronów może być taki, jak na rysunku. Poszczególne krzywe odpowiadają różnym drogom dyfuzji elektronów w silnie domieszkowanym GaAs typu p (od 0.6 μ do 0.1 μ). Pomiary drogi dyfuzji były wykonywane w materiałach w słabym polu optycznym (w porównaniu z laserowym), można się więc spodziewać, że droga dyfuzji elektronów w polu promieniowania laserowego o dużym natężeniu jest mniejsza, ponieważ czas życia pary elektron-dziura jest krótszy.

Gdy grubość skompensowanego obszaru typu p jest mniejsza od drogi dyfuzji elektronu, współczynnik wzmocnienia g wynikający z inwersji obsadzeń jest prawie stały. Ujemny współczynnik absorpcji związany z przejściami promienistymi wynosi

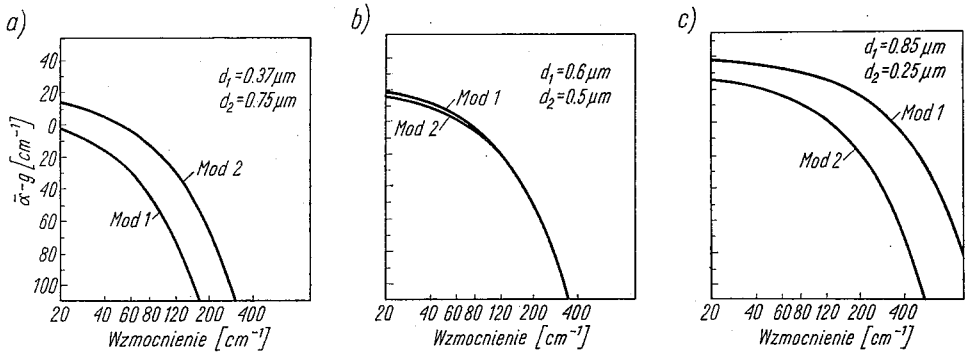
$$\alpha = -g + \alpha_{f.c.}, \tag{4.1}$$

gdzie $\alpha_{f.c.}$ (absorpcja na wolnych nośnikach) zależy od położenia jak na rysunku 10. Dla złączy dyfuzyjnych $\alpha_{f.c.}$ w pobliżu heterozłącza p^+-p jest dużo większe niż w pobliżu złącza $p-n$ (50 cm^{-1} w porównaniu z 5 cm^{-1} [19]), co powoduje, że współczynnik wzmocnienia netto jest większy w pobliżu złącza $p-n$.

Model zaproponowany przez Butlera różni się od modelu trójwarstwowego jedynie ilością warstw. Metoda rozwiązywania równań Maxwella jest więc tu analogiczna, jak opisana w części drugiej. Podobnie jak poprzednio zakłada się wykładniczy zanik pola promieniowania w obszarach ograniczających wnękę rezonansową, zaś zależność sinusoidalną od współrzędnej x wewnątrz wnęki. Ze względu na analogię postępowania nie będziemy tu rozważać szczegółowo równań Maxwella, lecz zajmiemy się od razu omówieniem wyników, do jakich prowadzi model Butlera.

4.2. Selekcja modów w laserze biheterozłączowym

Oznaczamy grubość warstwy czynnej w obszarze skompensowanego GaAs typu p przez d_2, zaś przez d_1 grubość pozostałej części wnęki, która pochłania promieniowanie. Rys. 11 przedstawia dla różnych wartości współczynnika wzmocnienia (tzn. ujemnego współczynnika absorpcji w ośrodku o grubości d_2 z jednorodną inwersją obsadzeń) stratę netto lub wzmocnienie netto we wnęce $\bar{\alpha}-g$ dla lasera biheterozłączowego: $\bar{\alpha}$ jest całkowitym współczynnikiem absorpcji uwzględniającym straty we wszystkich obszarach oprócz



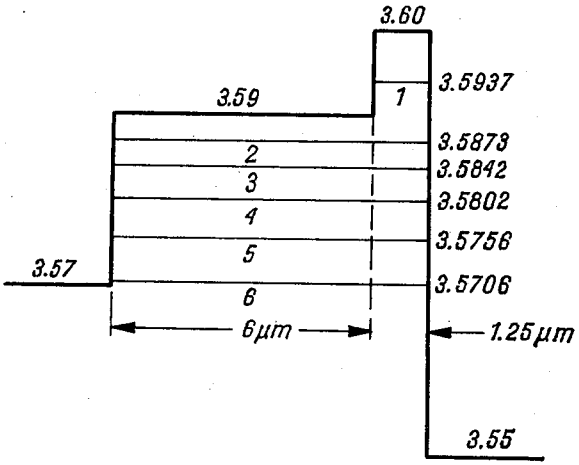
Rys. 11. Zależność pomiędzy współczynnikiem wzmocnienia w warstwie o grubości d_2 a wzmocnieniem lub stratami netto dla modu podstawowego (1) i drugiego rzędu (2) otrzymana w modelu J. K. Butlera W warstwie o grubości d_1 przyjęto współczynnik absorpcji 50 cm $^{-1}$. Grubość wnęki optycznej $T = d_1 + d_2 = 1.1$ μ [18]

strat na zwierciadłach rezonatora, natomiast g jest współczynnikiem wzmocnienia rozważanego modu. Wzmocnienie w materiale d_2 zależy tylko od wartości przepływającego prądu, podczas gdy $\bar{\alpha}$ i g zależą od całej struktury i rozkładu pola rozchodzącego się modu. Akcja laserowa pojawia się, gdy wzmocnienie netto $-(\bar{\alpha}-g)$ jest równe stratom na zwierciadłach.

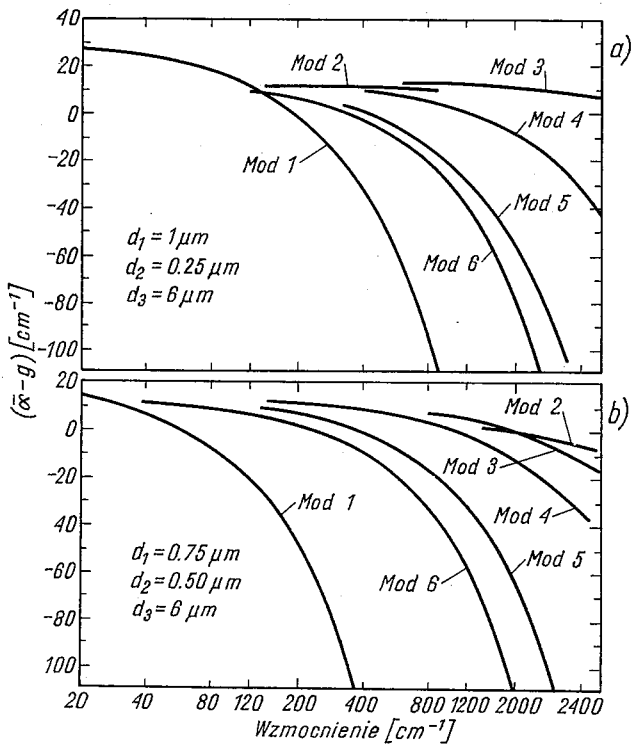
Z porównania rysunków 11 a÷c można stwierdzić, że w miarę zmniejszania się stosunku d_2/d_1, tzn. zwężania się obszaru czynnego, powiększa się próg na wzbudzenie modu podstawowego a obniża próg dla modu drugiego rzędu. Na ogół dla $d_2 < 0.5$ μ zawsze uprzywilejowane są mody wyższych rzędów.

4.3. Selekcja modów w laserze LOC

Przenormowując stałą propagacji q (2.9) jako $\frac{2\pi}{\lambda}(a'+ia'')$ można podać układ dozwolonych modów w rezonatorze. Dla modów rezonansowych $a'' = 0$, zaś a' ma interpretację efektywnego współczynnika załamania ośrodka dla rozchodzącej się w nim fali płaskiej. Pole elektromagnetyczne jest praktycznie ograniczone do obszaru rezonatora, jeśli a' jest mniejsze od współczynnika załamania we wnęce rezonansowej, lecz większe od współczynników załamania warstw sąsiednich. Rys. 12 przedstawia układ modów wzbudzających się w strukturze LOC. Mod podstawowy ma znormalizowaną stałą propagacji a' zawartą między wartościami współczynnika załamania n-GaAs i skompensowanego p-GaAs. Oznacza to, że energia pola promieniowania skoncentrowana jest głównie w wąskim obszarze skompensowanego p-GaAs, zatem charakterystyki lasera LOC pracującego na modzie podstawowym (takie jak rozkłady pola w strefie bliskiej i dalekiej, prąd progowy, wydajność kwantowa) powinny być podobne jak dla lasera monoheterozłączowego. Natomiast, gdy jest wzbudzony jeden z modów wyższego rzędu, pole w strefie bliskiej rozciąga się na całej szerokości wnęki (od heterozłącza p^+-p do heterozłącza n^+-n), jest więc podobne do pola promieniowania lasera biheterozłączowego o bardzo szerokiej wnęce. Jeśli warstwa skompensowanego GaAs typu p jest zbyt cienka, mogą



Rys. 12. Zależność współczynnika załamania od położenia w kierunku prostopadłym do złącza dla lasera LOC z zaznaczonymi położeniami modów rezonansowych (wartościami efektywnego współczynnika załamania a') [19]



Rys. 13. Krzywe wzmocnienia (analogiczne jak na rys. 11) dla poszczególnych modów rezonansowych w laserze LOC otrzymane w modelu J. K. Butlera

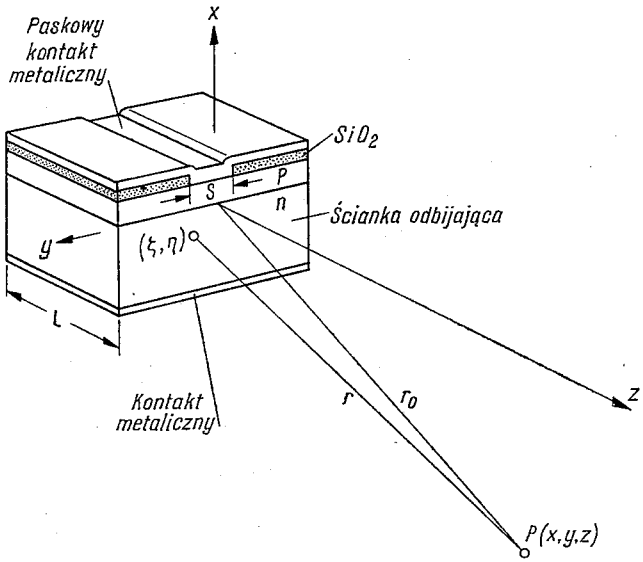
d_3 — odległość heterozłącza $n^+ - n$ od złącza $p—n$; obszar pomiędzy złączem $p—n$ a heterozłączem $p—p^+$ jest podzielony na dwie warstwy o grubościach d_1, d_2 i własnościach analogicznych jak w modelu lasera DH (por. rys. 9). Współczynnik absorpcji w *n-GaAs* przyjęto 10 cm^{-1}. Zmiany stosunku d_1/d_2 przesuwają przede wszystkim mod podstawowy, natomiast zmiany współczynnika absorpcji w *n*-GaAs wpływają głównie na położenie modów wyższych rzędów [19]

nie istnieć w ogóle mody rezonansowe, Jest to sytuacja identyczna jak w laserach monoheterozłączowych, gdy złącze $p-n$ leży zbyt blisko heterozłącza.

Na rys. 13 pokazane są otrzymane numerycznie krzywe wzmocnienia dla poszczególnych modów w laserze LOC. Jak widać, mod podstawowy osiąga próg najwcześniej. Jeśli zniszczymy zwierciadło rezonatora w obszarze pomiędzy heterozłączem p^+-p a złączem $p-n$ (np. przez lokalne wytrawienie zwierciadeł), wówczas współczynnik odbicia dla skoncentrowanego w tym obszarze modu podstawowego będzie dużo mniejszy niż dla modów wyższych rzędów, równomiernie rozłożonych w całym rezonatorze. Jak widać z rys. 13 powinien się wtedy wzbudzać mod o rzędzie najwyższym z dozwolonych, ponieważ ma on najniższy prąd progowy spośród wszystkich modów wyższych rzędów. Przyjęty przez Butlera model oddaje zatem poprawnie obserwowaną eksperymentalnie selekcję modów w laserach DH i LOC.

5. LASERY PASKOWE

Założenie, że w płaszczyźnie równoległej do złącza ośrodek jest jednorodny, prowadzi w modelu trójwarstwowym do rozwiązań nie oddających włóknistej natury promieniowania laserowego (tzn. ograniczenia obszaru promieniującego do drobnych plamek w płaszczyźnie złącza) [8, 51 ÷ 55]. Bardziej szczegółowy model powinien więc uwzględniać ogniskowanie promieniowania wzdłuż płaszczyzny złącza. Jednakże dobór odpowiedniego modelu ośrodka skupiającego jest trudny ze względu na przypadkowość występowania włókien w typowych laserach półprzewodnikowych. Emisja promieniowania pojawia się zwykle w dwóch lub więcej obszarach, których położenia wzdłuż płaszczyzny złącza nie można przewidzieć. W każdym z tych obszarów mogą się jednocześnie wzbudzać mody



Rys. 14. Schemat lasera z geometrią paskową kontaktu metalicznego z zaznaczonym układem współrzędnych
Przednie zwierciadło lasera leży w płaszczyźnie $z = 0$; $P(x, y, z)$ — punkt obserwacji pola w strefie dalekiej, odległy o r od punktu (ξ, η) na przednim zwierciadle

poprzeczne. Zależnie od odległości sąsiednich włókien emitujących promieniowanie, mody z tych włókien mogą być ze sobą sprzężone lub nie [77, 78]. Różne mody, których wzbudzenie nie może być kontrolowane przez eksperymentatora, interferują ze sobą, dając obrazy pola w strefie bliskiej, dalekiej oraz widma częstości na ogół inne dla każdej diody. Ze względu na tę przypadkowość, konstrukcja modelu lasera, w którym wzbudzające się mody mogłyby być sprawdzone doświadczalnie, jest zadaniem bardzo trudnym. Model taki musiałby uwzględniać przypadkowe niedoskonałości kryształu półprzewodnikowego.

Trudności te można ominąć zakładając, że emisja promieniowania zachodzi tylko w jednym, pojedynczym włóknie. Sytuacja taka jest realizowana w laserach z „geometrią paskową" (rys. 14). Obszar aktywny jest w nich ograniczony do bardzo wąskiego paska o szerokości zbliżonej do szerokości kontaktu metalicznego.

Dyfuzyjne homozłączowe lasery paskowe generują na ogół promieniowanie o symetrii hermitowsko-gaussowskiej w płaszczyźnie poziomej [20, 79]. J. C. Dyment [20] stwierdził, że występowanie tego typu symetrii świadczy o przestrzennej zależności współczynnika załamania w płaszczyźnie poziomej, powodującej zachowanie się ośrodka laserowego podobne jak soczewki.

5.1. Model paraboliczny

Model lasera paskowego zawierający prócz zwykłej zależności przenikalności dielektrycznej od współrzędnej x (prostopadle do złącza) dodatkowo zależność od współrzędnej y (w płaszczyźnie złącza) rozpatrzyli szczegółowo T. H. Zachos i J. E. Ripper [21]. Założyli oni, że współczynnik załamania jest największy na osi rezonatora i maleje ze wzrostem $|x|$ lub $|y|$:

$$n(x, y) = \bar{n}\left[1 - \left(\frac{x}{x_0}\right)^2 - \left(\frac{y}{y_0}\right)^2\right]^{1/2}; \tag{5.1}$$

$\bar{n}$ jest maksymalną wartością współczynnika załamania; stałe x_0, y_0 różnią się ze względu na znacznie silniejsze ogniskowanie w kierunku prostopadłym do złącza w porównaniu z równoległym.

Rozwiązania równań Maxwella dla takiej postaci współczynnika załamania były dyskutowane np. przez E. A. J. Marcatiliego [80]. Mody wzbudzające się w rezonatorze typu Fabry'ego-Pérota można charakteryzować przez podanie trzech liczb całkowitych (M, N, Q). Liczby N, Q wskazują, przez ile zer przechodzi natężenie składowej poprzecznej pola odpowiednio w kierunkach x, y. Liczba M określa ilość połówek fali mieszczących się we wnęce wzdłuż osi z. Równania Maxwella są rozwiązywane w sposób przybliżony przez pominięcie wyrazów rzędu $N\lambda/\pi x_0$ i $Q\lambda/\pi y_0$ w porównaniu z jednością. Wówczas składowe E_x i H_y znikają, problem sprowadza się do rozwiązania skalarnego równania falowego na E_y oraz H_x. Rozwiązań poszukuje się w postaci

$$E^{\pm}_{yNQM}(x, y, z, t) = \mathscr{E}_{NQM} X_N(x) Y_Q(y) e^{\pm ik\gamma_{NQM} z} e^{-i\omega t}, \tag{5.2}$$

gdzie:

$\mathscr{E}_{NQM}$ jest stałą,

$k = \dfrac{\omega}{c}$,

γ_{NQM} jest wielkością bezwymiarową charakteryzującą wektor falowy modów rozchodzących się we wnęce wzdłuż osi z.

Postać funkcji $X_N(x)$, $Y_Q(y)$ oraz γ_{NQM} wyznacza się przez podstawienie E_y do równania falowego. Okazuje się, że $X_N(x)$ oraz $Y_Q(y)$ są funkcjami hermitowsko-gaussowskimi z różnymi czynnikami skalującymi, a więc wyniki są zgodne z rezultatami eksperymentalnymi [20, 79, 81]:

$$X_N(x) = H_N\left(\sqrt{2}\frac{x}{w_{ox}}\right)\exp\left(-\frac{x^2}{w_{ox}^2}\right), \tag{5.3}$$

$$Y_Q(y) = H_Q\left(\sqrt{2}\frac{y}{w_{oy}}\right)\exp\left(-\frac{y^2}{w_{oy}^2}\right), \tag{5.4}$$

$$\gamma_{NQM} = \bar{n}\left(1 - \frac{2N+1}{\bar{n}x_0 k} - \frac{2Q+1}{\bar{n}y_0 k}\right)^{1/2}. \tag{5.5}$$

$H_m(\cdot)$ oznacza wielomian Hermite'a rzędu m. Czynniki skalujące w_{ox}, w_{oy} mają interpretację szerokości wiązki w kierunkach x, y dla modów podłużnych $(0, 0, M)$ i wynoszą

$$w_{ox} = \sqrt{\frac{\lambda x_0}{\pi\bar{n}}}, \qquad w_{oy} = \sqrt{\frac{\lambda y_0}{\pi\bar{n}}}; \tag{5.6a,b}$$

γ_{NQM} powinno spełniać znany warunek

$$\gamma_{NQM} = \frac{M\pi}{kL}, \tag{5.7}$$

co pozwala wyznaczyć częstości rezonansowe wnęki laserowej

$$\nu_{NQM} = \frac{c}{4\pi\bar{n}}\left(\frac{2N+1}{x_0} + \frac{2Q+1}{y_0}\right) + \frac{cM}{2L\bar{n}}\left\{1 + \left[\frac{L}{2\pi M}\left(\frac{2N+1}{x_0} + \frac{2Q+1}{y_0}\right)\right]^2\right\}^{1/2} \tag{5.8}$$

Różniczkując to równanie można znaleźć wyrażenie na separację modów rezonansowych. Jest ono dosyć skomplikowane, lecz dla modów poprzecznych niskiego rzędu, tzn. takich, że

$$\frac{L}{2\pi M}\left(\frac{2N+1}{x_0} + \frac{2Q+1}{y_0}\right) \ll 1 \tag{5.9}$$

przybiera prostszą postać:

$$L\frac{\Delta\lambda}{\lambda^2} = -\frac{1}{2\bar{n}_e}\left(\frac{L}{\pi x_0}\Delta N + \frac{L}{\pi y_0}\Delta Q + \Delta M\right), \tag{5.10}$$

gdzie $\bar{n}_e$ jest „równoważnym" współczynnikiem załamania uwzględniającym obecność ośrodka dyspersyjnego [48]:

$$\bar{n}_e = \bar{n}\left(1 - \frac{\lambda}{\bar{n}}\frac{d\bar{n}}{d\lambda}\right). \tag{5.11}$$

Gdy $\Delta N = \Delta Q = 0$, dostajemy wzór na separację modów rezonansowych we wnęce Fabry'ego-Pérota [82]

$$\frac{\Delta\lambda}{\Delta M} \simeq \frac{-\lambda^2}{2\bar{n}_e L}. \tag{5.12}$$

Separacja modów z $\Delta Q = \Delta M = 0$ wynosi

$$\frac{\Delta\lambda}{\Delta N} \simeq \frac{-\lambda^2}{2\pi\bar{n}_e x_0} \tag{5.13}$$

i wreszcie gdy $\Delta N = \Delta M = 0$

$$\frac{\Delta\lambda}{\Delta Q} \simeq \frac{-\lambda^2}{2\pi\bar{n}_e y_0}. \tag{5.14}$$

Widać, że separacja modów (N, Q, M) i $(N, Q+1, M)$ zależy od własności skupiających ośrodka wzdłuż osi y, zaś modów (N, Q, M) i $(N+1, Q, M)$ — od własności w kierunku prostopadłym do złącza.

Pole w strefie dalekiej w półprzestrzeni $z > 0$ otrzymujemy biorąc podwójną transformatę Fouriera pola na ściance $z = 0$ [83]

$$E_{NQM}(x, y, z) = B_{NQM}\frac{z}{r}\frac{e^{ikr}}{r}H_N\left(\sqrt{2}\frac{x}{w_x}\right)H_Q\left(\sqrt{2}\frac{y}{w_y}\right)\exp\left\{-\left[\left(\frac{x}{w_x}\right)^2 + \left(\frac{y}{w_y}\right)^2\right]\right\}; \tag{5.15}$$

r oznacza odległość punktu obserwacji od początku układu ($kr \gg 1$), B_{NQM} jest stałą, zawierającą współczynnik transmisji zwierciadła. Czynniki skalujące w_x, w_y dane są wzorami

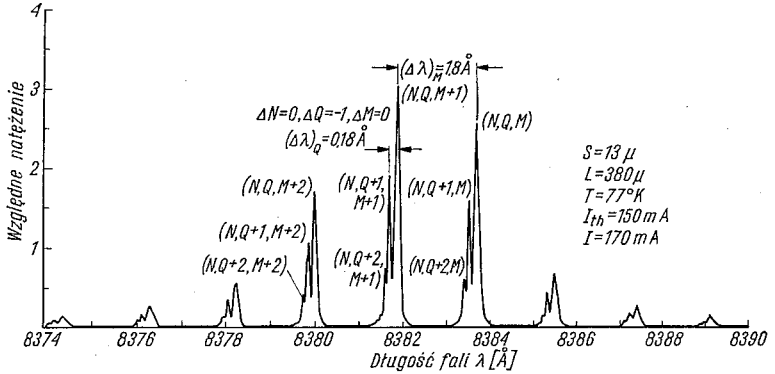
$$w_x = \frac{\lambda z}{\pi w_{ox}}, \tag{5.16a}$$

$$w_y = \frac{\lambda z}{\pi w_{oy}} \tag{5.16b}$$

są szerokościami wiązek modów podłużnych w kierunkach x, y w strefie dalekiej.

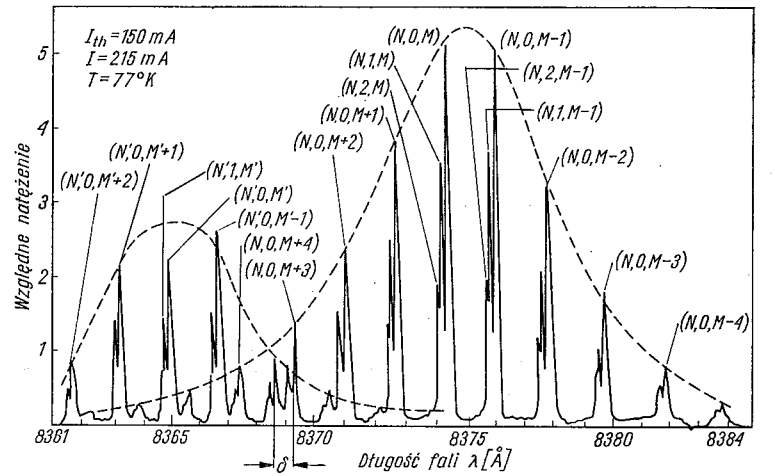
5.2. Widmo częstości promieniowania lasera paskowego

Opisany powyżej model lasera paskowego pozwolił dokładnie zbadać widmo częstości promieniowania laserowego. Obliczone z tego modelu częstości modów rezonansowych pokrywają się bardzo dobrze z otrzymanymi w wyniku pomiarów spektrometrycznych o wysokiej zdolności rozdzielczej. Typowe widma częstości dla pracującego w sposób ciągły lasera paskowego z GaAs przedstawione są na rysunkach 15 (dla prądu niewiele przewyższającego prąd progowy) i 16 (dla prądu znacznie przekraczającego wartość progową). Jak widać, obserwowane położenia maksimów pokrywają się z obliczonymi. Na rys. 16 występują dwie grupy modów. Wzajemne położenie kolejno wzbudzających się grup jest przypadkowe i zmienia się od diody do diody, ponieważ każda z grup koncentruje się wokół częstości, dla której średni współczynnik wzmocnienia przybiera maksimum.



Rys. 15. Typowe widmo częstości dla lasera paskowego z GaAs pracującego w sposób ciągły przy prądzie niewiele większym od progowego [21]

Widoczne są różne grupy modów osiowych. W każdej z grup liczby N, M pozostają stałe, natomiast liczba Q maleje ze wzrostem długości fali w obrębie danej grupy



Rys. 16. Typowe widmo częstości dla lasera paskowego z GaAs pracującego w sposób ciągły przy prądzie znacznie większym od progowego [22]

δ — zmieniająca się od diody do diody separacja sąsiednich modów z różnych grup

Współczynnik ten silnie zależy od częstości modu, wartości prądu oraz rozkładu domieszek w obszarze złącza, który jest inny dla każdej diody.

Ten sam model lasera paskowego dobrze opisuje promieniowanie heterozłączowego lasera paskowego w płaszczyźnie równoległej do złącza [72], co oznacza, że własności ogniskujące w obszarze czynnym w płaszczyźnie złącza nie zależą silnie od struktury lasera. Istotne różnice występują natomiast pomiędzy charakterystykami promieniowania lasera homo- i heterozłączowego w płaszczyźnie prostopadłej do złącza. Rozkład pola w tej płaszczyźnie nie wykazuje już symetrii hermitowsko-gaussowskiej, bardzo blisko progu wzbudzają się od razu mody wyższych rzędów, poszczególne grupy jednocześnie wzbudzonych modów wyższych rzędów położone są znacznie gęściej w widmie długości fali (odstęp pomiędzy sąsiednimi grupami wynosi poniżej 5 Å zamiast ok. 25 Å w lase-

rach dyfuzyjnych, gdzie ponadto mody różnych rzędów współistnieją dopiero daleko powyżej progu) [72]. Wszystkie te różnice oznaczają, że do opisu promieniowania heterozłączowego lasera paskowego w płaszczyźnie prostopadłej do złącza musi być przyjęty bardziej złożony model zależności przenikalności dielektrycznej od położenia niż prosta parabola (5.1) lub „studnia potencjału" (2.1).

6. ZAKOŃCZENIE

Zapoznaliśmy się z kilkoma modelami laserów półprzewodnikowych pozwalającymi uzyskać informacje o promieniowaniu laserowym. Na zakończenie należałoby dodać, że omówione modele są nadal ulepszane oraz że stale proponuje się nowe ich wersje. O aktualności zagadnień poruszonych w tym artykule może świadczyć fakt, że w trakcie jego opracowywania do druku ukazało się kilka prac z tej dziedziny. Dotyczą one zarówno laserów heterozłączowych [84, 87÷89], jak i laserów LOC [85] oraz paskowych [86].

LITERATURA CYTOWANA W TEKŚCIE

1. A. Halak, Heterozłączowe lasery półprzewodnikowe, Elektronika, **13**, zesz. 9 (1972), s. 360.
2. M. A. Herman Problemy fizyki złączowych laserów półprzewodnikowych, Postępy Fizyki (w przygotowaniu).
3. B. Mroziewicz, Lasery półprzewodnikowe, Nowa Technika, zeszyt 68, WNT, Warszawa 1967.
4. M. H. Pilkuhn, The Injection Laser, Phys. Status Solidi, 25 (1968), s. 9.
5. Gallium Arsenide Lasers, C. H. Gooch (ed.), Wiley-Interscience, London 1969.
6. D. F. Nelson, J. McKenna, Electromagnetic Modes of Anisotropic Dielectric Waveguides at p-n Junctions, J. Appl. Phys., 38 (1967), s. 38.
7. D. F. Nelson F. K. Reinhart, Light Modulation by the Electro-Optic Effect in Reverse-Biased GaP p-n Junctions, Appl. Phys. Letters, 5 (1964), s. 148.
8. R. N. Hall, G. E. Fenner, J. D. Kingsley, T. J. Soltys, R. O. Carlson, Coherent Light Emission from GaAs Junctions, Phys. Rev. Letters, 9 (1962), s. 366.
9. M. I. Nathan, W. P. Dumke, G. Burns, F. H. Dill Jr., G. J. Lasher, Stimulated Emission of Radiation from GaAs p-n Junctions, Appl. Phys. Letters, 1 (1962), s. 62.
10. T. M. Quist, R. H. Rediker, R. J. Keyes, W. E. Krag, B. Lax, A. L. McWhorter, Semiconductor Maser of GaAs, Appl. Phys. Letters, 1 (1962), s. 91.
11. W. L. Bond, B. G. Cohn, R. C. C. Leite, A. Yariv, Observation of the Dielectric-Waveguide Mode of Light Propagation in p-n Junctions, Appl. Phys. Letters, 2 (1963), s. 57.
12. G. J. Lasher, Threshold Relations and Diffraction Loss for Injection Lasers, IBM J. Res. Developm. (USA), 7 (1963), s. 58.
13. R. N. Hall, D. J. Olechna, Wave Propagation in an Active Dielectric Slab, J. Appl. Phys., 34 (1963), s. 2565.
14. A. L. McWhorter, H. J. Zeiger, B. Lax, Theory of Semiconductor Maser of GaAs, J. Appl. Phys., 34 (1963), s. 235.
15. A. Yariv, R. C. C. Leite, Dielectric-Waveguide Mode of Light Propagation in p-n Junctions, Appl. Phys. Letters, 2 (1963), s. 55.
16. N. E. Byer, J. K. Butler, Optical Field Distribution in Close-Confined Laser Structures, IEEE J. Quantum Electron., QE-6 (1970), s. 291.
17. M. J. Adams, M. Cross, Electromagnetic Theory of Heterostructure Injection Lasers, Solid State Electronics (GB), 14 (1971), s. 865.

18. J. K. Butler, H. S. Sommers Jr., H. Kressel, High Order Transverse Cavity Modes in Heterojunction Diode Lasers, Appl. Phys. Letters, 17 (1970), s. 403.
19. J. K. Butler, Theory of Transverse Cavity Mode Selection in Homo- and Heterojunction Semiconductor Diode Lasers, J. Appl. Phys., 42 (1971), s. 4447.
20. J. C. Dyment, Hermite-Gaussian Mode Patterns in GaAs Junction Lasers, Appl. Phys. Letters, 10 (1967), s. 84.
21. T. H. Zachos, J. E. Ripper, Resonant Modes of GaAs Junction Lasers, IEEE J. Quantum Electron., QE-5 (1969), s. 29.
22. T. L. Paoli, J. E. Ripper, T. H. Zachos, Resonant Modes of GaAs Junction Lasers — II: High-Injection Level, IEEE J. Quantum Electron., QE-5 (1969), s. 271.
23. T. H. Zachos, J. C. Dyment, Resonant Modes of GaAs Junction Lasers — III. Propagation Characteristics of Laser Beams with Rectangular Symmetry, IEEE J. Quantum Electron., QE-6 (1970), s. 317.
24. T. L. Paoli, J. E. Ripper, Direct Modulation of Semiconductor Lasers, Proc. IEEE, 58 (1970), s. 1457.
25. M. J. O. Strutt, Current Noise and Photon Noise of Semiconductor Diode Injection Lasers, Scientia Elec. (Switzerland), 17 (1971), s. 15.
26. A. R. Goodwin, P. R. Selway, Gain and Loss Processes in GaAlAs—GaAs Heterostructure Lasers, IEEE J. Quantum Electron., QE-6 (1970), s. 285.
27. L. A. D'Asaro, J. E. Ripper, Junction Lasers, Phys. Today, 24 (1971), s. 42.
28. W. Büchtemann, D. H. Höhn, Kohärenzmessungen an gepulsten GaAs-Laserdioden, Optik (Germany), 29 (1969), s. 401.
29. W. P. Dumke, The Injection Laser, w: Advances in Lasers, A. K. Levine (ed.), Vol. 2, Dekker, New York 1966.
30. F. Stern, Radiation Confinement in Semiconductor Lasers, w: Radiative Recombination in Semiconductors, Dunod, Paris 1964, s. 165.
31. M. M. Antonoff, Angular Distribution of Radiation from GaAs Injection Lasers, J. Appl. Phys., 35 (1964), s. 3623.
32. I. Hayashi, M. B. Panish, F. K. Reinhart, GaAs-$Al_xGa_{1-x}As$ Double Heterostructure Injection Laser, J. Appl. Phys., 42 (1971), s. 1929.
33. R. C. C. Leite, A. Yariv, On Mode Confinement in p-n Junctions, Proc. IEEE, 51 (1963), s. 1035.
34. A. L. McWhorter, Electromagnetic Theory of the Semiconductor Junction Laser, Solid State Electronics (GB), 6 (1963), s. 417.
35. W. W. Anderson, Mode Confinement and Gain in Junction Lasers, IEEE J. Quantum Electron., QE-1 (1965), s. 228.
36. H. Kressel, H. Nelson, F. Z. Hawrylo, Control of Optical Losses in p-n Junction Lasers by Use of a Heterojunction: Theory and Experiment, J. Appl. Phys., 41 (1970), s. 2019.
37. M. J. Adams, M. Cross, Wave-Guiding Properties of GaAs — $Al_xGa_{1-x}As$ Heterostructure Lasers, Phys. Letters, 32A (1970), s. 207.
38. H. Yonezu, Analysis on the Wave Confinement in the $Ga_xAl_{1-x}As$ — GaAs Laser Diode, Japan. J. Appl. Phys., 9 (1970), s. 1013.
39. Zh. I. Alferov, V. M. Andreev, V. I. Borodulin, G. T. Pak, E. L. Portnoi, V. I. Shveikin, Effect of the Heterostructure Parameters on Characteristics of Injection Lasers in AlAs — GaAs System, w: Proc. of the Int. Conf. on the Phys. and Chem. of Semicond. Heterojunctions and Layer Structures, Vol. 2, Akadémiai Kiadó, Budapest 1971, s. 159.
40. F. K. Reinhart, I. Hayashi, M. B. Panish, Mode Reflectivity and Waveguide Properties of DH Injection Lasers, J. Appl. Phys., 42 (1971), s. 4466.
41. H. Kressel, J. K. Butler, F. Z. Hawrylo, H. F. Lockwood, M. Ettenberg, Mode Guiding in Symmetrical (AlGa)As-GaAs Heterojunction Lasers with Very Narrow Active Regions, RCA Rev. (USA), 32 (1971), s. 393.
42. P. G. Jelisiejew, Ob optimalnoj wołnowodnoj strukturie inżekcjonnowo łazera, Kratkije Soobszcz. po Fizikie, No 4 (1970), s. 3.

43. F. Stern, Dispersion of the Index of Refraction Near the Absorption Edge of Semiconductors, Phys. Rev., 133A (1964), s. 1653.
44. J. Hatz, E. Mohn, Calculations of Intrinsic Threshold for TE and TM Mode in GaAs Laser Diodes, IEEE J. Quantum Electron., QE-3 (1967), s. 656.
45. R. G. Allachwerdian, A. N. Orajewskij, A. F. Suczkow, Wlijanije wołnowodnych swojstw p-n pierechoda na gienieraciju łazernych diodow iz arsenida gallija, Fiz. Techn. Połuprowodnikow, 4 (1970), s. 341.
46. W. J. Turner, W. E. Reese, Absorption Data of Laser-Type GaAs at 300°K and 77°K, J. Appl. Phys., 35 (1964), s. 350.
47. D. E. Hill, Infrared Transmission and Fluorescence of Doped Gallium Arsenide, Phys. Rev., 133A (1964), s. 866.
48. D. T. F. Marple, Refractive Index of GaAs, J. Appl. Phys., 35 (1964), s. 1241.
49. R. E. Collin, Prowadzenie fal elektromagnetycznych, WNT, Warszawa 1966, Rozdz. 11.
50. K. Unger, Moden in einem Halbleiterlaser, Ann. Phys. (Germany), 19 (1967), s. 64.
51. A. E. Michel, E. J. Walker, M. I. Nathan, Determination of the Active Region in Light-Emitting GaAs Diodes, IBM J. Res. Developm. (USA), 7 (1963), s. 70.
52. G. Burns, R. A. Laff, S. E. Blum, F. H. Dill Jr., M. I. Nathan, Directionality Effects of GaAs Light-Emitting Diodes, IBM J. Res Developm. (USA), 7 (1963), s. 62.
53. R. F. Broom, C. H. Gooch, C. Hilsum, D. J. Oliver, Onset of Stimulated Emission from Gallium Arsenide Semiconductor Optical Masers, Nature, 198 (1963), s. 368.
54. G. E. Fenner, J. D. Kingsley, Spatial Distribution of Radiation from GaAs Lasers, J. Appl. Phys., 34 (1963), s. 3204.
55. J. Hatz, Laser Diodes Made from Dislocation-Free GaAs Showing a Homogeneous Near-Field Pattern, IEEE J. Quantum Electron., QE-3 (1967), s. 643.
56. J. Ripper, informacja prywatna.
57. B. O. Seraphin, H. E. Bennett, Semiconductors and Semimetals, R. K. Willardson, A. C. Beer (eds.), Vol. 3, Academic Press, London 1967, s. 499.
58. T. S. Moss, Optical Properties of Semiconductors, Butterworth, London 1961.
59. K. Unger, Spontane und induzierte Emission in Laserdioden — II. Berücksichtigung der Zustandsdichteschwänze, Z. Phys., 207 (1967), s. 332.
60. Y. P. Varshni, Temperature Dependence of the Energy Gap in Semiconductors, Physica (Netherlands), 34 (1967), s. 149.
61. M. J. Adams, A Simple Approximation for High-Temperature Properties of the Injection Laser, Br. J. Appl. Phys. (J. Phys. D) (GB), 2 (1969), s. 1549.
62. H. C. Casey Jr., M. B. Panish, Composition Dependence of the $Ga_{1-x}Al_xAs$ Direct and Indirect Energy Gaps, J. Appl. Phys., 40 (1969), s. 4910.
63. A. E. Michel, E. J. Walker, Symp. Optical Masers, Polytechnic Institute of Brooklyn, New York 1963, s. 471.
64. K. Weiser, F. Stern, High-Order Transverse Modes in GaAs Lasers, Appl. Phys. Letters, 5 (1964), s. 115.
65. M. M. Antonoff, Electromagnetic Properties of the Semiconductor Injection Laser, Bull. Am. Phys. Soc., 9 (1964), s. 269.
66. Y. Nannichi, On the Polarization of Light from the GaAs-Diode Laser, Japan. J. Appl. Phys., 4 (1964), s. 360.
67. W. N. Aljamowskij, W. S. Bagojew, Ju. N. Bierozaszwili, B. M. Wuł, O polarizacji izłuczenija diodow iz arsenida gallija, Fiz. Tverdogo Tela, 8 (1966), s. 1091.
68. E. Mohn, J. Hatz, Ch. Deutsch, Die Polarization der Strahlung von GaAs-Laserdioden, Z. Angew Math. and Phys. (Switzerland), 17 (1966), s. 476.
69. M. J. Coupland, K. G. Hambleton, C. Hilsum, Measurement of Amplification in a GaAs Injection Laser, Phys. Letters, 7 (1963), s. 231.
70. I. Sakuma, H. Yonezu, K. Nishida, K. Kobayashi, F. Sato, Y. Nannichi, Continuous Operation of Junction Lasers at Room Temperature, Japan. J. Appl. Phys., 10 (1971), s. 282.

71. H. F. Lockwood, H. Kressel, H. S. Sommers Jr., F. Z. Hawrylo, An Efficient Large Optical Cavity Injection Laser, Appl. Phys. Letters, 17 (1970), s. 499.
72. J. E. Ripper, J. C. Dyment, L. A. D'Asaro, T. L. Paoli, Stripe-Geometry Double Heterostructure Junction Lasers: Mode Structure and C. W. Operation above Room Temperature, Appl. Phys. Letters, 18 (1971), s. 155.
73. M. J. Adams, M. Cross, On the Polarization of Radiaton from Double-Heterostructure Injection Lasers, Electron. Letters (GB), 7 (1971), s. 569.
74. Ż. I. Ałfiorow, W. M. Andrejew, E. L. Portnoj, M. K. Trukan, Inżekcjonnyje łazery na osnowie gieteropierechodow w sistemie AlAs-GaAs s nizkim porogom gienieracji pri komnatnoj tempierature, Fiz. Techn. Połuprowodnikow, 3 (1969), s. 1328.
75. M. B. Panish, I. Hayashi, S. Sumski, Double Heterostructure Injection Lasers with Room-Temperature Thresholds as Low as 2300 A/cm^2, Appl. Phys. Letters, 16 (1970), s. 326.
76. H. Kressel, F. Z. Hawrylo, Fabry-Pérot Structure $Al_xGa_{1-x}As$ Injection Lasers with Room-Temperature Threshold Current Densities of 2530 A/cm^2, Appl. Phys. Letters, 17 (1970), s. 169.
77. A. K. Jonscher, A. Boyle, Internat. Symposium on GaAs, Reading (1966), s. 78.
78. A. K. Jonscher, S. Sengupta, IEEE Device Meeting, Bad Nauheim (1967).
79. J. C. Dyment, T. H. Zachos, Injection Laser Far Field Patterns with Gaussian Profiles in the Junction Plane, J. Appl. Phys., 39 (1968), s. 2923.
80. E. A. J. Marcatili, Modes in a Sequence of Thick Astigmatic Lens-Like Focusers, Bell Syst. Techn. J. (USA), 43 (1964), s. 2887.
81. T. H. Zachos, Gaussian Beams from GaAs Junction Lasers, Appl. Phys. Letters, 12 (1968), s. 318.
82. G. Burns, M. I. Nathan, p-n Junction Lasers, Proc. IEEE, 52 (1964), s. 770.
83. F. Oberhettinger, Tabellen zur Fourier Transformation, Springer, Berlin 1957.
84. S. N. Stolarow, Wlijanie wołnowodnych swojstw gieteropierechodnych słojew na osnownyje charakteristiki inżekcjonnych łazerow, Kvantovaya Elektron., 2 (1972), s. 69.
85. J. K. Butler, H. Kressel, Transverse Mode Selection in Injection Lasers with Widely Spaced Heterojunctions, J. Appl. Phys., 43 (1972), s. 3403.
86. M. Cross, M. J. Adams, Wave-Guiding Properties of Stripe-Geometry Double Heterostructure Injection Lasers, Solid State Electronics (GB), 15 (1972), s. 919.
87. W. M. Andrejew, W. I. Borodulin, W. P. Koniajew, G. T. Pak, A. I. Pietrow, E. L. Portnoj, W. I. Szwejkin, Prostranstwiennoje raspredielenije izłuczenija gieterołazera, Fiz. Techn. Połuprowodnikow, 6 (1972), s. 1739.
88. P. A. Kirkby, G. H. B. Thompson, The Effect of Double Heterojunction Waveguide Parameters on the Far Field Emission Patterns of Lasers, Opto-Electronics (GB), 4 (1972), s. 323.
89. M. R. Matthews, R. B. Dyott, W. P. Carling, Filaments as Optical Waveguides in Gallium-Arsenide Lasers, Electron. Lett. (GB), 8 (1972), s. 505.

M. OSIŃSKI

ELECTROMAGNETIC THEORY OF INJECTION LASER RADIATION

Summary

The article reviews the results that have been obtained by treating the semiconductor *p-n* junction laser as a waveguide. In the simplest case this waveguide consists of several layers of different complex dielectric constant value. The active layer, characterised by negative conductivity (which is equivalent to negative absorption coefficient, therefore to amplification of electromagnetic radiation), is bounded by passive layers with positive conductivity. The refraction coefficient of the active layer is greater than those of passive layers, so the waveguide effect arises. By solving the wave equation for such a waveguide, one can obtain the electromagnetic field inside the laser. This enables us to estimate correctly the spectra of emitted radiation and also near- and far-field patterns.

This model can be improved by modifying position dependence of the dielectric constant. If we consider that, we find our results in good agreement with experimental data, such as the polarization of radia-

tion and the mode selection. If one takes into account the position dependence of the complex dielectric constant not only in the direction perpendicular to the *p-n* junction but also in the junction plane, the radiation generating region will be condensed to a narrow filament. Therefore electromagnetic theory explains in a simple way the main properties of the semiconductor injection laser radiaton.

M. OSIŃSKI

THÉORIE ÉLECTROMAGNETIQUE DE LA RADIATION DU LASER SEMICONDUCTEUR

Résumé

On a présenté dans l'article un aperçu des résultats de la théorie électromagnetique des lasers courants à jonction semiconductrice. On présume dans cette théorie que la jonction *p-n* constitue un guide des ondes. On a passé en revue des modèles différentes utilisées à la description du rayonnement laserien dans le cadre de cette théorie. On obtient en effet le spectre de rayonnement, les modèles de rayonnement des champs lointains et près, la polarisation de la lumière laserienne, la successivité de l'excitation des modes transversals et la structure filamenteuse de rayonnement, tous conformément à l'expériment.

M. OSIŃSKI

ELEKTROMAGNETISCHE THEORIE DER INJEKTIONSLASERSTRAHLUNGEN

Zusammenfassung

Es wurde eine Übersicht über die Ergebnisse, die bei Anwendung der elektromagnetischen Theorie der Halbleiter-Injektionslaser erhalten worden waren, vorgenommen. Diese Theorie setzt voraus, daß der *p-n*-Übergang einen Lichtwellenleiter bildet. Verschiedene theoretische Modelle, die im Rahmen der vorgetragenen Theorie genutzt worden waren, wurden bei Anwendung für homoepitaktische und heteroepitaktische Injektionslaser diskutiert. Interferenzmuster, Spektra und die Faserstruktur des von den Lasern ausgestrahlten elektromagnetischen Feldes sind als Ergebnis verschiedener theoretischer Modelle dargestellt worden.

М. ОСИНЬСКИ

ЭЛЕКТРОМАГНИТНАЯ ТЕОРИЯ ИЗЛУЧЕНИЯ ИНЖЕКЦИОННОГО ЛАЗЕРА

Резюме

Работа посвящена описанию результатов, которые можно получить рассматривая полупроводниковый лазер на *n-n* переходе как волновод. В простейшем случае этот волновод составлен из слоев отличающихся значением комплексной диэлектрической постоянной. Активный слой, характеризующийся отрицательной проводимостью (которая эквивалентна отрицательному коэффициенту поглощения, т. е. усилению электромагнитного излучения), находится в оболочке пассивных слоев с положительной проводимостью. Коэффициент преломления внутри активного слоя больше чем внутри пассивных слоев, что ведет к возникновению волноводного эффекта. Решая волновое уравнение для такого волновода можно найти электромагнитное поле внутри лазера. Это позволяет правильно оценивать спектральный состав излучения и распределения поля в ближней и дальней зонах.

Такую модель можно улучшить принимая более сложную зависимость диэлектрической постоянной от положения в направлении перпендикулярном к *p-n* переходу лучше соответствующую действительности. Тогда результаты расчетов поляризации излучения и селекции модов точнее совпадают с экспериментом. Дополнительный учет зависимости диэлектрической постоянной от положения в плоскости *p-n* перехода ведет к сужению области генерирующей излучение до узкой нити. Таким образом электромагнитная теория объясняет главные свойства излучения инжекционного ПКГ.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 2, ss. 359—388

Minimalizacja liczby stanów automatów skończonych

ALEKSANDER ALBICKI (WARSZAWA)

Instytut Teleelektroniki Politechniki Warszawskiej

Otrzymano 23.6.1973

W niniejszym artykule przedstawiony jest jeden z etapów syntezy układów logicznych, mianowicie minimalizacja liczby stanów automatów. Podane są podstawowe definicje teorii automatów oraz niektóre związki określone na zbiorze stanów automatów. W szczególności rozpatrywane są: relacja pokrycia automatów oraz relacja słabego homomorfizmu. Przy pomocy tych relacji buduje się ogólny algorytm minimalizacji liczby stanów. W dalszej części podane są twierdzenia i definicje wykorzystywane w trakcie praktycznego rozwiązywania problemu minimalizacji automatów niezupełnych. Sam proces minimalizacji ilustrowany jest przykładem. Podanych jest 5 algorytmów, na podstawie których można opracowywać konkretne programy maszynowe. Na zakończenie przeprowadzona jest dyskusja osiągniętych w niniejszej pracy wyników.

1. WSTĘP

Przy projektowaniu układów logicznych często wykorzystuje się model abstrakcyjny jakim jest automat. W wielu przypadkach problem syntezy układów logicznych, polegający na znalezieniu układu logicznego realizującego dane funkcje, rozwiązuje się w następujący sposób:

a) przy pomocy pewnych reguł, np. języka wyrażeń regularnych, tworzy się automat,

b) przeprowadza się minimalizację liczby stanów wewnętrznych automatu,

c) koduje się automat, w rezultacie otrzymuje się układ logiczny,

d) przeprowadza się syntezę strukturalną i elektryczną układu logicznego.

Praca niniejsza dotyczy punktu b.

Minimalizacją liczby stanów automatów zaczęli zajmować się jako pierwsi: E. F. Moore (Gedanken — Experiments on Sequential Machines, 1956) i D. A. Huffman (The Synthesis of Sequential Switching Circuits, 1954). Prace ich dotyczą wyłącznie automatów zupełnych.

Następnie G. H. Mealy (A Method for Synthesizing Sequential Circuits, 1955) oraz M. C. Paull i S. H. Unger (Minimizing the Number of States in Incompletly Specified Sequential Switching Functions, 1959) podali algorytmy wyznaczania zbiorów niesprzecznych i maksymalnych zbiorów niesprzecznych. Prace ich dotyczyły także układów niezupełnych, ale w ujęciu tradycyjnym — analogicznym do automatów zupełnych.

S. Ginsburg (prace w latach 1959—1962) zauważył, że minimalizacji liczby stanów dla automatów niezupełnych nie można przeprowadzać takimi metodami jak dla zupełnych. Podał on definicję słowa odpowiedniego dla danego stanu automatu. Następnie J. Betty i R. B. Miller znacznie rozszerzyli pojęcia wprowadzone przez Ginsburga i podali metodę minimalizacji liczby stanów dla automatów niezupełnych.

W r. 1966 Grasselli i Luccio wykorzystali osiągnięcie w dziedzinie minimalizacji form boolowskich (prace Mc. Cluskey'a) i podali własną metodę otrzymania co najmniej jednego automatu minimalnego. Był to pierwszy krok na drodze do efektywnej metody rozwiązania problemu minimalizacji w ogólnym przypadku, a więc także i dla automatów niezupełnych.

W ostatnich latach notuje się ogromne zainteresowanie matematyków teorią automatów. Plonem tych zainteresowań w odniesieniu do minimalizacji automatów są prace A. Ginzburga, M. Yoeli, R. Yeha, A. Pu i innych. Starają się oni stworzyć aparat matematyczny, który opisywałby proces minimalizacji. I tak Ginzburg i Yoeli wprowadzają pojęcie relacji słabego homomorfizmu, Yeh zajmuje się badaniem relacji zwrotnych, symetrycznych ale nieprzechodnich i dochodzi do podobnych wniosków co Ginzburg i Yoeli. A. Pu wprowadza pojęcie * — pokryć i na tej bazie wyprowadza dekompozycję automatów.

Jednocześnie z rozwojem teorii automatów mnożą się prace dotyczące praktycznych zastosowań pojęć i twierdzeń tej teorii przy syntezie układów logicznych. Prace te prowadzone są w dwóch kierunkach. Jeden kierunek to stworzenie algorytmów ułatwiających ręczne obliczenia — są to przeważnie algorytmy opracowane z zastosowaniem metod graficznych. Jako przykład może tu posłużyć praca Tse-Yun-Fenga [14]. Drugi kierunek to opracowanie maszynowych metod minimalizacji — z dostępnych w ostatnich latach publikacji to przede wszystkim prace E. M. Constantiniego [1] oraz J. Kella [5].

W ciągu kilku ostatnich lat w Pracowni Teletransmisyjnych Układów Logicznych Instytutu Teleelektroniki Politechniki Warszawskiej wykonano szereg prac z zakresu minimalizacji liczby stanów automatów skończonych. Praca niniejsza jest jedną z nich.

2. PODSTAWOWE POJĘCIA MATEMATYCZNE

Niech będzie dany zbiór G oraz binarna operacja ν, która przyporządkowuje parom złożonym z elementów zbioru G jako wynik elementy z tego samego zbioru. Zamiast pisać formalnie $\langle g_1, g_2\rangle\nu = g_3$, gdzie $g_1, g_2, g_3 \in G$, dalej będzie używany prostszy zapis: $g_1 g_2 = g_3$.

Definicja 1. Jeżeli operacja ν jest określona dla wszystkich par iloczynu kartezjańskiego $G \times G$, to nazywa się ją zupełną, a zbiór G z tą operacją $\langle G, \nu\rangle$ nazywa się grupoidem. Często zamiast pisać grupoid $\langle G, \nu\rangle$ będzie zapisane po prostu grupoid G.

Definicja 2. Jeżeli operacja w grupoidzie spełnia własność łączności

$$\bigwedge g_1, g_2, g_3 \in G \quad (g_1 g_2) g_3 = g_1 (g_2 g_3),$$

to grupoid taki nazywa się półgrupą.

Definicja 3. Półgrupa G nazywa się przemienną (lub abelową) wtedy i tylko wtedy, gdy $\bigwedge g_1, g_2 \in G \quad g_1 g_2 = g_2 g_1$.

Definicja 4. Niech F i K będą podzbiorami G. Wówczas

$$FK = \{fk | f \in F \wedge k \in K\}$$

nazywa się iloczynem względnym tych dwóch podzbiorów.

Definicja 5. Niepusty podzbiór $F \subseteq G$ (G jest półgrupą) nazywa się podpółgrupą, gdy $FF \subseteq F$.

Definicja 6. Niech $F \subseteq G$ (G jest półgrupą) i F^n będzie indukcyjnie zdefiniowane

$$F^1 = F, \quad F^{n+1} = FF^n.$$

Wówczas $F^* = \bigcup_{n \in N} F^n$ nazywa się podpółgrupą wygenerowaną przez F.

Definicja 7. Niech $F \subseteq G$ i G będzie półgrupą.
F nazywa się zbiorem generatorów G wtedy i tylko wtedy, gdy $F^* = G$.

Definicja 8. Niech G i F będą półgrupami.
Funkcja $\varphi: G \to F$ nazywa się homomorfizmem G w F wtedy i tylko wtedy, gdy

$$\bigwedge g_1, g_2 \in G \quad \varphi(g_1, g_2) = \varphi(g_1)\varphi(g_2). \tag{1}$$

Symbol Hom (G, F) oznacza zbiór wszystkich homomorfizmów G w F. Jeśli φ spełnia warunek (1) i jest bijekcją, to nazywa się ją izomorfizmem.

Definicja 9. Niech $F \subseteq G$. F nazywa się zbiorem wolnych generatorów wtedy i tylko wtedy, gdy

1) $F^* = G$,

2) dla każdej półgrupy H i dla każdej funkcji $\varphi_0: F \to H$ istnieje $\varphi \in \mathrm{Hom}(G, H)$ takie, że $\varphi|_F = \varphi_0$ ($\varphi|_F$ oznacza ograniczenie dziedziny funkcji tylko dla elementów zbioru F).

W tym przypadku G nazywa się wolną półgrupą ze zbiorem generatorów F.

Definicja 10. Element $e \in G$ nazywa się jednością w G wtedy i tylko wtedy, gdy $\bigwedge g \in G, \quad eg = ge = g$.
Półgrupa G z jednością nazywa się monoidem.

Twierdzenie. Każdy monoid G jest wiernie reprezentowany przez półgrupę odwzorowań K (K jest izomorficzne z G).

Dowód. patrz [3, rozdz. 1.5].

Definicja 11. Niech dana będzie relacja binarna $p \subseteq G \times G$. Jeśli ta relacja jest zwrotna, symetryczna i przechodnia, to nazywa się ją relacją równoważnościową.

Definicja 12. Niech ϱ będzie relacją równoważnościową określoną na półgrupie G:
ϱ nazywa się lewą kongruencją wtedy i tylko wtedy, gdy dla dowolnych

$$g_1, g_2, \gamma \in G \quad \langle g_1, g_2 \rangle \in \varrho \Rightarrow \langle \gamma g_1, \gamma g_2 \rangle \in \varrho,$$

ϱ nazywa się prawą kongruencją wtedy i tylko wtedy, gdy dla dowolnych

$$g_1, g_2, \gamma \in G \quad \langle g_1, g_2 \rangle \in \varrho \Rightarrow \langle g_1 \gamma, g_2 \gamma \rangle \in \varrho,$$

ϱ nazywa się kongruencją wtedy i tylko wtedy, gdy jest jednocześnie prawą i lewą kongruencją.

3. ELEMENTY TEORII ABSTRAKCYJNEJ AUTOMATÓW

Definicja 13. Półautomatem skończonym nazywamy trójkę $A = \langle S, X, \delta \rangle$, gdzie:

$S = \{s_0, s_1, \ldots, s_{n-1}\}$ — skończony zbiór stanów A,
$X = \{x_0, x_1, \ldots, x_{m-1}\}$ — skończony zbiór liter wejściowych A,
$\delta = \{\delta_{x_0}, \delta_{x_1}, \ldots, \delta_{x_{m-1}}\}$ — zbiór relacji $\delta_{x_i} \subseteq S \times S$, tzn. każdy element tego zbioru jest relacją.

Tak określony półautomat ze względu na to, że δ_{x_i} są relacjami nazywa się niedeterministycznym.

Definicja 14. Półautomatem deterministycznym jest trójka $A = \langle S, X, \delta \rangle$, w której S i X są jak poprzednio i $\delta = \{\delta_x\}_{x \in X}$, gdzie każda δ_{x_i} jest funkcją: $\delta_{x_i}: S \to S$. Deterministyczny półautomat jest więc szczególnym przypadkiem niedeterministycznego, w którym wszystkie relacje δ_x są funkcjami $S \to S$.

Rozpatrzmy działanie półautomatu deterministycznego; półautomat ten jest modelem abstrakcyjnym wiążącym zależności zachodzące w wolnej półgrupie słów X^* z zależnościami zachodzącymi w półgrupie półautomatu G_A (G_A jest skończonym monoidem funkcji generowanym przez funkcje δ_x i przez δ_0, gdzie 0 jest słowem pustym).

Każdemu słowu $x_i^* = x_{i1} x_{i2} \ldots x_{ik}$ odpowiada funkcja $\delta_{x_i^*} = \delta_{xi_1} \delta_{xi_2} \ldots \delta_{xi_k}$.

Relacja E na X^* zdefiniowana jako $x_1^* E x_2^* \Leftrightarrow \delta_{x_1^*} = \delta_{x_2^*}$ jest kongruencją (uwaga: dla konstrukcji półautomatu wystarczy prawa kongruencja), której klasy są złożone ze wszystkich takich słów w X^*, które indukują tę samą funkcję. Indeks E jest skończony, ponieważ S jest skończony i istnieje tylko skończona ilość różnych funkcji $\delta: S \to S$. Przy takiej konstrukcji półautomatu możliwe jest więc odwzorowywanie (istnieje homomorfizm $\varphi: X^* \to G_A$) zależności zachodzących między elementami X^*, których jest nieskończenie wiele, na zależności zachodzące w skończonej półgrupie G_A, przy czym G_A jest izomorficzne z X^*/E.

Na podstawie pojęcia półautomatu buduje się ważne, szczególnie z technicznego punktu widzenia, pojęcie automatu. Konstrukcję automatu przeprowadza się w ten sposób, że półautomat rozszerza się przez zaopatrzenie go w wyjścia. Uczynić to można na dwa sposoby, otrzymując dwie postacie automatu: automat Moore'a i automat Mealy'go.

Definicja 15. Automatem Moore'a jest piątka $\hat{A} = \langle S, X, Y, \delta, \lambda \rangle$, gdzie S, X i δ określone są jak poprzednio, Y — jest zbiorem liter wyjściowych, a λ jest relacją: $\lambda \subseteq S \times Y$. Jeśli każda δ_x i λ są funkcjami, to automat Moore'a jest deterministyczny, w przeciwnym wypadku automat jest niedeterministyczny.

Definicja 16. Automatem Mealy'ego jest piątka $\hat{A} = \langle S, X, Y, \delta, \lambda \rangle$, gdzie S, X, δ i Y określone są jak poprzednio, a $\lambda = \{\lambda_x\}_{x \in X}$, gdzie każda λ_x jest relacją $\lambda_x \subseteq S \times Y$. Jeżeli wszystkie δ_x i λ_x są funkcjami, to automat Mealy'ego jest deterministyczny. w przeciwnym wypadku automat jest niedeterministyczny.

Tak więc konstrukcja automatu Moore'a polega na tym, że niektórym jego stanom (możliwe, że wszystkim) przyporządkowuje się elementy zbioru wyjść Y. Konstrukcja automatu Mealy'ego polega na tym, że wyjścia w tym automacie (elementy zbioru Y) uzależnia się nie tylko od stanów, ale i od wejść, tzn. zamiast jednej wartości funkcji dla

jednego stanu s_i tak jak w automacie Moore'a, przyporządkowuje się temu stanowi zbiór wartości funkcji $\{s_i \lambda_x\}_{x \in \Lambda}$.

Widać stąd, że przy jednakowych zbiorach X i S automat Mealy'ego może wydawać więcej słów na wyjściu niż automat Moore'a (dokładniej odwzorowywać $X^* \to Y^*$), pomimo że w obydwu przypadkach zbiór wartości λ jest równy $Y (S\lambda = Y)$.

Definicja 17. Jeżeli $\bigwedge x \in X \quad pr_1 \delta_x = S \wedge pr_1 \lambda = S$,
to automaty takie nazywa się zupełnymi.

Przyjęte fest w literaturze technicznej pod pojęciem automat zupełny rozumieć automat deterministyczny zupełny; tak też przyjęte jest w niniejszej pracy.

Definicja 18. Automaty nie będące zupełnymi nazywa się niezupełnymi.

Poniżej będą rozpatrywane automaty niezupełne, w których δ_x lub λ_x mogą być relacjami funkcyjnymi lub funkcjami. W literaturze technicznej automaty takie nazywane są po prostu automatami niezupełnymi. I tak też w dalszej części niniejszej pracy będzie używany termin automat niezupełny. Dla większości stosowanych obecnie układów logicznych modelem abstrakcyjnym jest właśnie automat niezupełny. Należy również zwrócić uwagę na fakt, że automaty zupełne są szczególnym przypadkiem automatów niezupełnych.

Można wykazać, że określone wyżej postacie automatu Moore'a i Mealy'ego są sobie równoważne, tzn. dla każdego automatu Moore'a można skonstruować odpowiadający mu automat Mealy'ego i odwrotnie. Słowo „odpowiadający" oznacza tutaj, że obserwator z zewnątrz nie będzie mógł określić, z którą postacią automatu ma do czynienia, jeśli będzie nadawał na wejście automatu słowa wejściowe i odbierał z niego słowa wyjściowe. Dowód powyższego można przeprowadzić podając konstrukcję równoważnych sobie wzajemnie automatów Moore'a i Mealy'ego [10]. Poniżej, jeśli to nie będzie wyraźnie zaznaczone, rozpatrywane są automaty Mealy'ego.

Rozpatrzmy podstawowe zjawiska zachodzące w automacie. Niech dany będzie automat Mealy'ego $\hat{A} = \langle S, X, Y, \delta, \lambda \rangle$ i niech słowo $x^* \in X^*$ postaci $x^* = x_1 x_2 \dots x_k$, gdzie $x_i \in X$, będzie przyłożone do wejścia automatu. Wówczas relacje funkcyjne:

$$\delta_{x_1}, \delta_{x_1 x_2} = \delta_{x_1} \delta_{x_2}, \dots, \delta_{x_1 \dots x_k} = \delta_{x_1} \dots \delta_{x_k} \tag{2}$$

określają kolejne podzbiory stanów automatu $\hat{A}$ wykorzystywane przy przyłożeniu słowa x^*, a relacje funkcyjne

$$\lambda_{x_1}, \delta_{x_1} \lambda_{x_2}, \dots, \delta_{x_1 \dots x_{k-1}} \lambda_{x_k} \tag{2'}$$

określają kolejne wyjścia $\hat{A}$, gdy na wejście podane jest słowo x^*. W przypadku automatu niezupełnego obrazy relacji (2) lub (2') mogą być puste.

Powyższe relacje określają stany i słowa wyjściowe dla wszystkich możliwych stanów, do których słowo x^* jest przyłożone. Jeśli $\hat{A}$ jest w stanie s i przyłożymy słowo x^*, to kolejne wyjścia automatu określone będą następująco:

$$s\lambda_{x_1}, \quad s\delta_{x_1}\lambda_{x_2}, \quad \dots, \quad s\delta_{x_1 \dots x_{k-1}}\lambda_{x_k},$$

$s\lambda_{x^*} = s\delta_{x_1}\delta_{x_2} \dots \delta_{x_{k-1}}\lambda_{x_k}$ określa ostatnią literę wyjściową pojawiającą się w wyniku przyłożenia słowa x^* na wejście automatu $\hat{A}$ znajdującego się w stanie s.

Definicja 19. Słowo x^* nazywa się odpowiednim dla stanu s, jeśli $s\lambda_{x^*}$ jest określone.

Uwaga. W przypadku automatu Moore'a słowo wejściowe x^* jest odpowiednie dla stanu s, jeśli istnieje $s\delta_{x^*}\lambda$.

Wniosek. Dla automatów zupełnych wszystkie nadane słowa są odpowiednie.

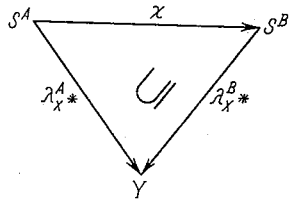
Pojęcie automatu może być interpretowane następująco: Automat $\hat{A} = \langle S, X, Y, \delta, \lambda\rangle$ jest tłumaczem słów półgrupy X^* na słowa ze zbioru Y^*. W przypadku automatów niezupełnych ze względu na definicję 19 wygodniej jest rozpatrywać automat jako tłumacz słów x^* na litery ze zbioru Y. Poniżej automat będzie rozpatrywany właśnie jako taki tłumacz. W zasadzie automat jest nawet zbiorem tłumaczy, gdyż słowo wyjściowe nie zależy li tylko od słowa wejściowego ale i od stanu $s \in S$, w którym znajduje się automat, gdy jest nadawana pierwsza litera słowa wejściowego.

Naturalnym problemem staje się szukanie prostszego automatu, który mógłby wykonać wszystkie te same zadania co i dany automat $\hat{A}$. Określenie „prostszy automat" jest nieprecyzyjne, jednakże można przyjąć, że automat prostszy jest taki, który ma mniej stanów wewnętrznych (kryterium 1). Znaczenie zdania „wykonywać wszystkie te same zadania co i dany automat $\hat{A}$" również należy dokładnie zinterpretować, mianowicie zdanie to znaczy: na słowa wejściowe odpowiednie dla $\hat{A}$ (w dowolnym stanie) odpowiadać tak jak i $\hat{A}$, a inne słowa odpowiedzi mogą być dowolne (kryterium 2). Powyższe rozumowanie prowadzi do znalezienia pewnej relacji (tzw. relacji pokrywania) między automatami określającej zależności porządkujące automaty według kryterium 2 („wykonywania więcej"), tzn. czy $\hat{A}$ może wykonać więcej niż $\hat{B}$ czy też odwrotnie.

Definicja 20. Automat $\hat{A} = \langle S^A, X, Y, \delta^A, \lambda^A\rangle$ jest pokryty przez automat $\hat{B} = \langle S^B, X, Y, \delta^B, \lambda^B\rangle$, co oznaczamy $\hat{A} \leqslant \hat{B}$, jeśli istnieje funkcja $\chi: S^A \to S^B$ taka, że

$$\bigwedge x^* \in X^*, \quad \lambda^A_{x^*} \subseteq \chi\lambda^B_{x^*}. \tag{3}$$

Zależność tę można przedstawić graficznie (rys. 1).



Rys. 1. Ilustracja graficzna relacji pokrywania automatów

W definicji 20 przyjęto założenie upraszczające, mianowicie: $X^A = X^B = X$ i $Y^A = Y^B = Y$. Te założenia nie zmniejszają ogólności definicji, gdyż zawsze można wprowadzić funkcje ze zbioru wejść pokrywanego automatu w zbiór wejść automatu pokrywającego oraz to samo uczynić dla wyjść. Najprościej to zrobić wprowadzając injekcje $i_1: X^A \to X^B$ oraz $i_2: Y^A \to Y^B$. Wyrażenie (3) oznacza, że dla każdego słowa $x^* \in X^*$ zbiór odpowiedzi automatu $\hat{A}$ jest zawarty w zbiorze odpowiedzi automatu $\hat{B}$. Dla każdego stanu $s^A \in S^A$ musi więc istnieć co najmniej jeden stan $s^B \in S^B$ taki,

że kiedy do automatu $\hat{B}$ w stanie s^B będą nadawane słowa x^* i do automatu $\hat{A}$ w stanie s^A będą nadawane te same słowa, to $\hat{B}$ wykona wszystkie te same tłumaczenia co i $\hat{A}$.

Własność 1. Relacja pokrywania jest zwrotna i przechodnia, lecz nie jest symetryczna:

$$\bigwedge \hat{A} \quad \hat{A} \leqslant \hat{A},$$
$$\bigwedge \hat{A}, \hat{B}, \hat{C} \quad \hat{A} \leqslant \hat{B} \wedge \hat{B} \leqslant \hat{C} \Rightarrow \hat{A} \leqslant \hat{C},$$
$$\bigwedge \hat{A}, \hat{B} \quad \hat{A} \leqslant \hat{B} \Rightarrow \hat{B} \leqslant \hat{A}.$$

Własność 2. Jeśli $\hat{A}$ jest zupełny, to wyrażenie (2) przyjmuje postać

$$\bigwedge x^* \in X^* \quad \lambda^A_{x^*} = \chi \lambda^B_{x^*}. \tag{3'}$$

Rzeczywiście: jeśli $\hat{A}$ jest zupełny, to $\lambda^A_{x^*}$ jest funkcją. Wówczas aby spełnić (3) to i $\chi\lambda^B_{x^*}$ musi być funkcją, a dwie funkcje określone na tym samym zbiorze i spełniające ponadto warunek (3) są sobie równe.

D e f i n i c j a 21. Jeśli $\hat{A} \leqslant \hat{B}$ i $\hat{B} \leqslant \hat{A}$, to automaty $\hat{A}$ i $\hat{B}$ są nieodróżnialne.

Lemat. Nieodróżnialność automatów jest relacją równoważnościową.

Dowód. Zwrotność i przechodniość relacji nieodróżnialności wynika bezpośrednio z własności 1. Założeniem definicji 21 jest symetryczność relacji pomiędzy A i B. Relacja nieodróżnialności jest przeto zwrotna, przechodnia i symetryczna, a więc jest równoważnością.

Twierdzenie 1. Jeśli automaty $\hat{A}$ i $\hat{B}$ są nieodróżnialne, to wykonują dokładnie takie same tłumaczenia.

Dowód.

$$\hat{A} \leqslant \hat{B} \Rightarrow \lambda^A_{x^*} \subseteq \chi_1 \lambda^B_{x^*}, \quad \text{gdzie} \quad \chi_1 : S^A \to S^B,$$
$$\hat{B} \leqslant \hat{A} \Rightarrow \lambda^B_{x^*} \subseteq \chi_2 \lambda^A_{x^*}, \quad \text{gdzie} \quad \chi_2 : S^B \to S^A$$
$$\text{a } \chi_1 \text{ i } \chi_2 \text{ są funkcjami.}$$
$$\hat{A} \leqslant \hat{B} \wedge \hat{B} \leqslant A \Rightarrow \lambda^A_{x^*} \subseteq \chi_1 \lambda^B_{x^*} \wedge \lambda^B_{x^*} \subseteq \chi_2 \lambda^A_{x^*},$$
$$\Rightarrow \lambda^A_{x^*} \subseteq \chi_1 \chi_2 \lambda^A_{x^*},$$
$$\Rightarrow \chi_1 \chi_2 = I_{S^A},$$
$$\Rightarrow \chi_1 = \chi_2^{-1},$$

więc χ_1 i χ_2 muszą być bijekcjami. Wówczas

$$\lambda^A_{x^*} \subseteq \chi_1 \lambda^B_{x^*} \Rightarrow \chi_2 \lambda^A_{x^*} \subseteq \chi_2 \chi_1 \lambda^B_{x^*} = \lambda^B_{x^*};$$

a przecież $\lambda^B_{x^*} \subseteq \chi_2 \lambda^A_{x^*}$ i stąd $\chi_2 \lambda^A_{x^*} = \lambda^B_{x^*}$, a χ_2 jest bijekcją.

Przeto dla każdego stanu $s^B \in S^B$ istnieje $s^A \in S^A$ taki, że gdy dowolne słowa wejściowe będą przykładane jednocześnie do obydwu tych stanów, to tłumaczenia wykonywane przez $\hat{A}$ i $\hat{B}$ będą jednakowe. Analogicznie można dowieść, że

$$\chi_1 \lambda^B_{x^*} = \lambda^A_{x^*},$$

a χ_1 jest bijekcją.

Przeto dla każdego stanu $s^A \in S^A$ istnieje $s^B \in S^B$ taki, że gdy dowolne słowa wejściowe będą przykładane jednocześnie do obydwu tych stanów, to tłumaczenia wykonywane przez $\hat{A}$ i $\hat{B}$ będą jednakowe. Stąd wynika, że wszystkie tłumaczenia wykonywane przez $\hat{A}$ są wykonywane przez $\hat{B}$ i odwrotnie, cbdo.

4. OGÓLNE ROZWIĄZANIE PROBLEMU MINIMALIZACJI LICZBY STANÓW

Dla każdego automatu $\hat{A}$ istnieje rodzina automatów pokrywająca go. Oznaczamy ją przez $\{\hat{A}_{pi}\}_{i\in I}$. Automaty, które należą do rodziny $\{\hat{A}_{pi}\}_{i\in I}$ i mają najmniej liczne zbiory S tworzą rodzinę automatów minimalnych. Oznaczamy ją przez $\{\hat{A}_{min.i}\}_{i\in I}$.

W szczególnym przypadku rodzina ta może zawierać jeden element i to właśnie $\hat{A}$.

Definicja 22. Automat nazywamy minimalnym do automatu $\hat{A}$ (oznaczamy go przez $\hat{A}_{\min}$) jeśli należy on do rodziny $\{\hat{A}_{min.i}\}_{i\in I}$:

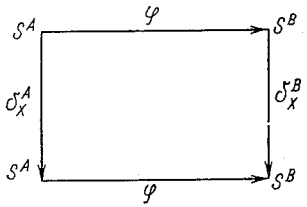
$$\hat{A}_{min} \in \{\hat{A}_{min.i}\}_{i\in I} \Leftrightarrow \hat{A} \leqslant \hat{A}_{min} \wedge [\neg \bigvee \hat{A}' \quad \hat{A} \leqslant \hat{A}' \wedge \operatorname{card} S^{A'} < \operatorname{card} S^{A_{min}}].$$

Definicja 20 umożliwia porównanie automatów między sobą. W celu znalezienia drogi do konstrukcji automatu pokrywającego dany $\hat{A}$ i jednocześnie takiego, aby był on prostszy od $\hat{A}$ (według kryterium 1), czyli do podania sposobu konstrukcji $\hat{B}$ takiego, że $\hat{A} \leqslant \hat{B}$ i $\operatorname{card} S^B \leqslant \operatorname{card} S^A$, należy szukać innych związków odpowiadających (3) ale przedstawionych algorytmicznie.

Intuicyjnie wydaje się, że związkami takimi powinny być morfizmy automatów, a w szczególności homomorfizm $\hat{A}$ na $\hat{B}$. Poniższe rozwiązania mają na celu znalezienie zależności pomiędzy relacjami homomorfizmu między automatami $\hat{A}$ i $\hat{B}$, a relacją pokrycia między tymi automatami.

Definicja 23. Funkcja (suriekcja) $\varphi: S^A \xrightarrow{\text{na}} S^B$ jest homomorfizmem $\hat{A}$ na $\hat{B}$, jeśli

$$\bigwedge x \in X \quad \delta_x^A \varphi \subseteq \varphi \delta_x^B \wedge \lambda_x^A \subseteq \varphi \lambda_x^B. \tag{4}$$



Rys. 2. Diagram obrazujący homomorfizm S^A na S^B

W przypadku automatów zupełnych warunek (4) upraszcza się do:

$$\bigwedge x \in X \quad \delta_x^A \varphi = \varphi \delta_x^B \wedge \lambda_x^A = \varphi \lambda_x^B. \tag{4'}$$

Jeśli φ jest bijekcją to automaty $\hat{A}$ i $\hat{B}$ są izoformiczne, tzn. różnią się tylko oznaczeniem stanów. Zależności powyższe wygodnie jest przedstawić na diagramach. Przemienność diagramu odpowiada własnościom morficznym i oznacza, że $\delta_x^A \varphi = \varphi \delta_x^B$ (rys. 2).

Definicja 24. Binarna relacja ψ, w której $pr_1\psi = S^A$ i $pr_2\psi = S^B$ jest słabym homomorfizmem $\hat{A}$ na $\hat{B}$, jeśli

$$\bigwedge x \in X \quad \psi^{-1}\delta_x^A \subseteq \delta_x^B\psi^{-1} \wedge \psi^{-1}\lambda_x^A \subseteq \lambda_x^B. \tag{5}$$

Można wykazać, że jeśli ψ jest funkcją, to warunki (4) i (5) są równoważne, tzn. każdy homomorfizm jest także słabym homomorfizmem i słaby homomorfizm, w którym ψ jest funkcją, jest homomorfizmem.

Twierdzenie 2. Jeśli ψ jest słabym homomorfizmen $\hat{A}$ na $\hat{B}$, to $\hat{A} \leqslant \hat{B}$.

Dowód. Załóżmy: $x^* = x_1 \ldots x_k$. Wtedy:

$$\psi^{-1}\delta_{x^*}^A = \psi^{-1}\delta_{x_1}^A \ldots \delta_{x_k}^A \subseteq \delta_{x_1}^B\psi^{-1}\delta_{x_2}^A \ldots \delta_{x_k}^A \subseteq \ldots \subseteq \delta_{x_1}^B \ldots \delta_{x_k}^B\psi^{-1} = \delta_{x^*}^B\psi^{-1}$$

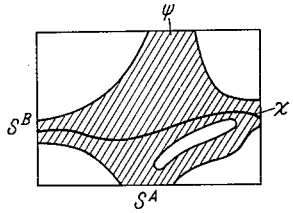
i

$$\psi^{-1}\lambda_{x^*}^A = \psi^{-1}\delta_{x_1}^A \ldots \delta_{x_{k-1}}^A\lambda_{x_k}^A \subseteq \delta_{x_1}^B\psi^{-1}\delta_{x_2}^A \ldots \delta_{x_{k-1}}^A\lambda_{x_k}^A \subseteq \ldots$$
$$\ldots \subseteq \delta_{x_1}^B \ldots \delta_{x_{k-1}}^B\psi^{-1}\lambda_{x_k}^A \subseteq \delta_{x_1}^B \ldots \delta_{x_{k-1}}^B\lambda_{x_k}^B = \lambda_{x^*}^B.$$

Ponieważ $pr_1\psi = S^A$ i $pr_2\psi = S^B$, to zawsze możliwe jest takie dobranie funkcji

$$\chi: S^A \to S^B, \quad \text{że} \quad \chi \subset \psi$$

(patrz rys. 3).



Rys. 3. Ilustracja graficzna dowodu twierdzenia 2

A więc

$$\bigwedge x^* \in X^*\psi^{-1}\lambda_{x^*}^A \subseteq \lambda_{x^*}^B \Rightarrow \chi\lambda_{x^*}^A \subseteq \lambda_{x^*}^B \Rightarrow \chi\chi^{-1}\lambda_{x^*}^A \subseteq \chi\lambda_{x^*}^B.$$

Ale χ jest funkcją, przeto $pr_1\chi = S^A$ i

$$\chi\chi^{-1} \subseteq I_{S^A}.$$

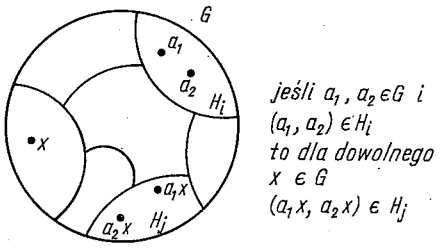
Podstawiając ten warunek wyżej otrzymamy

$$\lambda_{x^*}^A \subseteq \chi\lambda_{x^*}^B,$$

a to oznacza, że $\hat{A} \leqslant \hat{B}$ cbdo. Oczywiście $\hat{A} \leqslant \hat{B}$ nie implikuje istnienia ψ będącej słabym homomorfizmem $\hat{A}$ na $\hat{B}$.

Wprowadzone pojęcia homomorfizmu i słabego homomorfizmu będą służyć do następującej konstrukcji: mając dany automat $\hat{A}$ oraz π o własnościach podanych poniżej, utworzyć automat ilorazowy $\hat{A}/\pi$, tj. taki automat, którego stany odpowiadać będą podzbiorom zbioru stanów automatu $\hat{A}$ (istnieje bijekcja, przyporządkująca każdemu stanowi podzbiór zbioru stanów). W ten sposób można będzie otrzymać klasę automatów mniejszych od $\hat{A}$ i pokrywających go. Z tej klasy można będzie wybrać automaty minimalne. Dla zrozumienia zasady konstrukcji wydaje się celowym podanie kilku uzupełniających pojęć z algebry abstrakcyjnej.

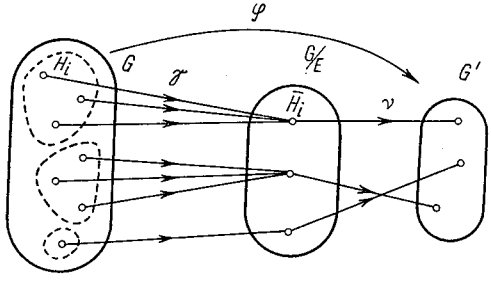
Pojęcie *podziału dopuszczalnego*. Podział $\pi = \{H_1, H_2, \ldots, H_k\}$ odpowiadający kongruencji E (E jest relacją równoważnościową) mający taką własność, że dla każdych dwóch bloków podziału H_1, H_2 istnieje jedyny blok H_3 taki, że $H_1 \cdot H_2 \subseteq H_3$, gdzie $H_1 \cdot H_2 = \{ab \mid a \in H_1 \wedge b \in H_2\}$ nazywa się podziałem dopuszczalnym.



Rys. 4. Podział dopuszczalny grupoidu G wyznaczony relacją kongruencji

Każda relacja kongruencji nad grupoidem G indukuje jego podział dopuszczalny (rys. 4). Związek między podziałem dopuszczalnym a homomorfizmem podaje poniższe twierdzenie.

Twierdzenie 3. Niech φ będzie homomorfizmem grupoidu G na grupoid G'. Wówczas $E = \varphi\varphi^{-1}$ jest kongruencją na G i istnieje izomorfizm $\nu: G/E \xrightarrow{\text{na}} G'$ taki, że $\varphi = \gamma\nu$, gdzie γ jest naturalnym homomorfizmem G na G/E (rys. 5).

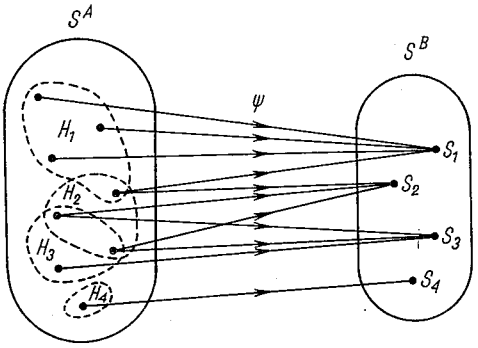


Rys. 5. Ilustracja graficzna homomorfizmu grupoidu G na grupoid G'

Dowód zamieszczony jest w [3].

Wniosek z twierdzenia 3. Każda kongruencja, tzn. każdy podział dopuszczalny grupoidu G, prowadzi do homomorficznego obrazu G, tzw. grupoidu ilorazowego G/E (jest to tzw. homomorfizm naturalny, który odwzorowuje każdy element G na jego klasę równoważności) i odwrotnie każdy homomorfizm G indukuje podział dopuszczalny grupoidu G, a odpowiadający mu grupoid ilorazowy jest izomorficzny z danym homomorficznym obrazem G.

Wyżej podane pojęcie podziału dopuszczalnego oraz jego związku z homomorfizmem zastosować można bezpośrednio do teorii automatów i rozszerzyć w sposób następujący: udowodnione było, że każdy homomorfizm indukuje podział dopuszczalny. Spróbujmy zatem sprawdzić co indukuje słaby homomorfizm $\psi: \hat{A} \xrightarrow{\text{na}} \hat{B}$. Okazuje się, jak to będzie wykazane niżej, że ψ indukuje dekompozycję, która dzieli zbiór stanów automatu $\hat{A}$ na rodziny jego podzbiorów: $\pi^d = \{H_i = s_i\psi^{-1}\}_{s_i \in S^B}$ (rys. 6).



Rys. 6. Ilustracja graficzna dekompozycji zbioru stanów automatu A

Ponieważ $pr_1\psi = S^A$, to każdy element S^A należy przynajmniej do jednego bloku dekompozycji π. Warunek ten dalej będzie nazywany warunkiem pokrycia. W szczególnym przypadku, gdy $H_i \cap H_j = \emptyset$ (dla $i \neq j$), to dekompozycja π staje się podziałem $S^A(\pi^d = \pi)$. Wychodząc bezpośrednio z określenia dekompozycji i biorąc pod uwagę to, że ψ jest słabym homomorfizmem, mamy:

$$\bigwedge x \in X \bigwedge H_i \in \pi^d \quad H_i\delta_x^A = s_i\psi^{-1}\delta_x^A \subseteq s_i\delta_x^B\psi^{-1} = s_j\psi^{-1} = H_j \tag{6}$$

oraz

$$\bigwedge x \in X \quad H_i\lambda_x^A = s_i\psi^{-1}\lambda_x^A \subseteq s_i\lambda_x^B, \tag{7}$$

gdzie

$$s_j = s_i\delta_x^B \quad \text{i} \quad s_i, s_j \in S^B.$$

Warunek (6) oznacza, że dla każdej litery x i każdego bloku H_1 dekompozycji π^d istnieje (co najmniej jeden) blok H_j, który zawiera $H_i\delta_x^A$. Warunek ten dalej będzie nazywany warunkiem zamkniętości. Taka dekompozycja nazywa się dopuszczalną dekompozycją S^A.

Warunek (7) oznacza, że wszelkie elementy w bloku H_i przy przyłożeniu jednakowych dopuszczalnych słów na wejście dają takie same odpowiedzi na wyjściu (może się zdarzyć, że wyjścia dla całego bloku są nieokreślone). O takiej dekompozycji mówi się, że jest wyjściowo-niesprzeczna.

Powyżej osiągnięte wyniki udowodniły następujące twierdzenie.

Twierdzenie 4. Słaby homomorfizm $\psi: \hat{A} \xrightarrow{\text{na}} \hat{B}$ indukuje dopuszczalną wyjściowo-niesprzeczną dekompozycję π^d (na zbiorze stanów S^A) taką, że bloki π^d są podzbiorami elementów s^A, które są w relacji ψ z tym samym elementem s^B.

Przez podanie sposobu konstrukcji można udowodnić twierdzenie odwrotne.

Twierdzenie 5. Dopuszczalna wyjściowo-niesprzeczna dekompozycja π^d (na zbiorze S^A) określa przynajmniej jeden automat $\hat{B} = \hat{A}/\pi^d$, który jest słabym homomorficznym obrazem $\hat{A}$.

Dowód. Konstrukcja automatu $\hat{B} = \hat{A}/\pi^d$. Dany $\hat{A} = \langle S^A, X^A, Y^A, \delta^A, \lambda^A\rangle$. Określić $\hat{B} = \hat{A}/\pi^d = \langle S^B, X^B, Y^B, \delta^B, \lambda^B\rangle$.

1. Przyjmujemy $X^B = X^A = X$ i $Y^B = Y^A = Y$.

2. Stanami B będzie zbiór: $S^B = \{\bar{H}_1\nu, \bar{H}_2\nu, \ldots, \bar{H}_k\nu\}$,
gdzie $\bar{H}_i$ wyznaczone są w sposób następujący: $\bigwedge H_i \in \pi^d \bigwedge s^A \in H_i \quad \bar{H}_i\nu = s^A\psi$ i ν jest bijekcją.

3. δ^B zdefiniujemy następująco: ponieważ $\bigwedge x \in X \bigwedge H_i \in \pi^d \bigwedge H_j \in \pi^d \quad H_i\delta_x^A \subseteq H_j$, to dla dowolnie wybranego H_j zdefiniujemy

$$\bar{H}_i\nu\delta_x^B = \bar{H}_j\nu.$$

4. λ^B definiujemy następująco: $\bar{H}_i\nu\lambda_x^B = H_i\lambda_x^A$.

Należy jeszcze udowodnić, że istnieje słaby homomorfizm $\hat{A}$ na $\hat{B} = \hat{A}/\pi^d$, a więc, że istnieje relacja ψ taka, że $pr_1\psi = S^A$ i $pr_2\psi = S^B$ i określona

$$\langle s^A, \bar{H}_i\nu\rangle \in \psi \Leftrightarrow s^A \in H_i,$$

gdzie

$$s^A \in S^A, \qquad \bar{H}_i\nu \in S^B$$

i spełniająca warunek (5).

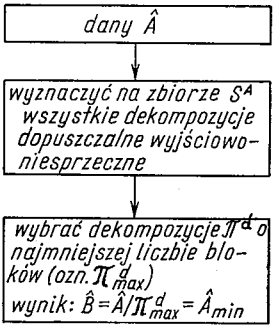
Sprawdzimy to:

$$\bigwedge x \in X \bigwedge \bar{H}_i\nu \in S^B \quad \bar{H}_i\nu\psi^{-1}\delta_x^A = H_i\delta_x^A \subseteq H_j = \bar{H}_j\nu\psi^{-1} = \bar{H}_i\nu\delta_x^B\psi^{-1}$$
$$\bar{H}_i\nu\psi^{-1}\lambda_x^A = H_i\lambda_x^A = \bar{H}_i\nu\lambda_x^B$$

Powyższe wyniki pokazują, że rzeczywiście ψ jest słabym homomorfizmem $\hat{A}$ na $\hat{B}$.

Przypadek szczególny: jeśli π^d jest dopuszczalnym wyjściowo-niesprzecznym podziałem S^A ($\pi^d = \pi$), to relacja ψ staje się funkcją i homomorfizmem $\hat{A}$ na $\hat{A}/\pi$. Przypadek ten obejmuje problem redukcji automatów zupełnych, a także redukcję wstępną automatów niezupełnych.

Na podstawie powyższego twierdzenia można bezpośrednio budować algorytmy minimalizacji liczby stanów automatów. Problem minimalizacji sprowadza się do poszukiwania na zbiorze stanów automatu dopuszczalnej wyjściowo-niesprzecznej dekompozycji takiej, aby liczba jej bloków była jak najmniejsza. Algorytm poszukiwania automatu minimalnego przedstawiony jest na diagramie (rys. 7).



Rys. 7. Diagram ogólnego algorytmu wyznaczania automatu minimalnego do danego

Wykonanie praktyczne takiego programu byłoby niezmiernie trudne ze względu na krytyczne przypadki, w których indukowałaby się olbrzymia liczba powstałych dekompozycji. W związku z tym nakłada się pewne warunki ograniczające, o których będzie mowa w następnym rozdziale.

5. ALGORYTM WYZNACZANIA DEKOMPOZYCJI MAKSYMALNYCH

5.1. Niezbędne definicje i twierdzenia

Techniczne sposoby wyszukiwania dekompozycji dopuszczalnych wyjściowo-niesprzecznych dla danego automatu $\hat{A}$ sprowadzają się do podania reguł tworzenia rodzin podzbiorów zbioru stanów automatu $\hat{A}$ takich, że te podzbiory będą blokami π^d.

Definicja 25. Para stanów (s_i, s_j) automatu $\hat{A}$ jest wyjściowo-niesprzeczna wtedy i tylko wtedy, gdy dla każdej litery wejściowej x przyłożonej jednocześnie do obydwu tych stanów litery wyjściowe są takie same jeśli obie są określone[1].

Wniosek. Para stanów jest wyjściowo-niesprzeczna wtedy i tylko wtedy, gdy należy ona do jądra relacji funkcyjnej $\lambda_x \lambda_x^{-1}$. (s_i, s_j) jest wyjściowo-niesprzeczne $\Leftrightarrow$

$$\Leftrightarrow \bigwedge x \in X \quad (s_i, s_j) \in \lambda_x \lambda_x^{-1} \Leftrightarrow \bigwedge x \in X \quad s_i \lambda_x \neq \emptyset \wedge s_j \lambda_x \neq \emptyset \Rightarrow s_i \lambda_x = s_j \lambda_x.$$

Definicja 26. Parę stanów (s_i, s_j) automatu $\hat{A}$ należącą do jednego bloku dekompozycji π^d nazywa się parą stanów niesprzecznych.

Każda para stanów niesprzecznych należy do jądra relacji ψ. Dalej relacja niesprzeczności oznaczana będzie $\psi\psi^{-1}$ albo „$\sim$". Każda para stanów (s_i, s_j) automatu $\hat{A}$ jest niesprzeczna wtedy i tylko wtedy, gdy dla każdej litery wejściowej x przyłożonej jednocześnie do obydwu tych stanów (s_i, s_j) jest wyjściowo-niesprzeczna i $(s_i \delta_x, s_j \delta_x)$ musi być parą niesprzeczną, a więc $(s_i \delta_x, s_j \delta_x)$ musi być wyjściowo-niesprzeczną i para stanów następnych do $s_i \delta_x$ i $s_j \delta_x$ też musi być niesprzeczną, a więc

$$\begin{aligned}(s_i, s_j) \in \psi\psi^{-1} \Leftrightarrow \bigwedge x \in X \quad & (s_i, s_j) \in \lambda_x \lambda_x^{-1} \wedge \\ & (s_i \delta_x, s_j \delta_x) \in \lambda_x \lambda_x^{-1} \wedge \\ & (s_i \delta_x \delta_x, s_j \delta_x \delta_x) \in \lambda_x \lambda_x^{-1} \wedge \\ & \ldots\ldots\ldots\ldots\ldots \\ & \ldots\ldots\ldots\ldots\ldots\end{aligned}$$

Wniosek. Każda para stanów (s_i, s_j) automatu $\hat{A}$ jest niesprzeczna wtedy i tylko wtedy, gdy dla każdego słowa $x^* \in X^*$ przyłożonego jednocześnie do obydwu tych stanów litery wyjściowe są takie same, jeśli obie są określone

$$\begin{aligned}(s_i, s_j) \in \psi\psi^{-1} &\Leftrightarrow \bigwedge x^* \in X^* \quad (s_i, s_j) \in \lambda_x \lambda_x^{-1} \\ &\Leftrightarrow \bigwedge x^* \in X^* \quad s_i \lambda_{x^*} \neq \emptyset \wedge s_j \lambda_{x^*} \neq \emptyset \Rightarrow s_i \lambda_{x^*} = s_j \lambda_{x^*}.\end{aligned} \tag{8}$$

Wniosek powyższy jest ogólnym i słusznym dla wszystkich automatów.

Dla dowolnego automatu skończonego card S jest skończona. Niech card $S = n$. Wówczas maksymalna ilość różnych par stanów niesprzecznych wynosi

$$\frac{n!}{(n-2)!} \cdot \frac{1}{2} = \frac{n(n-1)}{2}$$

[1] Jeśli $s\lambda_x = \emptyset$, to s tworzy parę wyjściowo-niesprzeczną z każdym stanem automatu.

(współczynnik 1/2 wynika z symetrii relacji niesprzeczności) i wtedy wyrażenie (8) przyjmie postać

$$(s_i, s_j) \in \psi\psi^{-1} \Leftrightarrow \bigwedge x^* \in X^* \quad \operatorname{card} x^* < \frac{n(n-1)}{2} \Rightarrow (s_i, s_j) \in \lambda_{x^*}\lambda_{x^*}^{-1}. \tag{8'}$$

Z wyrażenia (8') można bezpośrednio wyliczyć wszystkie pary stanów niesprzecznych danego automatu skończonego $\hat{A}$.

Relacja niesprzeczności jest zwrotna i, jak już wyżej powiedziano, symetryczna. W ogólnym przypadku nie jest ona przechodnia. Poniżej podane będą podstawowe definicje i twierdzenia, za pomocą których rozwiązany został maszynowy program minimalizacji stanów automatów skończonych.

Praktyczną ilustracją dla każdej definicji będzie rozwiązanie klasycznego już przykładu z artykułu Grasselliego i Luccio [4.].

Przykład. Automat zadany jest tablicą przejść-wyjść:

S \ X	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_1	x_2	x_3	x_4	x_5	x_6	x_7
a	a	—	d	e	b	a	—	0	—	0	1	0	—	—
b	b	d	a	—	a	a	—	0	1	—	—	—	1	—
c	b	d	a	—	—	—	g	0	1	1	—	—	—	0
d	—	e	—	b	b	—	a	—	—	—	—	0	—	—
e	b	e	a	—	b	e	a	—	—	—	—	—	—	1
f	b	c	—	h	f	g	—	0	—	1	1	1	0	—
g	—	c	—	e	—	g	f	—	1	—	1	—	0	0
h	a	e	d	b	b	e	a	1	0	1	0	—	—	1

$X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$, card $X = 7$

$S = \{a, b, c, d, e, f, g, h\}$, card $S = 8$

$Y = \{0, 1\}$, card $Y = 2$

Dalej ten automat nazywany będzie automatem *GL*.

Definicja 27. Podzbiór S' zbioru stanów S jest podzbiorem niesprzecznym (ozn. S^n) wtedy, gdy każda para stanów należąca do S jest niesprzeczna:

$$S' = S^n \Leftrightarrow S' \subseteq S \wedge \bigwedge s_i, s_j \in S' \quad (s_i, s_j) \in \psi\psi^{-1}.$$

Przykład. Wszystkie podzbiory niesprzeczne w automacie *GL*.

Uwaga. Ze względu na prostotę zapisu zamiast pary (s, s) zapisane jest wszędzie s.

$\{a,\}$
$\{b,\}$
$\{c,\}$
$\{d,\}$
$\{e,\}$
$\{f,\}$
$\{g,\}$
$\{h,\}$
$\{a, b,\}$

$\{a, d,\}$
$\{a, e,\}$
$\{a, g,\}$
$\{b, c,\}$
$\{b, d,\}$
$\{b, e,\}$
$\{c, d,\}$
$\{c, f,\}$
$\{c, g,\}$
$\{d, e,\}$
$\{d, h,\}$
$\{e, h,\}$
$\{f, g,\}$
$\{a, b, d,\}$
$\{a, b, e,\}$
$\{a, d, e,\}$
$\{b, c, d,\}$
$\{b, d, e,\}$
$\{c, f, g,\}$
$\{d, e, h,\}$
$\{a, b, d, e,\}$

Definicja 28. Podzbiór niesprzeczny S^n jest niesprzeczny ze stanem s_a (ozn. $s_a \sim S^n$) wtedy i tylko wtedy, gdy s_a jest niesprzeczny z każdym stanem należącym do S^n

$$s \sim S^n \Leftrightarrow \bigwedge s \in S^n \quad (s, s_a) \in \psi\psi^{-1}.$$

Przykład. Wybierzmy dowolny podzbiór niesprzeczny automatu GL i dowolny stan tegoż automatu. Np. niech $S^n = \{a, b\}$, a $s_a = d$. Ponieważ $a \sim d$ i $b \sim d$, to $d \sim \{a, b\}$.

Definicja 29. Podzbiór S^n zbioru stanu S jest maksymalnym podzbiorem niesprzecznym (ozn. S^{mn}), jeśli S^n jest podzbiorem niesprzecznym i w zbiorze stanów $(S-S^n)$ nie ma takiego stanu, który byłby niesprzeczny z S^n

$$S^n = S^{mn} \Leftrightarrow \neg\bigvee s \in (S-S^n) \quad s \sim S^n.$$

Przykład. Maksymalne podzbiory niesprzeczne automatu GL:

$\{a, b, d, e\}$
$\{b, c, d\}$
$\{c, f, g\}$
$\{d, e, h\}$
$\{a, g\}$

Znalezienie maksymalnych podzbiorów stanów niesprzecznych jest końcowym etapem minimalizacji liczby stanów automatów zupełnych [7]. W przypadku automatów niezupełnych, ze względu na nieprzechodniość relacji niesprzeczności, liczność rodziny maksymalnych podzbiorów stanów niesprzecznych może znacznie przekraczać wartość card S [10]. Należy więc szukać takich dekompozycji automatów niezupełnych, które zawierałyby najmniej bloków (oznaczamy je przez π^d_{max}). Aby wyznaczyć wszystkie dekompo-

zycje π^d_{max}, czyli pełną rodzinę dekompozycji maksymalnych (ozn. $\{\pi^d_{max}\}_p$), trzeba dokonać wyboru z rodziny podzbiorów niesprzecznych. Proces ten może być niezmiernie trudny do spełnienia, a to ze względu na dużą zwykle liczność rodziny podzbiorów niesprzecznych. Ponieważ uzyskanie tą drogą rozwiązania może być niemożliwe (ze względu na ograniczoność pamięci i szybkości działania maszyny cyfrowej) poniżej podana będzie metoda, dzięki której otrzymana będzie co najmniej jedna dekompozycja maksymalna. Metoda polega na selekcji podzbiorów niesprzecznych tak, że niektóre z nich mogą być wykluczone i nie brane pod uwagę przy wyborze π^d_{max}. Obecnie będą podane te metody selekcji, które były wykorzystywane przy opracowaniu algorytmów.

Dla każdego podzbioru niesprzecznego S^n istnieje rodzina zbiorów $\{S^n\delta_{x_l}\}_{x_l\in X}$ wyznaczona przez działanie funkcji δ_x automatu.

P r z y k ł a d. Wybierzmy dowolny podzbiór niesprzeczny automatu *GL*. Np. niech $S = \{a, d, e\}$; wtedy

$$\{S^n\delta_{x_l}\}_{x_l\in X} = \{\{a, d, e\}\delta_{x_l}\}_{x_l\in X} = = \{\{a, b\}, \{e\}, \{a, d\}, \{b, e\}, \{b\}, \{a, e\}, \{a\}\}.$$

D e f i n i c j a 30. Jeśli elementy rodziny $\{S^n\delta_{x_l}\}_{x_l\in X}$ spełniają poniższe 3 warunki:

1. $\operatorname{card} S^n\delta_{x_l} > 1$,
2. $S^n\delta_{x_l} \nsubseteq S^n$,
3. $\bigwedge i \neq j \quad S^n\delta_{x_l} \nsubseteq S^n\delta_{x_j}$,

to rodzinę tę nazywa się rodziną implikowaną przez S^n i oznacza przez P.

P r z y k ł a d. Rodziny implikowane przez podzbiory maksymalne niesprzeczne automatu *GL*:

S^{mn}	P
$\{a, b, d, e\}$	$\emptyset$
$\{b, c, d\}$	$\{\{a, b\}, \{a, g\}, \{d, e\}\}$
$\{c, f, g\}$	$\{\{c, d\}, \{e, h\}\}$
$\{d, e, h\}$	$\{\{a, b\}, \{a, d\}\}$
$\{a, g\}$	$\emptyset$

Niesprzeczny zbiór S^n_β jest wykluczony przez niesprzeczny zbiór S^n_α, jeśli

$$S^n_\alpha \supset S^n_\beta \quad \text{i} \quad P_\alpha \subseteq P_\beta.$$

P r z y k ł a d. Wybierzmy dowolne podzbiory niesprzeczne automatu *GL*.
Np. niech $S^n_\alpha = \{d, e, h\}$ a $S^n_\beta = \{e, h\}$, wtedy:

$$P_\alpha = \{\{a, b\}, \{a, d\}\},$$
$$P^n = \{\{a, b\}, \{a, d\}\}.$$

Ponieważ $\{d, e, h\} \supset \{e, h\}$, i $\{\{a, b\}, \{a, d\}\} \subseteq \{\{a, b\}, \{a, d\}\}$, to $\{e, h\}$ jest wykluczony przez $\{d, e, h\}$.

D e f i n i c j a 31. Niesprzeczny zbiór, który nie jest wykluczony przez żaden inny zbiór niesprzeczny, nazywa się klasą prostą[2]. Dla oznaczenia klas prostych stosowany będzie dalej symbol C (ewentualnie z indeksem).

[2] Termin klasa jest tutaj synonimem terminu zbiór.

Przykład. Klasy proste automatu *GL*:

$\{a, b, d, e\}$
$\{b, c, d\}$
$\{c, f, g\}$
$\{d, e, h\}$
$\{a, g\}$
$\{b, c\}$
$\{c, d\}$
$\{c, f\}$
$\{c, g\}$
$\{d, h\}$
$\{f, g\}$
$\{f\}$

Na podstawie pojęcia klas prostych można wyprowadzić rodzinę dekompozycji maksymalnych oznaczoną $\{\pi_{max}^{dKP}\}_p$, (członkami każdej dekompozycji π_{max}^{dKP} są wyłącznie klasy proste), przy czym oczywiście $\{\pi_{max}^{dKP}\}_p \subseteq \{\pi_d^{max}\}_p$.

Twierdzenie 6. Istnieje przynajmniej jedna dekompozycja maksymalna zawierająca wyłącznie klasy proste.

Dowód zamieszczony jest w [4].

Można wykazać, że w celu wyznaczenia dekompozycji maksymalnych π_{max}^{dKP} nie wszystkie klasy proste muszą być wykorzystywane. Poniżej podane będą warunki pozwalające na eliminację niektórych z nich.

Definicja 32. Klasa prosta, która nie może być członem dekompozycji maksymalnych, nazywa się nadmiarową klasą prostą. Poniżej podane będą warunki, które pozwolą na usunięcie ze zbioru klas prostych, klas prostych nadmiarowych.

Definicja 33. Zbiór implikujący bezpośrednio (ozn. M_α) klasę prostą C_α jest zbiorem wszystkich klas prostych, takich które mają klasę C jako jedną z ich implikowanych klas prostych

$$M_\alpha = \{C_i | \bigvee x \in X \quad \operatorname{card} C_i\delta_x > 1 \wedge C_i\delta_x \subseteq C_\alpha \wedge i \neq \alpha\}.$$

Przykład. Wybierzmy dowolną klasę prostą automatu *GL*. Np. niech $C_\alpha = \{a, b, d, e\}$. Wówczas $M_\alpha = \{\{b, c, d\}, \{d, e, h\}, \{c, d\}\}$.

Definicja 34. Zbiór implikujący (ozn. M_α^*) klasę prostą C_α jest zbiorem wszystkich klas prostych takich, które mają klasę C_α jako jedną z ich implikowanych nie tylko bezpośrednio klas prostych:

$$M_\alpha^* = \{C_i | \bigvee x^* \in X^* \quad \operatorname{card} C_i\delta_{x^*} > 1 \wedge C_i\delta_{x^*} \subseteq C_\alpha \wedge i \neq \alpha\}.$$

Przykład. Wybierzmy klasę prostą z poprzedniego przykładu: Wówczas

$$M_\alpha^* = \{\{b, c, d\}, \{d, e, h\}, \{c, d\}, \{c, f, g\}, \{c, f\}, \{c, g\}, \{f, g\}\}.$$

Definicja 35. Zbiór implikowany bezpośrednio (ozn. N_α) przez klasę prostą C_α jest zbiorem wszystkich klas prostych, które są implikowane przez C_α:

$$N_\alpha = \{C_i | \bigvee x \in X \quad \operatorname{card} C_\alpha \delta_x > 1 \wedge C_\alpha \delta_x \subseteq C_i \wedge i \neq \alpha\}.$$

Przykład. Wybierzmy dowolną klasę prostą automatu GL. Np. niech $C_\alpha = \{c, f\}$. Wówczas $N_\alpha = \{\{b, c, d\}, \{c, d\}\}$.

Definicja 36. Zbiór implikowany (ozn. N_α^*) przez klasę prostą C_α jest zbiorem wszystkich klas prostych, które są implikowane nie tylko bezpośrednio ale i niebezpośrednio przez C_α:

$$N_\alpha^* = \{C_i | \bigvee x^* \in X^* \quad \operatorname{card} C_\alpha \delta_{x^*} > 1 \wedge C_\alpha \delta_{x^*} \subseteq C_i \wedge i \neq \alpha\}.$$

Przykład. Wybierzmy klasę prostą z poprzedniego przykładu. Wówczas $N_\alpha^* = \{\{b, c, d\}, \{c, d\}, \{a, b, d, e\}, \{a, g\}, \{d, e, h\}\}$.

Definicja 37. Klasa prosta nazywa się klasą prostą inicjalną wtedy i tylko wtedy jeśli jej zbiór implikujący jest pusty:

$$C_\alpha = C_{inicjal} \Leftrightarrow M_\alpha = \emptyset.$$

Przykład. Wśród klas prostych automatu GL klasami prostymi inicjalnymi są: $\{c, g\}$ i $\{c, f\}$.

Definicja 38. Klasa prosta nazywa się klasą prostą końcową wtedy i tylko wtedy, gdy zbiór przez nią implikowany jest pusty:

$$C_\alpha = C_{końc} \Leftrightarrow N_\alpha = \emptyset.$$

Przykład. Wśród klas prostych automatu GL klasami prostymi końcowymi są:

$$\{a, b, d, e\}, \{b, c\}, \{d, h\}, \{a, g\}, \{f\}.$$

Definicja 39. Klasę prostą nazywa się pojedynczą klasą prostą wtedy i tylko wtedy, gdy zbiór ją implikujący i zbiór przez nią implikowany są puste:

$$C_\alpha = C_{pojedyn} \Leftrightarrow M_\alpha = N_\alpha = \emptyset.$$

Przykład. Wśród klas prostych automatu GL pojedynczymi klasami prostymi są:

$$\{b, c\}, \{d, h\}, \{f\}.$$

Definicja 40. Klasa prosta nazywa się podstawową klasą prostą wtedy i tylko wtedy, gdy zawiera element nie zawarty w żadnej innej klasie prostej:

$$C_\alpha = C_{podst} \Leftrightarrow \bigvee s \in S \quad [s \in C_\alpha \wedge \neg \bigvee i \in I (i \neq \alpha \Rightarrow s \in C_i)].$$

Przykład. Wśród klas prostych automatu GL nie ma klas prostych podstawowych.

Wniosek. Każda klasa prosta podstawowa jest zawsze członem dekompozycji maksymalnej.

D e f i n i c j a 41. Podłańcuchem dla klasy prostej C_α (ozn. C_α^{pL}) nazywać będziemy podzbiór zbioru N_α^* taki, że dla każdego elementu C_i tego podzbioru istnieje dokładnie jeden element C_j i $C_i\delta_x \subseteq C_j$:

$$C_\alpha^{pL} = \{C_i | C_i \in N_\alpha^* \wedge \bigvee ! j \in I \quad C_j \in (N_{\alpha_i}^* \cup C_\alpha) \wedge C_i\delta_x \subseteq C_j\}.$$

P r z y k ł a d. Wybierzmy dowolną klasę prostą automatu GL. Np. niech $C_\alpha = (\{c, f, g\}$. Podłańcuchami do tej klasy prostej będą:

$$C_{\alpha 1}^{pL} = \{\{b, c, d\},\ \{a, b, d, e\}\},$$
$$C_{\alpha 2}^{pL} = \{\{b, c, d\},\ \{a, g\}\},$$
$$C_{\alpha 3}^{pL} = \{\{d, e, h\},\ \{a, b, d, e\}\},$$
$$C_{\alpha 4}^{pL} = \{\{b, c, d\},\ \{d. e, h\},\ \{a, b, d, e\}\},$$

Z df. 41 wynikają bezpośrednio następujące wnioski.

Wniosek 1. Każdy podłańcuch jest zbiorem klas prostych, który nie musi spełniać warunków zamkniętości.

Wniosek 2. Maksymalna długość podłańcucha dla danej klasy prostej wynosi $l = \text{card}\ \{C_i\}_{i\in I} - 1$.

Wniosek 3. Podłańcucha nie można podzielić na takie części, które byłyby podłańcuchami.

D e f i n i c j a 42. Łańcuchem dla klasy prostej C_α (ozn. C_α^L) nazywać będziemy taką sumę podłańcuchów tej klasy, że każdy element P_α jest zawarty w co najmniej jednej klasie prostej należącej do tej sumy:

$$C_\alpha^L = \{\bigcup_{i\in I} C_{\alpha_i}^{pL} | \bigwedge S^n \in P_\alpha \bigvee j \in I \quad S^n \subseteq C_j \wedge C_j \in \bigcup_{i\in I} C_{\alpha_i}^{pL}\}.$$

P r z y k ł a d. Łańcuchami dla klasy prostej z poprzedniego przykładu są:

$$C_{\alpha_1}^L = \{\{b, c, d\},\ \{d, e, h\},\ \{a, b, d, e\}\},$$
$$C_{\alpha_2}^L = \{\{b, c, d\},\ \{a, g\},\ \{d, e, h\},\ \{a, b, d, e\}\}.$$

Z df. 42 wynikają bezpośrednio następujące wnioski.

Wniosek 1. Każdy łańcuch jest zbiorem klas prostych, który nie musi spełniać warunków zamkniętości.

Wniosek 2. Maksymalna długość łańcucha dla danej klasy prostej wynosi $l = \text{card}\ \{C_i\}_{i\in I} - 1$.

D e f i n i c j a 43. Łańcuch dla klasy prostej C_α nazywać będziemy prostym łańcuchem dla tej klasy, jeśli nie można go podzielić na inne łańcuchy.

P r z y k ł a d. Wśród łańcuchów z poprzedniego przykładu prostym łańcuchem jest:

$$C_\alpha^L = \{\{b, c, d\},\ \{d, e, h\},\ \{a, b, d, e\}\}.$$

Z df. 43 wynika bezpośrednio następujący wniosek.

Wniosek 1. Łańcuchem prostym dla każdej klasy prostej pojedynczej lub końcowej jest zbiór Ø.

Definicja 44. Łańcuch prosty dla klasy prostej C_α nazywa się zamkniętym, gdy

$$C_i \in C_\alpha^{\mathcal{L}} \Rightarrow \bigwedge S^n \in P_i \bigvee j \in I \quad C_j \in C_\alpha^{\mathcal{L}} \wedge S^n \subseteq C_j.$$

Twierdzenie 7. Zbiór klas prostych wraz z ich prostymi łańcuchami zamkniętymi spełnia zawsze warunki zamkniętości.

Dowód. Ponieważ każdy łańcuch dla każdej klasy prostej jest zamknięty, to bezpośrednio z definicji zamkniętości wynika, że zbiór klas prostych wraz z ich łańcuchami zamkniętymi jest zamknięty cbdo.

Twierdzenie 8. Każda dekompozycja π_{max}^{dKP} jest zbiorem klas prostych wraz z ich prostymi zamkniętymi łańcuchami.

Dowód wynika bezpośrednio z df. 44 i z tw. 7.

Poniżej podane są warunki ułatwiające wybór takich klas prostych i łańcuchów do nich aby otrzymać dekompozycję maksymalną.

Definicja 45. Każdy łańcuch dla klasy prostej C_α nazywać będziemy nadmiarowym, jeśli nie może on być członem dekompozycji maksymalnej π_{max}^{dKP}:

$$C_\alpha^{\mathcal{L}} = C_{\alpha\,\mathrm{nadm}}^{\mathcal{L}} \Leftrightarrow \neg \bigvee \pi_{max}^{dKP} \quad C_\alpha^{\mathcal{L}} \in \pi_{max}^{dKP}.$$

Lemat. Każdy łańcuch jest nadmiarowym łańcuchem, jeśli jego długość jest większa od $\min(n, m)-1$, gdzie n — ilość stanów w zbiorze S, a m — ilość maksymalnych podzbiorów niesprzecznych.

Dowód jest oczywisty [7].

Wartość $\min[(n, m)]-1]$ nazywana będzie górnym ograniczeniem.

Twierdzenie 9. Klasa prosta inicjalna $C_{\alpha in}$ jest nadmiarową klasą prostą, jeśli jej wszystkie elementy są pokryte przez każdy z jej prostych zamkniętych łańcuchów:

$$C_{\alpha in} = C_{\alpha nadm} \Leftrightarrow \bigwedge i \in I \bigvee j \in I' \quad s \in C_{\alpha in} \Rightarrow s \in C_j \wedge C_j \in C_{\alpha ini}^{\mathcal{L}}.$$

Dowód. Załóżmy, że $C_{\alpha in} \in \pi_{max}^{dKP}$.

Wówczas $\bigvee i \in I \quad C_{\alpha in} \in \pi_{max}^{dKP} \Rightarrow C_{\alpha ini}^{\mathcal{L}} \in \pi_{max}^{dKP}$. Niech $C_{\alpha ini}^{\mathcal{L}}$ pokrywa wszystkie elementy $C_{\alpha in}$.

Również $M_{\alpha in} = \emptyset$, a więc usunięcie $C_{\alpha in}$ z π_{max}^{dKP} nie narusza warunków pokrycia i zamkniętości i $C_{\alpha in}$ jest nadmiarową klasą prostą, cbdo.

Przykład. W zbiorze klas prostych automatu GL według tw. 9 nie ma klas prostych nadmiarowych.

Twierdzenie 10. Klasa prosta C_α jest nadmiarową klasą prostą, jeśli istnieje klasa prosta C_β taka, że $C_\alpha \supset C_\beta$, $M_\alpha = M_\beta$, $N_\alpha \supseteq N_\beta$ i elementy $s \in (C_\alpha - C_\beta)$ są pokryte przez każdy z prostych zamkniętych łańcuchów klasy prostej C_β:

$$C_\alpha = C_{\alpha\,\mathrm{nadm}} \Leftrightarrow \bigvee \beta \in I \quad C_\alpha \supset C_\beta \wedge M_\alpha = M_\beta \wedge N_\alpha \supset N_\beta \wedge$$
$$\wedge \bigwedge i \in I' \bigvee j \in I \quad s \in (C_\alpha - C_\beta) \Rightarrow s \in C_j \wedge C_j \in C_{\beta i}^{\mathcal{L}}.$$

Dowód. Warunek $M_\alpha = M_\beta$ i $N_\alpha \supseteq N_\beta$ zapewniają, że po usunięciu C_α z π_{max}^{dKP} warunki zamknięcia nie zmienią się.

Wszystkie elementy C_α pokryte są albo przez C_β, albo przez każdy łańcuch do C_β, cbdo.

Przykład. Spośród klas prostych automatu GL wybierzmy

1) $C_\alpha = \{b, c, d\}$ i $C_\beta = \{c, d\}$
$\{b, c, d\} \supset \{c, d\}$
2) $M_\alpha = \{\{c, f, g\}, \{c, f\}, \{c, g\}\}$
$M_\beta = \{\{c, f, g\}, \{c, f\}, \{c, g\}\}$
$M_\alpha = M_\beta$
3) $N_\alpha = \{\{a, b, d, e\}, \{a, g\}, \{d, e, h\}\}$
$N_\beta = \{\{a, g\}, \{a, b, d, e\}, \{d, e, h\}\}$
$N_\alpha \supseteq N_\beta$
4) $C_{\beta_1}^{L} = \{\{a, g\}, \{a, b, d, e\}\}$
$C_{\beta_2}^{L} = \{\{a, g\}, \{d, e, h\}, \{a, b, d, e\}\}$
stan b jest pokryty przez $C_{\beta_1}^{L}$ i $C_{\beta_2}^{L}$
$C_\alpha = \{b, c, d\}$ jest klasą prostą nadmiarową.

Łatwo można sprawdzić, że również klasa prosta $\{c, f, g\}$ jest klasą prostą nadmiarową.

Uwaga 1. Po zastosowaniu tw. 10 i usunięciu klas prostych nadmiarowych może się okazać, że niektóre klasy proste nie będące dotychczas nadmiarowymi, stają się w myśl tw. 9 nadmiarowymi.

Przykład. Po usunięciu ze zbioru klas prostych $\{b, c, d\}$ i $\{c, f, g\}$ można zastosować tw. 9 i usunąć klasę prostą $\{c, g\}$. Sprawdźmy to:

1) $\{c, g\}$ jest klasą prostą inicjalną,
2) $\{cg\}_1^{L} = \{\{c, d\}, \{a, g\}, \{a, b, d, e\}, \{f, g\}, \{d, e\}\}$ i stany c i g są pokryte przez jedyny teraz zamknięty łańcuch prosty do klasy $\{c, g\}$; $\{c, g\}$ jest klasą prostą nadmiarową. Również nadmiarową staje się klasa prosta $\{c, f\}$.

Uwaga 2. Z pewną stratą co do ogólności rozwiązania problemu minimalizacji — mianowicie w przypadku gdy celem jest znalezienie nie wszystkich a tylko niektórych dekompozycji maksymalnych π_{max}^{dKP} — można złagodzić założenia twierdzeń 9 i 10. Otrzymamy wówczas tw. 9a i 10a.

Twierdzenie 9a. Klasa prosta inicjalna $C_{\alpha in}$ jest nadmiarową klasą prostą, jeśli jej elementy są pokryte przez implikowane przez nią klasy proste:

$$C_{\alpha in} = C_{\alpha nadm} \Leftrightarrow \bigvee i \in I \quad s \in C_{\alpha in} \Rightarrow s \in C_i \wedge C_i \in N_\alpha^*.$$

Dowód. Załóżmy, że jeśli $C_{\alpha in} \in \pi_{max}^{dKP}$, to każda $C_i \in N_\alpha^*$ też jest elementem π_{max}^{dKP}. Wtedy usunięcie $C_{\alpha in}$ z π_{max}^{dKP} nie zmieni warunków pokrycia i zamkniętości. cbdo.

Przykład. Wśród klas prostych automatu GL klasą nadmiarową (według tw. 9a) jest tylko $\{c, g\}$.

Twierdzenie 10a. Klasa prosta C_α jest nadmiarową klasą prostą, jeśli istnieje klasa prosta C_β taka, że $C_\alpha \supset C_\beta$, $M_\alpha = M_\beta$, $N_\alpha \supseteq N_\beta$ i elementy $s \in (C_\alpha - C_\beta)$ są pokryte przez N_β^*:

$$C_\alpha = C_{\alpha nadm} \Leftrightarrow \bigvee \beta \in I \quad C_\alpha \supset C_\beta \wedge M_\alpha = M_\beta \wedge N_\alpha \supseteq N_\beta$$
$$\wedge \bigvee j \in I \quad s \in (C_\alpha - C_\beta) \Rightarrow s \in C_j \wedge C_j \in N_{\beta_1}^*.$$

Dowód. Warunki $M_\alpha = M_\beta$ i $N_\alpha \supseteq N_\beta$ zapewniają, że po usunięciu C_α warunki zamknięcia nie zmienią się. Elementy C_α są pokryte przez C_i albo N_i^*, cbdo.

Przykład. W zbiorze klas prostych automatu GL nadmiarowymi klasami prostymi (według tw. 10a) są $\{b, c, d\}$, $\{c, f, g\}$, $\{c, f\}$.

W wielu przypadkach optymalnym rozwiązaniem problemu minimalizacji jest wyznaczenie takiej dekompozycji maksymalnej, że w tablicy wyjść odpowiadającego jej automatu będzie jak najwięcej „—", tzn. takich przypadków, że $\bigvee x \in X \bigvee s \in S\ s\lambda_x = \emptyset$. Zbiór takich przypadków dla danego automatu $\hat{A}$ oznaczamy przez S_k; $S_k = \{\langle s, x\rangle | s\lambda_x = \emptyset\}$.

Dokompozycja maksymalna taka, że w odpowiadającym jej automacie $\operatorname{card} S_k = \max$ nazywana będzie dekompozycją maksymalną wyjściowo-optymalną i oznaczana będzie $\pi^d_{max(-)}$. Wyznaczanie rodziny dekompozycji $\{\pi^d_{max(-)}\}_p$ w ogólnym przypadku możliwe jest tylko wtedy, gdy — jak to już było zaznaczone wcześniej — wyboru dokonywać będziemy ze wszystkich podzbiorów niesprzecznych. Można podać metodę pozwalającą na otrzymywanie większej ilości rozwiązań niż w przypadku wyboru z klas prostych, ale mniej niż w przypadku wyboru z podzbiorów niesprzecznych. Dzięki tej metodzie w przypadkach praktycznych otrzymuje się $\pi^d_{max(-)}$, albo wyniki zbliżone do $\pi^d_{max(-)}$.

Poniżej podana będzie metoda wyznaczania dla danej dekompozycji π^{dKP}_{max}, rodziny dekompozycji takich, których członami będą pewne podzbiory o właściwościach podanych niżej.

Definicja 46. Podzbiór S^n_α klasy prostej C_α nazywa się klasą prostą dodatkową do klasy prostej C_α wtedy i tylko wtedy, gdy

$$S^n_\alpha \subseteq C_\alpha \wedge P_{S^n_\alpha} = P_\alpha.$$

Klasę prostą dodatkową do klasy prostej C_α oznaczać będziemy $C_{\alpha d}$

$$S^n_\alpha = C_{\alpha d} \Leftrightarrow S^n_\alpha \subseteq C_\alpha \wedge P_{S^n_\alpha} = P_\alpha.$$

Z df. 46 wynika, że zbiór wszystkich klas prostych wraz z klasami prostymi dodatkowymi do nich spełnia warunki zamkniętości i pokrywa zbiór stanów S.

Twierdzenie 11. Jeśli dana jest $\pi^{dKP}_{max} = \{C_1, C_2, \ldots, C_k\}$, to $\pi^{dKP}_{max(-)} = \{C_{1d}, C_{2d}, \ldots, \ldots, C_{kd}\}$ jest taką dekompozycją, której członami są klasy proste dodatkowe i $\bigwedge i \leqslant k$ $\operatorname{card} C_{id} = \min$ (liczność każdej klasy prostej dodatkowej jest jak najmniejsza).

Dowód. $\pi^{dKP}_{max(-)}$ jest taką dekompozycją, w której $\operatorname{card} S_k = \max$. Jeśli dana jest $\pi^{dKP}_{max} = \{C_1, C_2, \ldots, C_k\}$ i jeśli C_i zastąpimy C_{id}, to ilość kresek w tablicy wyjść albo pozostanie taka sama albo wzrośnie. Wynika to bezpośrednio z warunku $C_i \supset C_{id}$.

Gdy każdą C_i zastąpimy C_{id} (oczywiście każde takie zastąpienie nie może naruszyć warunków pokrycia) i gdy $\bigwedge i \leqslant k$ card $C_{id} = \min$, to dla tak utworzonego automatu $\operatorname{card} S_k = \max$, cbdo.

Wszystkie klasy proste dodatkowe wyznaczyć można bezpośrednio ze zbioru wszystkich podzbiorów niesprzecznych. Można również wyznaczyć klasy proste dodatkowe ze zbioru klas prostych. Możliwe jest przy tym także podanie kilku warunków ułatwiających znalezienie $\pi^{dKP}_{max(-)}$.

Twierdzenie 12. Klasa prosta nadmiarowa w zbiorze wszystkich klas prostych jest również nadmiarową w zbiorze wszystkich klas prostych wraz z ich dodatkowymi klasami prostymi.

Dowód. Wprowadzenie do zbioru klas prostych klas prostych dodatkowych nie zmienia warunków tw. 9 i 10.

Twierdzenie 13. Jeśli klasa prosta jest nadmiarową w zbiorze wszystkich klas prostych, to dodatkowe do niej klasy proste również są nadmiarowe w zbiorze wszystkich klas prostych wraz z dodatkowymi do nich.

Dowód. Należy udowodnić $C_{\alpha n} \notin \pi_{max}^{dKP} \Rightarrow C_{\alpha nd} \notin \pi_{max}^{dKP}$.
Rzeczywiście:

a) $C_\alpha \supset C_{\alpha d}$,

b) $P_\alpha = P_{\alpha d} \Rightarrow N_\alpha^* = N_{\alpha d}^*$,
$C_\alpha \supset C_{\alpha d} \Rightarrow M_\alpha \subseteq M_{\alpha d}$

i jeśli C_α jest usunięta z π_{max}^d, to i usunięcie każdej $C_{\alpha d}$ nie zmieni warunków pokrycia i zamkniętości. A więc $C_{\alpha d}$ są nadmiarowe, cbdo.

5.2. Algorytmy

Opierając się na podanych w p. 5.1 definicjach i twierdzeniach, podane są poniżej algorytmy poszukiwania dekompozycji maksymalnych.

Algorytm 1. Wyznaczenie dla danego automatu $\hat{A}$ pełnej rodziny dekompozycji maksymalnych $\{\pi_{max}^d\}_p$.

Wersja 1.

1. Wyznaczyć wszystkie pary stanów niesprzecznych.
2. Wyznaczyć wszystkie podzbiory niesprzeczne.
3. Dla każdego podzbioru niesprzecznego wyznaczyć łańcuchy proste.
4. Dla każdego stanu $s \in S^A$ wyznaczyć zbiór podzbiorów niesprzecznych, w których ten stan jest zawarty.
5. Pokryć wszystkie stany zbioru S^A podzbiorami niesprzecznymi wraz z ich łańcuchami otrzymanymi w p. 3.
6. Z otrzymanych w p. 5 „ciągów" podzbiorów niesprzecznych wraz z ich łańcuchami wybrać ciągi najkrótsze.

Wersja 2.

1. Wyznaczyć wszystkie pary stanów niesprzecznych.
2. Wyznaczyć wszystkie podzbiory niesprzeczne.
3. Wprowadzić zmienną k ($k \in N$ i zmienia się od 1 do n, gdzie $n = \text{card}\, S$).
4. Ustawić k i sprawdzić czy k zbiorów niesprzecznych pokrywa zbiór S.
5. Jeśli p. 4 jest spełniony, to sprawdzić czy wybrane zbiory niesprzeczne spełniają warunek zamkniętości.
6. Jeśli p. 4 nie jest spełniony, to wykonywać dalej p. 4, wybierając wszystkie kombinacje k zbiorów niesprzecznych oraz zmieniając wartość k.
7. Jeśli p. 5 jest spełniony, to k zbiorów niesprzecznych tworzy dekompozycję.
8. Jeśli p. 5 nie jest spełniony, to wykonywać dalej p. 4.
9. Jeśli $k = n$, to koniec.

Algorytm 2. Wyznaczenie dla danego automatu $\hat{A}$ pełnej rodziny dekompozycji maksymalnych wyznaczonej na podstawie klas prostych $\{\pi_{max}^{dKP}\}_p$.

1. Wyznaczyć wszystkie pary stanów niesprzecznych.
2. Wyznaczyć wszystkie podzbiory niesprzeczne.

3. Wyznaczyć klasy proste.
4. Eliminować klasy proste nadmiarowe (tw. 9 i tw. 10).
5. Dla każdej klasy prostej otrzymanej po wykonaniu p. 4 utworzyć łańcuchy proste.
6. Eliminować łańcuchy proste nadmiarowe.
7. Dla każdego stanu $s \in S^A$ wyznaczyć zbiór klas prostych (wraz z ich łańcuchami prostymi), w których ten stan jest zawarty.
8. W zbiorach otrzymanych w p. 7 (dla każdego oddzielnie) eliminować te klasy proste wraz z ich łańcuchami, które zawierają właściwie inne klasy proste wraz z ich łańcuchami.
9. Pokryć wszystkie stany zbioru S^A klasami prostymi wraz z ich łańcuchami otrzymanymi w p. 6.
10. Wyznaczyć najkrótsze „ciągi" otrzymane w p. 8.

Algorytm 3. Wyznaczenie dla każdego automatu $\hat{A}$ co najmniej jednej dekompozycji maksymalnej π^d_{max}.

1. Wyznaczyć wszystkie pary stanów niesprzecznych.
2. Wyznaczyć wszystkie podzbiory niesprzeczne.
3. Wyznaczyć klasy proste.
4. Eliminować klasy proste nadmiarowe (tw. 9a i 10a).
5. Dla każdej klasy prostej C_α otrzymanej po wykonaniu p. 4 utworzyć N^*_α.
6. Dla każdego stanu $s \in S^A$ wyznaczyć zbiór klas prostych (wraz z ich zbiorami implikowanymi N^*), w których ten stan jest zawarty.
7. W zbiorach otrzymanych w p. 6 (dla każdego oddzielnie) eliminować te klasy proste wraz z ich zbiorami implikowanymi N^*, które zawierają właściwie inne klasy proste wraz z ich zbiorami implikowanymi.
8. Pokryć wszystkie stany zbioru S^A klasami prostymi wraz z ich zbiorami implikowanymi wyznaczonymi w p. 5.
9. Wyznaczyć najkrótsze „ciągi" otrzymane w p. 7.

Algorytm 4. Wyznaczenie dla danego automatu A pełnej rodziny dekompozycji maksymalnych wyjściowo-optymalnych wyznaczonej na podstawie klas prostych dodatkowych $\{\pi^{dKP\,D}_{max(-)}\}$.

1. Wyznaczyć wszystkie pary stanów niesprzecznych.
2. Wyznaczyć wszystkie podzbiory niesprzeczne.
3. Wyznaczyć klasy proste.
4. Eliminować klasy proste nadmiarowe (tw. 9 i tw. 10).
5. Dla każdej klasy prostej otrzymanej po wykonaniu p. 4 utworzyć łańcuchy proste.
6. Eliminować łańcuchy proste nadmiarowe.
7. Dla każdego stanu $s \in S^A$ wyznaczyć zbiór klas prostych (wraz z ich łańcuchami prostymi), w których ten stan jest zawarty.
8. W zbiorach otrzymanych w p. 7 dla każdego oddzielnie eliminować te klasy proste wraz z ich łańcuchami, które zawierają właściwie inne klasy proste wraz z ich łańcuchami.
9. Pokryć wszystkie stany zbioru S^A klasami prostymi wraz z ich łańcuchami otrzymanymi w p. 8.

10. Wybrać najkrótsze „ciągi" klas prostych wraz z ich łańcuchami prostymi otrzymane w p. 9. Te najkrótsze „ciągi" są dekompozycjami maksymalnymi π_{max}^{dKP}.
11. Dla każdej klasy prostej członu π_{max}^{dKP} wyznaczyć klasy proste dodatkowe i wykonać ponownie punkty $5 \div 10$.
12. Wśród π_{max}^{dKPD} wyznaczonych w p. 11 wybrać takie, które zawierają jak najmniej stanów $s \in S^A$. Wynikiem są dekompozycje maksymalne wyjściowo-optymalne $\pi_{max(-)}^{dKPD}$.

P r z y k ł a d. Zastosujmy wyżej przedstawiony algorytm do minimalizacji liczby stanów automatu *GL*. Pierwsze cztery czynności zostały już poprzednio wykonane, załóżmy więc, że dane są klasy proste nienadmiarowe:

$\{a, b, d, e\}$
$\{d, e, h\}$
$\{b, c\}$
$\{c, d\}$
$\{f, g\}$
$\{d, h\}$
$\{a, g\}$
$\{f\}$

Wykonajmy punkt 5:

Klasy proste	Łańcuchy proste
$\{a, b, d, e\}$	Ø
$\{d, e, h\}$	$\{a, b, d, e\}$
$\{b, c\}$	Ø
$\{c, d\}$	$\{\{a, g\}, \{d, e, h\}, \{a, b, d, e\}\}$
	$\{\{a, g\}, \{a, b, d, e\}\}$
$\{f, g\}$	$\{\{d, e, h\}, \{a, b, d, e\}\}$
$\{d, h\}$	Ø
$\{a, g\}$	Ø
$\{f\}$	Ø

Wykonajmy p. 6: tylko dla klasy prostej $\{c, d\}$ eliminować można łańcuchy proste $\{\{a, g\}, \{d, e, h\}, \{a, b, d, e\}\}$.

Wykonajmy p. 7:

stan	klasy proste (wraz z ich łańcuchami) pokrywające ten stan (łańcuchy proste nie są tu oznaczone)
a	$\{a, b, d, e\}$
	$\{a, g\}$
b	$\{a, b, d, e\}$
	$\{b, c\}$
c	$\{c, d\} \rightarrow \{\{a, g\}, \{a, b, d, e\}\}$
	$\{b, c\}$
d	$\{a, b, d, e\}$
	$\{d, e, h\} \rightarrow \{a, b, d, e\}$
	$\{c, d\} \rightarrow \{\{a, g\}, \{a, b, d, e\}\}$
	$\{d, h\}$

Stan	klasy proste (wraz z ich łańcuchami) pokrywające ten stan (łańcuchy proste nie są tu oznaczone)
e	$\{a, b, d, e\}$
	$\{d, e, h\} \rightarrow \{a, b, d, e\}$
f	$\{f, g\} \rightarrow \{\{d, e, h\}, \{a, b, d, e\}\}$
	$\{f\}$
g	$\{a, g\}$
	$\{f, g\} \rightarrow \{\{d, e, h\}, \{a, b, d, e\}\}$
h	$\{d, e, h\} \rightarrow \{a, b, d, e\}$
	$\{d, h\}$

Wykonajmy punkt 8: wyeliminowane zostają następujące klasy proste wraz z ich łańcuchami:

dla stanu d $\{d, e, h\} \rightarrow \{a, b, d, e\}$

$\{c, d\} \rightarrow \{\{a, g\}, \{a, b, d, e\}\}$

dla stanu e $\{d, e, h\} \rightarrow \{a, b, d, e\}$

Wykonajmy punkt 9:

$$[\{a, b, d, e\} \vee \{a, g\}]\ [\{a, b, d, e\} \vee \{b, c\}]\ [\{c, d\}\ \{a, g\}\ \{a, b, d, e\} \vee \{b, c\}]$$
$$[\{a, b, d, e\} \vee \{d, h\}]\ [\{a, b, d, e\}]\ [\{f, g\}\ \{d, e, h\}\ \{a, b, d, e\} \vee \{f\}]$$
$$[\{a, g\} \vee \{f, g\}\ \{d, e, h\}\ \{a, b, d, e\}]\ [\{d, e, h\}\ \{a, b, d, e\} \vee \{d, h\}]$$

Uwaga. Jeśli jakiś stan zawarty jest w kilku klasach prostych, to pokrywamy go alternatywnie wszystkimi tymi klasami prostymi.

Wykonajmy punkt 10: istnieje tylko jedna

$$\pi_{max}^{dKP} = \{\{a, b, d, e\}, \{b, c\}, \{f, g\}, \{d, e, h\}\}.$$

Wykonajmy punkt 11. klasami prostymi dodatkowymi są:

dla klasy prostej	$\{a, b, d, e\}$	$\{a\}$
		$\{b\}$
		$\{d\}$
		$\{e\}$
		$\{d, e\}$
„ „ „	$\{b, c\}$	$\{b\}$
		$\{c\}$
„ „ „	$\{f, g\}$	$\emptyset$
„ „ „	$\{d, e, h\}$	$\{e, h\}$.

Wykonajmy teraz operacje 5÷10 na zbiorze:

$\{a, b, d, e\}$
$\{d, e, h\}$
$\{b, c\}$
$\{f, g\}$
$\{e, h\}$
$\{d, e\}$
$\{a\}$
$\{b\}$
$\{c\}$
$\{d\}$
$\{e\}$

Wyniki:

$$\pi_{max_1}^{dKPD} = \{\{a, b, d, e\}, \{b, c\}, \{f, g\}, \{d, e, h\}\},$$

$$\pi_{max_2}^{dKPD} = \{\{a, b, d, e\}, \{c\}, \{f, g\}, \{d, e, h\}\},$$

$$\pi_{max_3}^{dKPD} = \{\{a, b, d, e\} \{b, c\}, \{f, g\}, \{e, h\}\},$$

$$\pi_{max_4}^{dKPD} = \{\{a, b, d, e\}, \{c\}, \{f, g\}, \{e, h\}\}.$$

Wybieramy $\pi_{max(-)}^{dKPD}$

$$\pi_{max(-)}^{dKPD} = \{\{a, b, d, e\}, \{c\}, \{f, g\}, \{e, h\}\}.$$

Algorytm 5. Wyznaczenie dla danego automatu zupełnego $\hat{A}$ podziału maksymalnego π_{max}.

1. Wyznaczyć wszystkie pary stanów niesprzecznych.
2. Wyznaczyć wszystkie podzbiory niesprzeczne.
3. Wyznaczyć maksymalne podzbiory niesprzeczne. Rodzina tych podzbiorów tworzy podział maksymalny π_{max}.

5.3. Dyskusja

A. Nie jest wcale oczywiste, że automatowi o najmniejszej liczbie stanów odpowiada najprostszy układ logiczny.

W celu porównywania złożoności układów logicznych należałoby wprowadzić pewne kryteria, na podstawie których każdemu układowi logicznemu realizowanemu przyporządkowywałoby się jakiś współczynnik. Najbardziej ogólnym kryterium, ale i być może najistotniejszym jest cena realizacji. W tym przypadku uchwycenie, który układ logiczny jest tańszy od innego jest sprawą żmudnych obliczeń. Przy korzystaniu z tego kryterium może się okazać, że układ o możliwie najmniejszej liczbie jednostek pamięci może mieć tak rozbudowane funkcje wzbudzeń, że będzie on droższy od układu o większej liczbie jednostek pamięci i o prostszych funkcjach wzbudzeń. Przy innych kryteriach, np. ilości połączeń między układami elementarnymi (takie kryterium jest bardzo istotne przy realizacji układów logicznych z zastosowaniem układów scalonych), może wystąpić ten sam problem.

W związku z powyższym może nasunąć się pytanie: czy konieczna jest minimalizacja liczby stanów automatów skoro nie zawsze doprowadza ona do optymalnego (np. najtańszego) układu logicznego. Odpowiedź na to pytanie jest natychmiastowa: znalezienie w ogólnym przypadku „optimum" będzie możliwe jedynie przy zastosowaniu maszynowej syntezy układów logicznych. Jednym z etapów tej syntezy będzie napewno minimalizacja liczby stanów automatów. Oczywiście etap ten w powiązaniu z innymi będzie sterowany tak, że jeśli konieczne będzie przebadanie automatów nie tylko minimalnych, to w wyniku etapu minimalizacji liczby stanów otrzymamy automaty o $\operatorname{card} S = S_{min} + k$ (gdzie $k = 1, \ldots, n$). W związku z tym rozwiązanie maszynowych metod minimalizacji można traktować jako rozwiązanie jednego z etapów realizacji maszynowej syntezy układów logicznych.

B. Załóżmy, że tylko wśród automatów minimalnych znajdują się automaty optymalne, tj. takie, których realizacja techniczna przy określonym kryterium (np. cena, ilość układów scalonych, ilość połączeń między nimi) jest najprostsza.

Wtedy $\{\hat{A}_{opt}\}_p \subseteq \{\hat{A}_{min}\}_p$ i przy wyznaczeniu $\{\hat{A}_{opt}\}_p$ należałoby brać pod uwagę każdy z automatów osiągniętych w wyniku zastosowania algorytmu 1. W przeważającej liczbie przypadków można założyć, że $\hat{A}_{opt\,i} \in \{\hat{A}_{min}^{KPD}\}_p$. Ponieważ $\{\hat{A}_{min}^{KPD}\}_p \subseteq \{\hat{A}_{min}\}_p$, to wybór $\hat{A}_{opt\,i}$ z $\{\hat{A}_{min}^{KPD}\}_p$ może okazać się łatwiejszym niż poszukiwanie $\hat{A}_{opt\,i}$ z $\{\hat{A}_{min}\}_p$.

Tylko w nielicznych przypadkach (specjalnie w tym celu konstruowanych przez autora) okazało się, że $\hat{A}_{opt\,i} \notin \{\hat{A}_{min}^{KPD}\}_p$. W tych samych przypadkach okazywało się, że automaty wyszukiwane tylko z $\{\hat{A}_{min}^{KPD}\}$ różniły się niewiele od $\hat{A}_{opt\,i} \in \{\hat{A}_{opt}\}_p$. Biorąc to pod uwagę wydaje się, że stosowanie algorytmów 2 lub 4 jest celowe. Wyniki algorytmu 4 pokrywają wyniki algorytmu 2. Niekiedy potrzebnym wynikiem minimalizacji jest osiągnięcie co najmniej jednego automatu minimalnego. Wówczas należy stosować algorytm 3.

Wyznaczenie $\pi_{max(-)}^{dKPD}$ nie zawsze oznacza, że automat odpowiadający tej dekompozycji jest optymalnym. Optymalne będą funkcje wyjściowe, ale funkcje wzbudzeń układów pamięci mogą być (łatwo jest skonstruować takie przypadki np. [10]) nie najprostsze. Dlatego też należy szukać „optimum" wśród $\{\pi_{max}^{dKPD}\}$, a nie tylko wśród $\{\pi_{max(-)}^{dKPD}\}$. Oczywiście $\{\pi_{max(-)}^{dKPD}\}_p \subseteq \{\pi_{max}^{dKPD}\}_p$.

LITERATURA CYTOWANA W TEKŚCIE

1. B. M. Constantini, Un algoritmo programmato per la minimizzazione degli stati nelle machine sequenziali incompletamente specificate, Note Recensioni e Notice, Marzo-Aprile 1971 (307—321).
2. S. C. de Sarker, A. K. Basu, Choudhury, Simplification of Incompletly Specified Flow Table with the Help of Primo Closed Sets. IEEE Transactions on Computers, October 1969.
3. A. Ginzburg, Algebraic Theory of Automata, Academic Press, New York — London 1968.
4. A. Grasselli, F. Luccio, A Method Minimizing the Number of Internal State in Incompletly Specified Sequential Networks, IEEE on EC, June 1965 (350).
5. J. Kella, State Minimization of Incompletly Specified Sequential Machines, IEEE Transations on Computers, Vol. C. 19, No. 4, April 1970.
6. W. G. Lazariew, F. U. Pijl, Sintez uprawlajuszczych awtomatow, Energia, Moskwa 1970.
7. W. Majewski, Zagadnienie minimalizacji liczby stanów automatów skończonych, Rozprawy Elektrotechniczne, t. XV, z. 2, 1968.
8. W. Majewski, Układy logiczne, Warszawa 1972.
9. W. S. Meisel, A Note on Internal State Minimization in Incompletly Specified Sequential Networks, IEEE Trans. on EC., August 1967 (p. 508).
10. R. E. Miller, Switching Theory, Vol. 1 — Combinational Circuits, Vol. 2 — Sequential Circuits and Machines, John Wiley and Sons, New York 1965 (także w tłumaczeniu rosyjskim).
11. M. C. Paul, S. H. Unger, Minimizing the Number of States in Incompletly Specified Sequential Switching Functions, IEEE Trans. Electron. Comp., Sept. 1959 (356—367).
12. A. T. Pu, Generalized Decomposition on Incomplete Finite Automata, Information and Control, 13, 1968 (1—19).
13. W. Traczyk, Syntezy automatów asynchronicznych, Wyd. Politechniki Warszawskiej, Warszawa 1969.
14. Tse-Yun-Feng. On Minimal Covering of Incompletly Specified Sequential Network, Journal of the Franklin Institute, Vol. 292, No. 2, August 1971.
15. R. Yeh, On Relation Homomorphims of Automata Information and Control, 13, 1968 (140—155).

A. ALBICKI

STATE MINIMIZATION OF AUTOMATA

Summary

This paper presents one of the steps of sequential circuits synthesis referred to as state minimization of automata. There are given fundamental definitions of automata theory and some relations are defined on set of states. In particular there are discussed cover relation and weak homorphism relation. These relations are useful to build the generalized algorithm of state minimization.

In further parts there are given theorems and definitions used for practical solutions of state minimization illustrated by an example. There are given 5 algorithms which are the base for computer programs.

In the final part results of this paper are discussed.

A. ALBICKI

LA MINIMALISATION DE LA QUANTITÉ DES ÉTATS DE L'AUTOMATE

Résumé

L'article présente une des étapes de la synthèse des raisons logiques — la minimalisation de la quantité des états de l'automate. Sont présentées les définitions fondamentales de la théorie des automates, ainsi que quelques unions définies par l'ensemble des états des automates. Particulièrement sont pris en considération: la rélation de la counverture de l'automate et la rélation du faible homomorphisme. Ces rélations aident à construire un algorytme général pour la minimalisation de la quantité des états.

Dans la suite l'article présente les théorèmes et définitions qui peuvent etre utilisés pour la solution pratique du problème de la minimalisation des automates incomplets. Le procès de la minimalisation est illustré par un exemple. Sont donnés 5 algorytmes à l'aide desquels on peut réaliser avec un computer les programmes concrèts.

A la fin on a présenté une discussion sur les résultats obtenus.

A. ALBICKI

REDUKTION DER ZAHL INNERER AUTOMATENZUSTÄNDE

Zusammenfassung

In diesem Artikel wird eine von den Etappen der Synthese logischer Systeme — die Zustandsreduktion für Automaten — dargestellt. Es werden Grunddefinitionen der Automatentheorie und manche Verbindungen, die auf der Automatensammlung bestimmt werden, angeführt. Insbesondere sind Relationen der Automatendeckung und des schwachen Homomorphismus untersucht worden. Mit Hilfe dieser Relationen ist ein allgemeiner Algorithmus für die Zustandsreduktion erbaut worden.

Im weiteren Teil sind auch Theoreme und Definitionen vorgetragen worden, die bei der praktischen Lösung des Reduktionsproblems unvollständig bestimmter Automaten benutzt werden können. Anhand eines Beispiels wird der Reduktionsprozeß gezeigt. 5 Algorithmen sind gegeben, womit konkrete Programme für Computer bearbeitet werden können.

Zum Schluß werden die erreichten Resultate diskutiert.

A. АЛЬБИЦКИ

МИНИМАЛИЗАЦИЯ ЧИСЛА ВНУТРЕННИХ СОСТОЯНИЙ АВТОМАТОВ

Резюме

В настоящей статье представляется один из этапов синтеза логических схем — минимализация числа внутренних состояний автоматов.

Представлены основные понятия теорий автоматов, а также некоторые отношения определенные на множестве состояний автоматов. В частности рассматриваются: отношение покрытия и отношение слабого гомоморфизма. При помощи этих отношений разработан основной алгоритм минимализации числа состояний. Далее представлены дефиниции и теоремы применяемые при практической разработке проблемы минимализации не вполне определенных автоматов. Сам процесс минимализации иллюстрируется в приведенном примере. Приведены также 5 алгоритмов, при использовании которых можно разработать конкретные машинные программы.

В заключение дискутируются результаты полученные в этой работе.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 2, ss. 389—396

621.396.962.23

Ocena wpływu sygnału zakłócającego na pomiar prędkości metodą dopplerowską

ANDRZEJ LIZOŃ (WARSZAWA)

Instytut Radioelektroniki Politechniki Warszawskiej

Otrzymano 20.4. 1973

W pracy omówiono zagadnienie pomiaru prędkości obiektu metodą dopplerowską przy jednoczesnym odbiorze sygnału pochodzącego od innego poruszającego się obiektu. Problem rozwiązano przeprowadzając analizę częstotliwości chwilowej wypadkowego sygnału odbieranego dla przypadku, gdy prędkości poszczególnych obiektów różnią się nieznacznie. Uzyskane wyniki wskazują, że przy długim czasie pomiaru częstotliwości dopplerowskiej w porównaniu z okresem wahań tej częstotliwości zmierzona wartość prędkości odpowiada prędkości obiektu, od którego sygnał odbierany ma większą amplitudę. Przy krótkim czasie pomiaru tej częstotliwości wyniki mogą wahać się począwszy od średniej arytmetycznej prędkości obu obiektów do wartości znacznie odchylającej się w kierunku przeciwnym od prędkości obiektu dającego sygnał o większej amplitudzie.

1. WSTĘP

Przy pomiarze prędkości za pomocą radaru dopplerowskiego o fali ciągłej powstaje zagadnienie rozróżniania poruszających się obiektów znajdujących się w obszarze obejmowanym przez charakterystykę kierunkową anteny radaru.

W przypadku trudności rozdzielenia sygnałów odbieranych pochodzących od tych obiektów istotna jest ocena wpływu sygnału zakłócającego na wynik pomiaru prędkości wybranego obiektu. Zagadnienie to poruszone zostało przez Koryu Ishii [1], nie zostało ono jednak rozwiązane poprawnie, wnioski natomiast i uogólnienia przedstawione zostały błędnie. Poniżej rozpatrzono to zagadnienie ponownie, rozwiązując je dla przypadku dwóch obiektów poruszających się jednocześnie, gdy różnica ich prędkości jest znacznie mniejsza od prędkości każdego z tych obiektów.

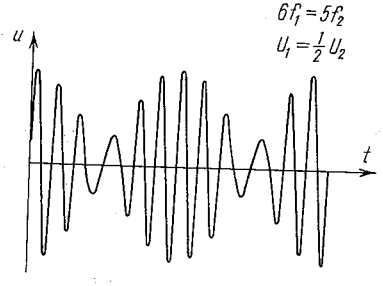
2. ZASADA POMIARU PRĘDKOŚCI

Prędkość poruszającego się obiektu określa się za pomocą radaru dopplerowskiego mierząc częstotliwość dopplerowską sygnału odebranego, odbitego lub retransmitowanego przez ten obiekt. Częstotliwość ta dla prędkości obiektu małych w porównaniu

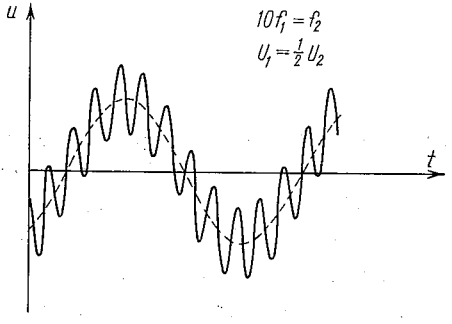
z prędkością rozchodzenia się fal elektromagnetycznych określona jest w przybliżeniu zależnością

$$f_D \cong \frac{2F_n}{C} v_r \tag{1}$$

gdzie:
F_n — częstotliwość sygnału wysyłanego z radaru,
C — prędkość rozchodzenia się fal elektromagnetycznych,
v_r — prędkość radialna obiektu względem radaru.
Dla sygnału sinusoidalnego pochodzącego od jednego obiektu zależność ta wiąże jednoznacznie prędkość radialną obiektu z częstotliwością dopplerowską sygnału odbieranego. Mierząc tę częstotliwość można określić szukaną wartość prędkości radialnej.



Rys. 1. Suma dwóch sinusoid o nieznacznie różniących się częstotliwościach



Rys. 2. Suma dwóch sinusoid o znacznie różniących się częstotliwościach

Jeśli w obszarze obejmowanym przez charakterystykę kierunkową anteny radaru poruszają się inne obiekty, wypadkowy sygnał odbierany jest sumą sygnałów sinusoidalnych pochodzących od poszczególnych obiektów. Wpływ sygnałów zakłócających na wynik pomiaru prędkości wybranego obiektu zależy nie tylko od ich amplitud i częstotliwości, ale także od sposobu pomiaru częstotliwości sygnału wypadkowego. Najczęściej pomiar ten odbywa się na podstawie pomiaru ilości przejść przez zero sygnału wypadkowego w określonym czasie. Otrzymany jednak wówczas wynik dla sygnałów składowych o nieznacznie różniących się częstotliwościach jest inny niż w przypadku, gdy różnica częstotliwości tych sygnałów jest porównywalna z częstotliwością jednego z sygnałów składowych.

Słuszność tego stwierdzenia wyjaśniają rysunki 1 i 2. Na rysunkach tych pokazane są sygnały wypadkowe będące wynikiem złożenia dwóch sygnałów składowych o częstotliwościach: w pierwszym przypadku $6f_1 = 5f_2$ oraz w drugim $10f_1 = f_2$. W obu przypadkach amplituda sygnału o częstotliwości mniejszej jest dwukrotnie większa od amplitudy sygnału o częstotliwości większej. W pierwszym przypadku sygnał wypadkowy można rozpatrywać jako sygnał wąskopasmowy o zmieniającej się w czasie częstotliwości. Można wykazać, że częstotliwość jego zmierzona na podstawie pomiaru ilości przejść przez zero sygnału wypadkowego w określonym przedziale czasu równa jest średniej wartości częstotliwości chwilowej tego sygnału wyznaczonej w tym samym przedziale czasu. W miarę skracania czasu pomiaru można zbliżyć się do pomiaru częstotliwości chwilowej. W drugim przypadku, gdy sygnału wypadkowego nie można traktować jako sygnału wąskopasmowego, ten sposób oceny jego częstotliwości jest niesłuszny.

Poniżej rozpatrzony zostanie dokładnie przypadek dwóch sygnałów składowych o nieznacznie różniących się częstotliwościach i różnych amplitudach. Przypadek dwóch sygnałów składowych jest najprostszym przypadkiem nadającym się do szczegółowej analizy. Różne amplitudy tych sygnałów przyjęto jako przypadek najistotniejszy w praktyce. Uzyskane wyniki wyjaśnią wpływ niepożądanego sygnału zakłócającego na wynik pomiaru prędkości. Wyników tych nie można jednak uogólniać na przypadek znacznie różniących się częstotliwości sygnałów składowych, ani na przypadek większej liczby sygnałów zakłócających.

3. CZĘSTOTLIWOŚĆ CHWILOWA SYGNAŁU WYPADKOWEGO

Sygnał wypadkowy, będący sumą dwóch przebiegów sinusoidalnych o różnych amplitudach i nieznacznie różniących się częstotliwościach

$$u_1 = U_1 \sin(\omega_1 t + \varphi_1), \tag{2}$$

$$u_2 = U_2 \sin(\omega_2 t + \varphi_2), \tag{3}$$

$$U_1 < U_2, \tag{4}$$

$$|\omega_1 - \omega_2| \ll \omega_1, \tag{5}$$

wygodnie przedstawić w postaci (por. Dodatek)

$$u = u_1 + u_2 = U \sin(\omega_2 t + \varphi), \tag{6}$$

gdzie:

$$tg\,\varphi = \frac{K \sin(\omega_r t + \varphi_1) + \sin\varphi_2}{K \cos(\omega_r t + \varphi_1) + \cos\varphi_2}, \tag{7}$$

$$K = \frac{U_1}{U_2} < 1, \tag{8}$$

$$\omega_r = \omega_1 - \omega_2 \tag{9}$$

Częstotliwość chwilowa przebiegu wypadkowego równa jest zatem

$$f = \frac{1}{2\pi}\left(\omega_2 + \frac{d\varphi}{dt}\right), \tag{10}$$

natomiast odchyłka częstotliwości chwilowej od częstotliwości sygnału o większej amplitudzie wynosi (por. Dodatek)

$$\Delta f = \frac{1}{2\pi}\frac{d\varphi}{dt} = \frac{\omega_r}{2\pi}\frac{K^2+K\cos(\omega_r t+\varphi_1-\varphi_2)}{1+K^2+2K\cos(\omega_r t+\varphi_1-\varphi_2)}. \tag{11}$$

Zależność (11) określa zmienność w czasie częstotliwości chwilowej sygnału wypadkowego. Z zależności tej wynika, że częstotliwość chwilowa zmienia się okresowo z okresem

$$T = \frac{2\pi}{|\omega_r|} = \frac{2\pi}{|\omega_1-\omega_2|}. \tag{12}$$

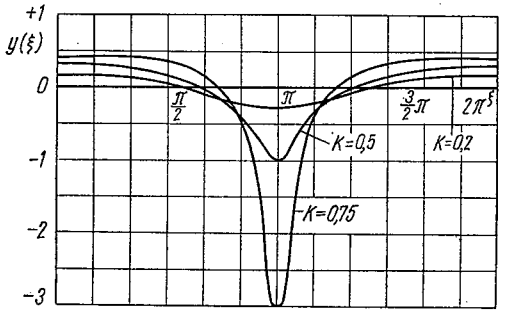
Zmiany te odbywają się wokół średniej wartości częstotliwości równej częstotliwości sygnału o większej amplitudzie $\frac{\omega_2}{2\pi}$, średnia bowiem wartość odchyłki częstotliwości Δf wyznaczona w okresie jej zmian równa jest zeru (por. Dodatek).

Chwilową wartość odchyłki częstotliwości można określić z zależności (11) pomijając w niej zależności fazowe. Można ją wówczas wyrazić jako

$$\Delta f = \frac{\omega_r}{2\pi}y(\xi), \tag{13}$$

gdzie

$$y(\xi) = \frac{K^2+K\cos\xi}{1+K^2+2K\cos\xi}. \tag{14}$$



Rys. 3. Wykres funkcji $y(\xi)$ w zakresie jednego okresu

Wykres funkcji $y(\xi)$ w zakresie jednego okresu dla różnych wartości współczynnika K pokazany jest na rys. 3. Maksymalna wartość tej funkcji wynosi

$$y_{\max} = \frac{K}{1+K} \tag{15}$$

i dla $0 < K < 1$ przyjmuje wartości dodatnie od zera do 0,5. Minimalna wartość tej funkcji równa jest

$$y_{\min} = -\frac{K}{1-K} \tag{16}$$

i przyjmuje wartości ujemne od zera do bardzo dużych w miarę jak K zbliża się do jedności.

Wartość odchyłki częstotliwości zależy zatem nie tylko od różnicy częstotliwości sygnałów składowych, ale także od stosunku amplitud tych sygnałów.

Jeśli amplitudy te różnią się znacznie ($K \ll 1$), wahania częstotliwości chwilowej sygnału wypadkowego są bardzo małe $\left(\Delta f_{\max} \ll \frac{\omega_r}{2\pi}\right)$ i symetryczne względem średniej wartości częstotliwości. Jeśli jednak amplitudy te są sobie bliskie ($K \leqslant 1$), wahania częstotliwości chwilowej są niesymetryczne względem średniej wartości częstotliwości; odchyłka częstotliwości może się wówczas zmieniać od połowy różnicy częstotliwości sygnałów składowych do wartości znacznie przekraczającej różnicę tych częstotliwości.

Asymetria zmian częstotliwości chwilowej wokół średniej wartości ma zawsze następującą własność. Mniejsze odchyłki częstotliwości występują w kierunku częstotliwości sygnału o mniejszej amplitudzie, większe natomiast w kierunku przeciwnym. Można to wykazać w sposób następujący. Jeśli $\omega_1 > \omega_2$, wówczas $\omega_r > 0$, częstotliwość chwilowa odchyla się mniej w kierunku do częstotliwości $\frac{\omega_1}{2\pi}$, tzn. w kierunku do częstotliwości sygnału o mniejszej amplitudzie. Jeśli zaś $\omega_1 < \omega_2$, wówczas $\omega_r < 0$, mniejsze odchyłki częstotliwości zachodzą nadal w kierunku do częstotliwości $\frac{\omega_1}{2\pi}$.

4. DYSKUSJA OTRZYMANYCH WYNIKÓW

Na podstawie oceny własności częstotliwości chwilowej sygnału wypadkowego można wysnuć następujące wnioski, dotyczące pomiaru prędkości obiektu metodą dopplerowską przy jednoczesnym odbiorze sygnału pochodzącego od innego poruszającego się obiektu, gdy prędkości obiektów różnią się nieznacznie.

1. Jeśli czas pomiaru częstotliwości dopplerowskiej sygnału wypadkowego jest duży w porównaniu z okresem zmian tej częstotliwości, zmierzona prędkość odpowiada zawsze prędkości obiektu, którego sygnał odbierany ma większą amplitudę.

2. Jeśli czas pomiaru częstotliwości dopplerowskiej sygnału wypadkowego jest mały w porównaniu z okresem zmian tej częstotliwości, zmierzona wartość prędkości oscyluje wokół wartości średniej, równej prędkości obiektu, którego sygnał odbierany ma większą amplitudę. Odchyłki od wartości średniej są proporcjonalne do różnicy prędkości poszczególnych obiektów, zależą jednak także od stosunku amplitud sygnałów odbieranych od tych obiektów. Przy znacznie różniących się amplitudach poszczególnych sygnałów odchyłki te są niewielkie i w przybliżeniu symetryczne względem średniej wartości zmierzonej prędkości. Przy porównywalnych amplitudach sygnałów składowych odchyłki te są niesymetryczne. Mniejsza wartość odchyłki występuje zawsze w kierunku wartości prędkości obiektu dającego sygnał o mniejszej amplitudzie i nie przekracza nigdy wartości połowy różnicy prędkości obu obiektów. Większa wartość odchyłki występuje zawsze w kierunku przeciwnym i może przyjmować wartości znacznie większe od różnicy prędkości tych obiektów. Zmierzona zatem wówczas wartość prędkości może wahać się począwszy od średniej arytmetycznej obu prędkości do wartości znacznie odchylającej się w przeciwnym kierunku od prędkości obiektu dającego sygnał o większej amplitudzie.

Dodatek

Sumę dwóch przebiegów sinusoidalnych o różnych częstotliwościach i różnych amplitudach

$$u_1 = U_1 \sin(\omega_1 t + \varphi_1), \tag{17}$$

$$u_2 = U_2 \sin(\omega_2 t + \varphi_2), \quad U_2 > U_1, \tag{18}$$

można sprowadzić do postaci

$$u = u_1 + u_2 = U \sin(\omega_2 t + \varphi). \tag{19}$$

Wyrażenie to tylko wówczas przedstawia sobą przebieg o zmiennej amplitudzie i zmiennej częstotliwości, gdy różnica częstotliwości sygnałów składowych jest znacznie mniejsza od każdej z tych częstotliwości, tzn. gdy

$$|\omega_1 - \omega_2| \ll \omega_1. \tag{20}$$

Wartość kąta φ można wyznaczyć podstawiając do wyrażeń (17) i (18) oznaczenie

$$\omega_r = \omega_1 - \omega_2 \tag{21}$$

i korzystając z prostych przekształceń trygonometrycznych. Wówczas

$$u_1 = U_1 \sin[(\omega_2 + \omega_r)t + \varphi_1] = U_1[\sin\omega_2 t \cos(\omega_r t + \varphi_1) + \cos\omega_2 t \sin(\omega_r t + \varphi_1)], \tag{22}$$

$$u_2 = U_2 \sin(\omega_2 t + \varphi_2) = U_2[\sin\omega_2 t \cos\varphi_2 + \cos\omega_2 t \sin\varphi_2], \tag{23}$$

$$u = U \sin(\omega_2 t + \varphi) = U[\sin\omega_2 t \cos\varphi + \cos\omega_2 t \sin\varphi]. \tag{24}$$

Z zależności tych wynika, że spełnione jest tożsamościowo

$$\operatorname{tg}\varphi \equiv \frac{K\sin(\omega_r t + \varphi_1) + \sin\varphi_2}{K\cos(\omega_r t + \varphi_1) + \cos\varphi_2}, \tag{25}$$

gdzie

$$K = \frac{U_1}{U_2} < 1. \tag{26}$$

Częstotliwość chwilowa przebiegu (19) równa jest zatem

$$f = \frac{1}{2\pi}\left(\omega_2 + \frac{\mathrm{d}\varphi}{\mathrm{d}t}\right). \tag{27}$$

Dla wyrażenia jej wystarczy obliczyć odchyłkę częstotliwości

$$\Delta f = \frac{1}{2\pi}\frac{\mathrm{d}\varphi}{\mathrm{d}t}. \tag{28}$$

Wartość jej można wyznaczyć różniczkując wyrażenie (25) i wprowadzając proste przekształcenia trygonometryczne. W wyniku tego otrzymuje się

$$\Delta f = \frac{\omega_r}{2\pi}\frac{K^2 + K\cos(\omega_r t + \varphi_1 - \varphi_2)}{1 + K^2 + 2K\cos(\omega_r t + \varphi_1 - \varphi_2)} = \frac{\omega_r}{2\pi} y(\xi), \tag{29}$$

gdzie

$$y(\xi) = \frac{K^2 + K\cos\xi}{1 + K^2 + 2K\cos\xi}. \tag{30}$$

Funkcja $y(\xi)$ jest okresowa z okresem $\xi = 2\pi$ i symetryczna względem $\xi = \pi$. Wykres tej funkcji w zakresie jednego okresu jest pokazany na rys. 3 dla kilku wartości K. Funkcja ta w zakresie jednego okresu osiąga maksimum

$$y_{\max} = \frac{K}{1+K} \quad \text{dla} \quad \xi = 0{,}2\ \pi \tag{31}$$

oraz minimum

$$y_{\min} = -\frac{K}{1+K} \quad \text{dla} \quad \xi = \pi. \tag{32}$$

Średnia jej wartość za okres wynosi

$$y_{śr} = \frac{1}{\pi}\int_0^{\pi} y(\xi)\,d\xi = \frac{K^2}{\pi}\int_0^{\pi}\frac{d\xi}{1+K^2+2K\cos\xi} + \frac{K}{\pi}\int_0^{\pi}\frac{\cos\xi\,d\xi}{1+K^2+2K\cos\xi}. \tag{33}$$

Poszczególne całki tego wyrażenia można obliczyć korzystając z podstawienia

$$a = \frac{2K}{1+K^2} \tag{34}$$

i z zależności [2]

$$\int_0^{\pi}\frac{\cos nx\,dx}{1+a\cos x} = \frac{\pi}{\sqrt{1-a^2}}\left(\frac{\sqrt{1-a^2}-1}{a}\right)^n, \quad a^2 < 1. \tag{35}$$

Otrzymuje się wówczas dla $n = 0$ i $n = 1$

$$\int_0^{\pi}\frac{dx}{1+a\cos x} = \frac{\pi}{\sqrt{1-a^2}}, \tag{36}$$

$$\int_0^{\pi}\frac{\cos x\,dx}{1+a\cos x} = \frac{\pi}{\sqrt{1-a^2}}\,\frac{\sqrt{1-a^2}-1}{a} \tag{37}$$

oraz

$$y_{śr} = \frac{K^2}{\pi(1+K^2)}\int_0^{\pi}\frac{d\xi}{1+a\cos\xi} + \frac{K}{\pi(1+K^2)}\int_0^{\pi}\frac{\cos\xi\,d\xi}{1+a\cos\xi} = \frac{K^2}{\pi(1+K^2)}\,\frac{\pi}{\sqrt{1-a^2}} +$$
$$+ \frac{K}{\pi(1+K^2)}\,\frac{\pi}{\sqrt{1-a^2}}\,\frac{\sqrt{1-a^2}-1}{a} = \frac{K}{(1+K^2)\sqrt{1-a^2}}\left(K+\frac{\sqrt{1-a^2}-1}{a}\right). \tag{38}$$

Po podstawieniu wyrażenia (34) otrzymuje się

$$y_{śr} = 0. \tag{39}$$

LITERATURA CYTOWANA W TEKŚCIE

1. T. K. Ishii, Analysis of Target-Speed Determination with Doppler Radar, IEEE Trans. on Instrumentation and Measurement, Vol. JM-19, p. 86-91, No. 2, May 1970.
2. J. S. Gradsztejn, J. M. Ryżik, Tablicy integrałow, sum, rjadow i proizwiedienij, Gos. Izd. Fiz. Mat. Lit., Moskwa 1962.

A. LIZOŃ

ANALYSIS OF DISTURBING SIGNAL INFLUENCE UPON THE VELOCITY MEASUREMENT WITH DOPPLER METHOD

Summary

Target-speed determination by means of the Doppler method with simultaneous reception of the signal coming from another moving target was discussed. The problem was solved by analyzing the instantaneous frequency of the received resultant signal when the speed of the targets differs slightly. The obtained

results show that with a long measurement time of the Doppler frequency in comparison to the period of frequency variations the measured values correspond to the speed of the target from which the received signal is of greater amplitude. With a short measurement time the results may vary from the mean speed of both targets to the value which deviates considerably in the direction opposite from the speed of the target which gives a greater amplitude signal.

A. LIZOŃ

BEURTEILUNG DES STÖRENDEN SIGNALEINFLUSSES AUF DIE GESCHWINDIGKEITSMESSUNGEN MITTELS DOPPLER-METHODE

Zusammenfassung

Im Aufsatz wurde die Frage der Geschwindigkeitbemessung eines Gegenstandes mittels dopplerischer Methode beim gleichzeitigen Empfang eines von einem anderen sich in Bewegung befindenden Gegenstand herrührenden Sygnals erörtert. Das Problem wurde gelöst, indem eine Analyse der Augenblicksfrequenz eines resultierenden Signals für den Fall durchgeführt wurde, wofür der Geschwindigkeitsunterschied der einzelnen Gegenstände nur gering war. Die erhaltenen Ergebnisses deuten darauf, daß bei langer Meßzeit dopplerischer Frequenz im Vergleich mit der Schwankungsperiode dieser Frequenz der gemessene Geschwindigkeitswert derjenigen Geschwindigkeit des Gegenstandes entspricht, dessen empfangenes Signal die größte Amplitude hat. Bei kurzer Meßzeit dieser Frequenz können die Ergebnisse schwanken, und zwar im Bereich vom arithmetischen Mittelwert der Geschwindigkeit beider Gegenstände bis zu einem Wert, der von der Geschwindigkeit des das Signal ausstoßenden Gegenstandes mit größerer Amplitude in entgegengesetzter Richtung bedeutend abweicht.

А. ЛИЗОНЬ

ОЦЕНКА ВЛИЯНИЯ СИГНАЛА ПОМЕХ НА ИЗМЕРЕНИЕ СКОРОСТИ ДВИЖЕНИЯ ПРИ ИСПОЛЬЗОВАНИИ ЭФФЕКТА ДОППЛЕРА

Резюме

В работе изложена проблема измерения скорости движения объекта в доплеровских системах при одновременном приеме сигналов от полезного и мешающего движущихся объектов. Проблема эта решена путем проведения анализа мгновенной частоты результирующего принимаемого сигнала когда разность скорости этих объектов незначительна. Полученные результаты указывают, что при длинном интервале времени измерения доплеровской частоты по сравнению с периодом отклонений этой частоты, измеренная скорость определяет скорости объекта, которого сигнал имеет большую амплитуду. При коротком интервале времени измерения этой частоты результаты могут колебаться, начиная от средней арифметической скорости обоих объектов до величины значительно отличающейся но в обратном направлении скорости объекта дающего сигнал большой амплитуды.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 2, ss. 397—408

621.313.333

Samohamowność dwufazowych silników indukcyjnych o masywnym wirniku ferromagnetycznym

JACEK GIERAS (POZNAŃ)

Instytut Elektrotechniki Przemysłowej Politechniki Poznańskiej
Zakład Maszyn Elektrycznych

Otrzymano 5.6.1973

W artykule określono kryteria samohamowności dwufazowych silników wykonawczych o jednorodnym masywnym wirniku ferromagnetycznym dla zwartego oraz rozwartego uzwojenia sterowania. Wyznaczono dopuszczalne maksymalne średnice wirnika, dla których silnik jeszcze spełnia warunek samohamowności przy rozwartym uzwojeniu sterowania.

1. WPROWADZENIE

W układach automatycznej regulacji oprócz wykonawczych dwufazowych silników indukcyjnych o wirnikach klatkowym lub para- albo diamagnetycznym kubkowym stosuje się również silniki o masywnym wirniku ferromagnetycznym. Teoria tych ostatnich nie jest dostatecznie zbadana i wymaga wielu uzupełnień. Jednym z problemów, który dotychczas nie był analizowany są warunki samohamowności silników o wirniku masywnym. Sformułowanie kryteriów samohamowności, czyli określenie warunku koniecznego do występowania samohamowności polega na wyznaczeniu minimalnej rezystancji lub modułu impedancji wirnika (w przypadku masywnych wirników ferromagnetycznych), przy której występuje jeszcze rozważana właściwość silnika. Samohamowność dwufazowych silników wykonawczych o wirnikach klatkowym i kubkowym jest przedmiotem między innymi prac [1], [2], [4], [5]. Stwierdzono, że silniki te są samohamowne przy rozwartym uzwojeniu sterowania, jeżeli:

$$R_{2s} \geqslant X_{\mu s} + X_{2s} \tag{1}$$

oraz przy zwartym uzwojeniu sterowania, jeżeli:

$$R_{2s} \geqslant X_{\mu s} \sqrt{\frac{R_{1s}^2 \left(1 + \frac{X_{2s}}{X_{\mu s}}\right)^2 + X_{1s}^2 \left(1 + \frac{X_{2s}}{X_{\mu s}} + \frac{X_{2s}}{X_{1s}}\right)^2}{R_{1s}^2 + (X_{\mu s} + X_{1s})^2}} \tag{2}$$

przy czym:

R_{1s} — rezystancja uzwojenia sterowania,

X_{1s} — reaktancja rozproszenia uzwojenia sterowania,

$X_{\mu s}$ — reaktancja główna sprowadzona do uzwojenia sterowania,

R_{2s} — rezystancja wirnika niezależna od poślizgu sprowadzona do uzwojenia sterowania,

X_{2s} — reaktancja rozproszenia wirnika niezależna od poślizgu, sprowadzona do uzwojenia sterowania.

Rezystancja i reaktancja gałęzi wirnikowej schematu zastępczego silnika o masywnym wirniku ferromagnetycznym, sprowadzone do uzwojenia sterowania — rys. 1, dla pola magnetycznego współbieżnego odpowiednio wynoszą [8]:

$$R_{2s}^{+} = a_R k_z Z_{cs} \sqrt{\frac{\mu_{rs}^{+}}{s}}, \tag{3}$$

$$X_{2s}^{+} = a_X k_z Z_{cs} \sqrt{\frac{\mu_{rs}^{+}}{s}}, \tag{4}$$

gdzie:

a_R, a_X — współczynniki zmiany odpowiednio rezystancji i reaktancji wirnika na skutek nieliniowości przenikalności magnetycznej w głąb wirnika oraz strat histerezowych według Nejmana [6],

k_z — współczynnik tzw. *efektu krańcowego* [2], [3], [7], [8],

Z_{cs} — stała wartość modułu impedancji wirnika sprowadzonej do uzwojenia sterowania,

μ_{rs} — względna przenikalność magnetyczna na powierzchni wirnika,

s — poślizg wirnika względem pola wirującego współbieżnego (dla przeciwbieżnego $2-s$).

Dla pola przeciwbieżnego zależności (3), (4) przyjmują postać:

$$R_{2s}^{-} = a_R k_z Z_{cs} \sqrt{\frac{\mu_{rs}^{-}}{2-s}}, \tag{5}$$

$$X_{2s}^{-} = a_X k_z Z_{cs} \sqrt{\frac{\mu_{rs}^{-}}{2-s}} \tag{6}$$

W stanie zwarcia moduły impedancji wirnika dla pola magnetycznego współbieżnego i przeciwbieżnego są jednakowe:

$$|\underline{Z}_2|_{s=1} = \sqrt{a_R^2 + a_X^2}\, k_z Z_{cs} \sqrt{\mu_{rs}}. \tag{7}$$

Współczynnik *efektu krańcowego* jednorodnego wirnika masywnego wyraża się następującymi półempirycznymi zależnościami:

— według Gibbsa [3]

$$k_z \approx 1 + \frac{2}{\pi}\frac{\tau}{L_2}, \tag{8a}$$

— według Panasienkowa [7]

$$k_z \approx 1 + 0{,}5\frac{\tau}{L_2}, \tag{8b}$$

— według autora

$$k_z \approx 1 + \frac{\pi}{\pi + p\dfrac{\tau}{L_2}}\left(\frac{\tau}{L_2}\right)^2, \tag{8c}$$

gdzie:

τ — podziałka biegunowa,

L_2 — długość wirnika,

p — liczba par biegunów.

Wzór (8c) został wyprowadzony przy założeniu rozpływu prądów w masywnym wirniku według torów prostokątnych oraz założeniu, że przenikalność magnetyczna na wszystkich powierzchniach walca stanowiącego wirnik jest jednakowa.

Stała wartość modułu impedancji sprowadzona do uzwojenia sterowania, niezależna od poślizgu i zmiennej przenikalności magnetycznej wynosi:

$$Z_{cs} = \frac{4m_1(z_{1s}\zeta_1)^2}{\pi D_2} L_2 \sqrt{\frac{\omega_1 \mu_0}{2\gamma_2}}, \tag{9}$$

gdzie:

m_1 — liczba faz stojana,

$z_{1s}\zeta_1$ — efektywna liczba zwojów szeregowych uzwojenia sterowania,

D_2 — średnica wirnika,

$\omega_1 = 2\pi f_1$ — pulsacja prądu w uzwojeniach stojana,

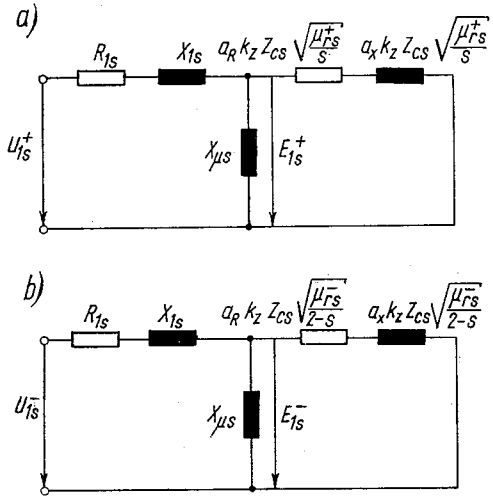
μ_0 — przenikalność magnetyczna próżni,

γ_2 — konduktywność materiału wirnika.

2. ZAŁOŻENIA

Kryteria samohamowności dwufazowych silników indukcyjnych o masywnym wirniku ferromagnetycznym zostaną określone przy następujących założeniach:

a) wirnik jest wykonany z jednorodnego izotropowego materiału w postaci pełnego walca lub cylindra o grubości ścianki większej od równoważnej głębokości wnikania fali elektromagnetycznej,



Rys. 1. Schematy zastępcze dwufazowego silnika indukcyjnego o masywnym wirniku ferromagnetycznym (dla uzwojenia sterowania): a) dla pola magnetycznego współbieżnego, b) dla pola magnetycznego przeciwbieżnego

b) zmienna przenikalność magnetyczna w głąb masywnego wirnika ferromagnetycznego i straty histerezowe są uwzględnione półempiryczną metodą Nejmana [6], według której rezystancję wirnika (przy stałej przenikalności magnetycznej i bez występowania strat na histerezę) mnoży się przez współczynnik $a_R \approx 1{,}3...1{,}5$, reaktancję zaś — przez współczynnik $a_X \approx 0{,}8...0{,}9$,

c) straty mocy w rdzeniu stojana praktycznie nie występują (rezystancja strat dąży do nieskończoności) i w schemacie zastępczym gałąź poprzeczna zawiera tylko reaktancję główną $X_{\mu s}$ — rys. 1,

d) silnik jest sterowany amplitudowo lub amplitudowo-fazowo, tzn. sygnał sterujący może być podawany na uzwojenie sterowania z potencjometru, autotransformatora, urządzenia stykowego, selsyna transformatorowego, wzmacniacza elektronicznego lub magnetycznego.

Zostaną rozpatrzone graniczne przypadki zamknięcia obwodu sterowania poprzez impedancję zewnętrzną, tzn. stan zwarcia (wartość impedancji zewnętrznej równa zeru) oraz przy otwartym obwodzie sterowania (wartość impedancji zewnętrznej dąży do nieskończoności).

3. CHARAKTERYSTYKI MECHANICZNE SILNIKA INDUKCYJNEGO O MASYWNYM WIRNIKU FERROMAGNETYCZNYM

Na podstawie schematu zastępczego — rys. 1, w myśl teorii składowych symetrycznych układów dwufazowych [2], składowa zgodna prądu wirnika sprowadzonego do uzwojenia sterowania jest określona zależnością:

$$\underline{I}^{+}_{2s} = \frac{\underline{U}^{+}_{1s}}{R_{1s}+jX_{1s}+\underline{\sigma}(a_R+ja_X)k_zZ_{cs}\sqrt{\frac{\mu^{+}_{rs}}{s}}} \tag{10}$$

gdzie:

$$\underline{\sigma} = 1+\frac{R_{1s}+jX_{1s}}{jX_{\mu s}} = 1+\frac{X_{1s}}{X_{\mu s}}-j\frac{R_{1s}}{X_{\mu s}}.$$

Przechodząc do wartości skutecznych otrzymuje się:

$$I^{+}_{2s} = \frac{U^{+}_{1s}}{\sqrt{\left\{R_{1s}+\left[a_R\left(1+\frac{X_{1s}}{X_{\mu s}}\right)+a_x\frac{R_{1s}}{X_{\mu s}}\right]k_zZ_{cs}\sqrt{\frac{\mu^{+}_{rs}}{s}}\right\}^2+\left\{X_{1s}+\left[a_x\left(1+\frac{X_{1s}}{X_{\mu s}}\right)-\right.\right.}} \tag{11}$$

$$\left.\left.-a_R\frac{R_{1s}}{X_{\mu s}}\right]k_zZ_{cs}\sqrt{\frac{\mu^{+}_{rs}}{s}}\right\}^2}$$

Moment elektromagnetyczny wywołany polem współbieżnym wynosi:

$$M^{+} = \frac{P^{+}_{w}}{2\pi n_1} = \frac{(I^{+}_{1s})^2a_Rk_zZ_{cs}\sqrt{\frac{\mu^{+}_{rs}}{s}}}{\pi n_1}. \tag{12}$$