4. ROZWIĄZANIE ZADANIA

Problem poszukiwania optymalnego w sensie proponowanego kryterium (3.1) systemu przesyłania informacji, którego model matematyczny zapisano równaniami (2.11), daje się sprowadzić do następującego zadania poszukiwania minimum funkcji kosztów jednej zmiennej z ograniczeniami:

$$\min_x \{K(x)\} = \min_x \left\{ f(n,k,s) + w_B B_0 + \frac{2nR_c}{k(1-x)^n - T_0 R_c} (2w_B - w_c \ln 2x) \right\} \tag{4.1}$$

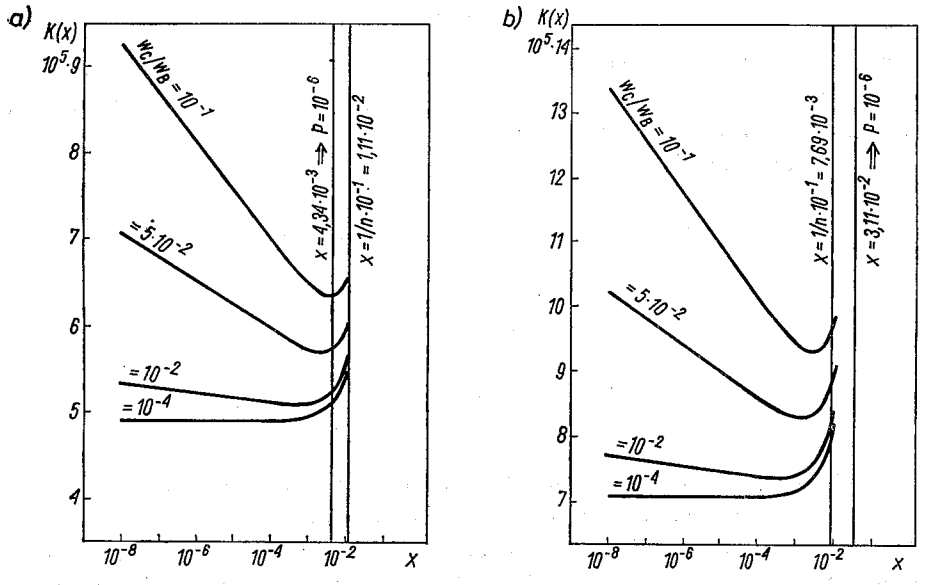
$$\left. \begin{aligned} &\sum_{j=2s+1}^{n} \frac{C_n^{j-s}}{C_j^s} \left(\frac{x}{1-x} \right)^j \leqslant P_c \\ &\frac{1}{2} e^{-\frac{2C_{max}}{B_{min}-B_0}} \leqslant x \leqslant \frac{1}{2}. \end{aligned} \right\} \tag{4.2}$$

Ponieważ próby poszukiwania analitycznego rozwiązania powyższego zadania nastręczają sporo kłopotów natury czysto obliczeniowej, dlatego celem lepszej interpretacji otrzymanych wyników posłużono się metodą graficzną. Funkcję kryterialną (4.1) wraz z ograniczeniami (4.2) przebadano dla różnych wartości prawdopodobieństwa błędu elementarnego $0 < x \ll 1/n$, przy wzajemnych relacjach kosztów jednostkowych pasma w_B i stosunku mocy sygnał/szum w_c typu: $w_B \gg w_C$, $w_B \approx w_C$, $w_B \ll w_C$, oraz ustalonych wartościach:

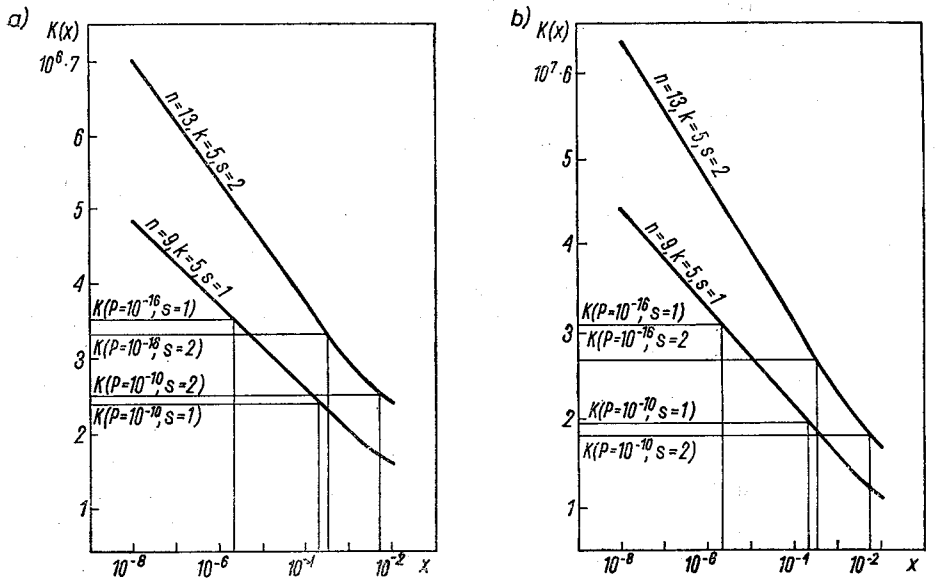
a) kosztów kodowania i dekodowania informacji $K_E = const$ i $K_E \ll (K_B + K_C)$,
b) pasma częstotliwości $B_0 = const$ i czasu odbioru $T_0 = const$ sygnału potwierdzenia w kanale sprzężenia zwrotnego,
c) długości ciągu kodowego $n = const$ i liczby pozycji informacyjnych $k = const$ wybranego kodu nadmiarowego (n, k),
d) prędkości transmisji informacji $R_c = const$.
e) prawdopodobieństwa błędnego odbioru wiadomości $P_c = const$.

Przypadek $w_B \gg w_C$ oznacza, że nakłady finansowe łożone na wygenerowanie sygnału odpowiednio dużej mocy są znacznie mniejsze od nakładów ponoszonych na pasmo częstotliwości zajmowane przez sygnały w systemie. Analizując dla tego przypadku zachowanie się funkcji (4.1) w przedziale $0 < x \ll 1/n$, można stwierdzić, że w znacznym obszarze zmienności x jej szybkość opadania, a następnie narastania jest bardzo mała. Praktycznie można uznać, że koszt systemu jest stały i nie zależy od prawdopodobieństwa błędu elementarnego x, co ilustrują konkretne przykłady zamieszczone na rys. 1 i 3.

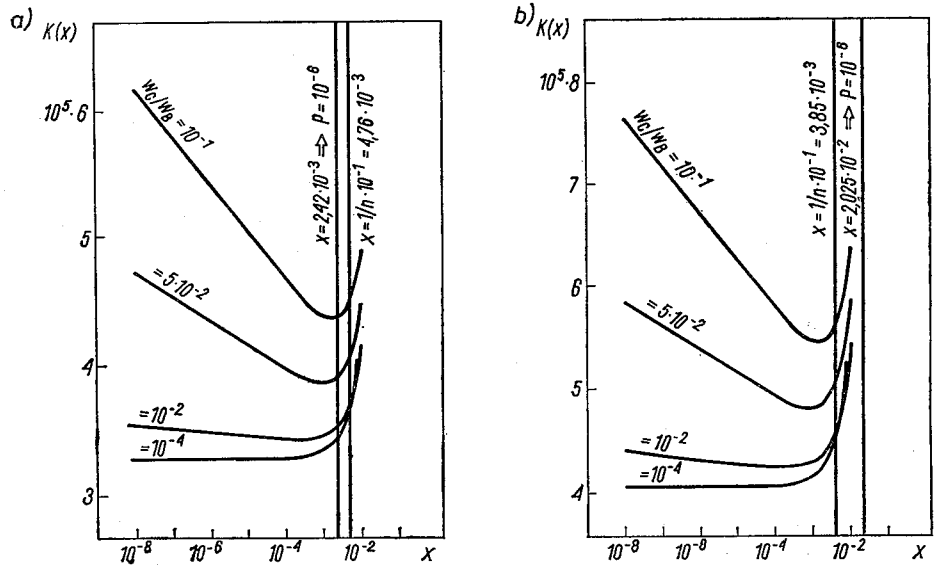
W sytuacji, gdy koszty jednostkowe generowania sygnału dużej mocy rosną ($w_B > w_C$) i stają się porównywalne ($w_B \approx w_C$) z kosztami jednostkowymi pasma, funkcja kryterialna maleje szybko ze wzrostem x i to tym szybciej im większe stają się koszty jednostkowe stosunku mocy sygnał/szum w porównaniu z kosztami jednostkowymi pasma ($w_B < w_C$). Przy zbliżaniu się z prawdopodobieństwem x do obszaru, w którym wartości x stają się porównywalne z $1/n$, wartości funkcji kosztów (4.1) zaczynają gwałtownie rosnąć. Na rysunkach 1 do 4 przedstawiono wykresy funkcji (4.1) dla wybranych przykładów systemów cyfrowych. Rysunki 1 i 2 przedstawiają wykresy funkcji kosztów (4.1) dla systemów z kodem systematycznym (9, 5) i (13, 5), zaś rysunki 3 i 4 wykresy funkcji



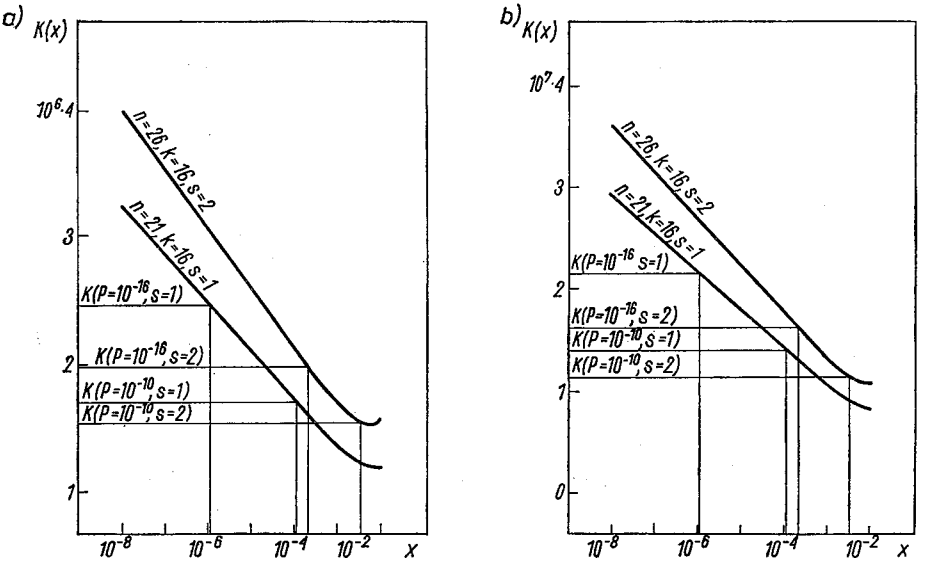
Rys. 1. Wykresy funkcji kosztów $K(x)$ badanego systemu transmisji, wykorzystującego kod systematyczny: a) (9,5), $s = 1$; b) (13,5), $s = 2$



Rys. 2. Wykresy funkcji kosztów $K(x)$ badanego systemu transmisji, przy wzajemnych relacjach kosztów jednostkowych stosunku mocy i pasma: a) $w_c/w_B = 1$; b) $w_c/w_B = 10$; dla kodów (9,5), $s = 1$ i (13,5), $s = 2$



Rys. 3. Wykresy funkcji kosztów $K(x)$ badanego systemu transmisji, wykorzystującego kod systematyczny: a) (21,16), $s = 1$; b) (26.16), $s = 2$



Rys. 4. Wykresy funkcji kosztów $K(x)$ badanego systemu transmisji, przy wzajemnych relacjach kosztów jednostkowych stosunku mocy i pasma: a) $w_c/w_B = 1$; b) $w_c/w_B = 10$; dla kodów (21,16), $s = 1$ i (26,16), $s = 2$

kosztów (4.1) dla systemów z kodem systematycznym (21, 16) i (26, 16). We wszystkich przykładach przyjęto, że żądana prędkość transmisji wynosi $R_c = 600 \frac{\text{bit}}{\text{s}}$, czas odbioru sygnału potwierdzenia $T_0 = 10^{-3}$ s, zaś prawdopodobieństwo błędnego odbioru nie powinno być gorsze niż $P_c = 10^{-6}$.

Przedstawione wykresy funkcji kosztów pozwalają sądzić, że istnieje realna możliwość znalezienia optymalnych wartości prawdopodobieństwa x, a tym samym na bazie równań (2.11) optymalnych wartości pasma i stosunku mocy sygnał/szum, przy których koszt systemu można ustalić na jak najniższym poziomie, przy jednoczesnym zapewnieniu żądanych wartości parametrów informacyjnych, tj. prędkości transmisji R_c i prawdopodobieństwa błędnego odbioru wiadomości P_c.

5. WNIOSKI I UWAGI KOŃCOWE

Z przeprowadzonej w rozdziale 4 analizy oceny kosztów wybranego systemu przesyłania informacji cyfrowych z decyzyjnym sprzężeniem zwrotnym i rysunków 1 do 4 wynika, że właściwy dobór kodu nadmiarowego pozwala ustalić w oparciu o proponowane kryterium (3.1) koszty systemu na możliwie niskim poziomie, przy jednoczesnym zachowaniu wymaganych wartości prędkości transmisji oraz prawdopodobieństwa błędnego odbioru wiadomości. Na potwierdzenie tego wniosku rozważmy następujący przykład. Niech wybrany przez nas system transmisji należy do klasy systemów opisanej we wstępnych rozdziałach niniejszej pracy. Żądamy, by prędkość transmisji $R_c = 600 \frac{\text{bit}}{\text{s}}$, a prawdopodobieństwo błędnego odbioru wiadomości $P_c = 10^{-10}$, $T_0 = 10^{-3}$, zaś liczba pozycji informacyjnych kodu $k = 5$. Załóżmy, że mamy zdecydować, który z dwu kodów systematycznych (9, 5) i (13, 5) będzie kodem lepszym z punktu widzenia kryterium kosztów (3.1). Niech koszty jednostkowe pasma i stosunku mocy sygnał/szum pozostają w następującej relacji: wariant pierwszy $w_C/w_B = 1$, wariant drugi $w_c/w_B = 10$. Dla kodu $k = 5$, $n = 9$, $s = 1$ prawdopodobieństwo błędu elementarnego oszacowane z pierwszej nierówności układu (4.2) przyjmuje wartości $x_1 \simeq 2{,}03 \cdot 10^{-4}$, dla kodu $k = 5$, $n = 13$, $s = 2$ zaś $x_2 \simeq 5{,}09 \cdot 10^{-3}$, przy czym obie te wartości spełniają drugą nierówność układu (4.2). Z rysunków 2a (wariant pierwszy) i 2b (wariant drugi) możemy odczytać koszty systemów z kodem (9, 5) i (13, 5). W wariancie pierwszym, gdy koszty jednostkowe pasma są porównywalne z kosztami jednostkowymi stosunku mocy sygnał/szum, koszt systemu z kodem (9, 5) jest mniejszy od kosztu systemu z kodem (13, 5). Natomiast w wariancie drugim, gdy koszty jednostkowe stosunku mocy sygnał/szum przewyższają dziesięciokrotnie koszty jednostkowe pasma, koszt systemu z kodem (9, 5) jest większy od kosztu systemu z kodem (13, 5). Na rysunkach 2a i 2b zaznaczono także koszty systemów wykorzystujących te dwa kody dla przypadku, gdy żądamy by prawdopodobieństwo błędnego odbioru wiadomości $P_c = 10^{-16}$. Podobną analizę kosztów przeprowadzono na rysunkach 4a i 4b dla kodów (21, 16) i (26, 16). Wynika z niej, że dla obu wariantów kosztów $w_C/w_B = 1$ i $w_C/w_B = 10$ oraz zarówno dla $P_c = 10^{-10}$ jak i $P_c = 10^{-16}$ koszty systemu z kodem (26, 16) są mniejsze od kosztów systemu z kodem (21, 16).

Bibliografia

1. E. R. Berlekamp; *The technology of error-correcting codes.* Proc. IEEE, Nr 5, 1980. pp. 564–593
2. M. Rasiukiewicz, A. Leśnicki: *Podstawy systemów horyzontowych linii radiowych.* Warszawa: WKiŁ, 1983
3. E. G. Serwinski: *Optymizacija sistem peredaci diskretnoj informacji.* Moskwa: Svjaz, 1974
4. B. Van der Veen: *Wstęp do teorii badań operacyjnych.* Warszawa: PWN, 1970

W. LUDWIN

APPLICATION OF THE COST CRITERION IN THE ASSESSMENT OF A CERTAIN SYSTEM OF DATA TRANSMISSION WITH DECISION FEEDBACK

Summary

A digital transmission system with a decision feedback is considered. This system is examined by means of the criterion of costs. The cost function depends on the signal bandwidth, the signal/noise ratio and the encoding, decoding costs. According to the proposed mathematical model of this system and to the criterion of costs the optimisation problem is set. This optimisation problem is successfully solved. It is demonstrated that optimum values of the signal bandwidth and the signal/noise ratio for which transmission costs in the decision feedback system are minimum, can be found.

W. LUDWIN

APPLICATION DU CRITÈRE DE FRAIS POUR ESTIMATION D'UN CERTAIN SYSTÈME A RETOUR DE TRANSMISSION DES DONNÉES

Résumé

On a considéré le système de transmission des données avec boucle. En se basant sur le critère de frais, on a examiné le coût de la transmission des données dans ce système. A partir du modèle mathématique proposé du système et du critère de frais, on a construit le problème de recherche du minimum de la fonction de coût de la transmission avec limitations. On a examiné aussi les différentes variantes de la résolution du problème conçu de la sorte. En tenant compte des asspects économiques de la construction et de l'exploitation des systèmes de transmissions des données, on a montré qu'ile est possible de trouver les systèmes optimaux de transmission des données conformes au critère de frais proposé.

W. LUDWIN

ANWENDUNG DES KOSTENKRITERIUMS ZUR EINSCHÄTZUNG EINES NACHRICHTENÜBERTRAGUNGSSYSTEMS MIT ENTSCHEIDUNGSRÜCKFÜHRUNG

Zusammenfassung

Das digitale Nachrichtenübertragungssystem mit Entscheidungsrückführung wurde erwogen. Nach ökonomischem Kriterium wurden Kosten der Informationstransmission bei Anwendung dieses Systems überprüft. In Anlehnung an das vorgeschlagene mathematische Modell des Systems sowie das Kostenkriterium wurde die Aufgabe des Nachsuchensnachdem Minimum der Kostenfunktion der Transmission mit Beschränkungen konstruiert. Verschiedene Lösungsvarianten für die derartig gestaltete Aufgabe wurden überprüft. Unter Berücksichtigung der ökonomischen Aspekte des Bauens und und Betreibens der Übertragunssysteme von digitalen Nachrichten wurde nachgewiesen, daß es möglich ist, optimale Systeme für die Informationsübertragung im Sinne des vorgeschlagenen Kostenkriteriums zu finden.

В. ЛЮДВИН

ПРИМЕНЕНИЕ КРИТЕРИЯ СТОИМОСТИ К ОЦЕНКЕ ОПРЕДЕЛЕННОЙ СИСТЕМЫ ТРАНСМИССИИ ИНФОРМАЦИИ С РЕШАЮЩЕЙ ОБРАТНОЙ СВЯЗЬЮ

Р е з ю м е

Рассмотрена система передачи дискретной информации с решающей обратной связью. Эта система проанализирована используя критериальную функцию стоимости. Исходя из математической модели системы и критериальной функции стоимости построена задача оптимизации которая потом решена. Показано, что можно найти для системы с обратной связью оптимальные значения отношения мощности сигнала к помехе и полосе частот для которых стоимисть системы согласно критернальной функции достигает минимума.

ROZPRAWY ELEKTROTECHNICZNE, 1988, 34, z. 3, ss. 1007—1029

519.6

Jednomagistralowy system wielomikroprocesorowy

ANTONI ZABŁUDOWSKI, RYSZARD DIRSKA, ANDRZEJ JURKIEWICZ

Wydział Telekomunikacji i Elektrotechniki, Akademia Techniczno-Rolnicza w Bydgoszczy

Otrzymano 1986.05.07

Autoryzowano do druku 1987.11.11

Opisano system wielomikroprocesorowy zbudowany z wykorzystaniem mikroprocesorów MCY 7880. Podano ogólną koncepcję rozwiązania systemu, jego rzeczywiszą realizację oraz model matematyczny opisujący sposób działania — pozwalający na wyznaczenie niezbędnych charakterystyk systemu. Praca składa się ze wstępu i czterech części. W pierwszej zdefiniowano pojęcie tego systemu, przedstawiono jego organizację logiczną, a także pokazano i omówiono chatakterystykę mocy obliczeniowej. W pierwszej części omówiono systemy ze wspólną magistralą. Druga część zawiera opis rzeczywistej realizacji systemu, w którym wykorzystano naturalny podział cyklu maszynowego na cykle zegarowe. Trzecia i czwarta część pracy poświęcona jest analizie zaproponowanego rozwiązania systemu wielomikroprocesorowego. Zaproponowany model matematyczny jest łańcuchem Markowa, w którym stany reprezentowane są w postaci wektora o pięciu składowych, wyznaczających liczby procesorów znajdujących się w poszczególnych taktach zegarowych μP MCY 7880. Dla oceny pracy systemu porównano wyniki uzyskane metodą analityczną i metodą symulacyjną, a także wyniki uzyskane drogą symulacyjną dla systemów pracujących z jedną magistralą i pamięcią główną oraz dodatkowymi magistralami i pamięciami lokalnymi.

WSTĘP

Jednym z efektów rewolucji technologicznej w dziedzinie mikroelektroniki, a w szczególności w dziedzinie projektowania i wytwarzania układów LSI oraz VLSI, stało się powszechne wykorzystanie systemów mikroprocesorowych. Koszty produkowanych obecnie systemów mikroprocesorowych maleją szybko, tak, że możliwe staje się wykorzystanie tych systemów w bardzo wielu różnych zastosowaniach, w szczególności jako systemów obliczeniowych czy sterujących. Rośnie więc zapotrzebowanie na tworzenie nowych systemów mikroprocesorowych o zwiększonej mocy obliczeniowej. Współczesne systemy o zwiększonej mocy obliczeniowej mogą być tworzone dwoma różnymi sposobami [6]:

a) na drodze technologicznej poprzez działanie mające na celu usprawnienie architektury procesora, a tym samym zwiększenie możliwości obliczeniowych pojedynczego mikroprocesora,

b) poprzez usprawnienie pracy istniejących mikroprocesorów na poziomie systemowym, które polega na tworzeniu złożonych systemów mogących pracować współbieżnie.

Działania usprawniające pracę mikroprocesora, które odbywają się na poziomie technologicznym przeprowadzane muszą być u producenta takich systemów. Z reguły produkt wypuszczany na rynek stanowi pewien kompromis pośród wymagań stawianych przez różne grupy użytkowników. Dlatego też poza tworzonymi, do ściśle określonych zastosowań specjalizowanych układów sterujących, większość produkowanych systemów mikroprocesorowych budowana jest jako systemy uniwersalne. Stąd też działania na poziomie użytkowników systemów mikroprocesorowych związane być muszą z wykorzystaniem koncepcji przetwarzania współbieżnego. Idea przetwarzania współbieżnego polega na zdekomponowaniu (segmentacji) pierwotnego problemu (procesu) na szereg mniejszych zadań, które realizowane w sposób równoległy według określonej kolejności pozwalają na zmniejszenie całkowitego czasu przetwarzania [7]. Wykorzystanie systemów mikroprocesorowych do przetwarzania współbieżnego polega na tworzeniu systemów złożonych z kilku mikroprocesorów lub mikrokomputerów, przy czym najczęściej tworzy się systemy wielomikroprocesorowe złożone z wielu mikroprocesorów pracujących równolegle. Koncepcje dotyczące sposobu pracy systemów wielomikroprocesorowych zostały w sposób bezpośredni przeniesione ze znanych rozwiązań zrealizowanych z wykorzystaniem dużych systemów procesorowych [4, 5].

Zasadniczym czynnikiem, który zdecydował o tym, że systemy wielomikroprocesorowe stały się opłacalne był fakt, że znaczna poprawa charakterystyk systemu odbywała się przy relatywnie niewielkim wzroście kosztu sprzętu wykorzystanego do realizacji takiego systemu. Poza zwiększeniem mocy obliczeniowej systemu wielomikroprocesorowego, uległa również zwiększeniu niezawodność takiego systemu, a sam system mógł być wykorzystany do rozwiązywania szerszej klasy problemów, a więc stawał się bardziej elastyczny.

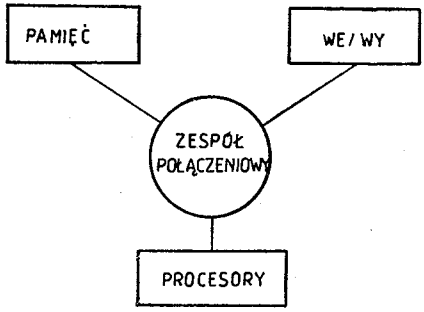
Podstawowa definicja systemu wielomikroprocesorowego zaproponowana przez Enslowa [4, 5] jest następująca:

System wieloprocesorowy jest takim, który charakteryzuje się następującymi cechami:

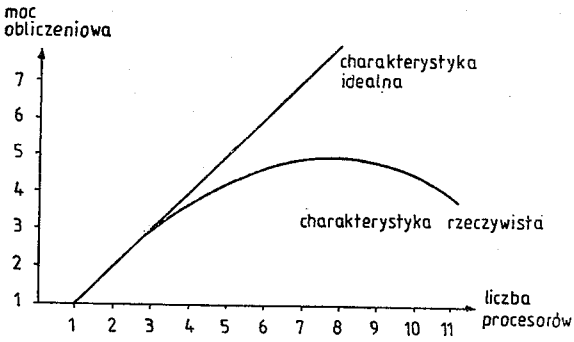
a) zawiera dwa lub więcej procesorów o zbliżonych, choć niekoniecznie, możliwościach obliczeniowych,
b) wszystkie procesory mają dostęp do wspólnych zasobów systemu, a więc pamięci (operacyjnej i zewnętrznej) oraz urządzeń WE-WY,
c) pracą całego systemu steruje wyspecjalizowany blok sterujący, zaś współpraca między poszczególnymi procesorami odbywać się może na różnych poziomach wykonywanych zadań.

Podstawowa organizacja logiczna systemu wieloprocesorowego przedstawiona została na rys. 1. O tym jaka jest fizyczna (konkretna) organizacja systemu wieloprocesorowego decyduje rodzaj użytego zespołu połączeniowego, który zapewnia komunikację między procesorami a pamięcią (pamięciami) oraz urządzeniami WE-WY. Stosowane są trzy zasadnicze organizacje fizyczne systemu wieloprocesorowego:

— systemy ze wspólną magistralą (szyną) i podziałem czasu na magistrali,
— systemy z przełącznicą krzyżową,
— systemy wielomagistralowe z wielobramowymi (wieloportowymi) pamięciami i urządzeniami WE-WY.



Rys. 1. Organizacja logiczna systemu wieloprocesorowego

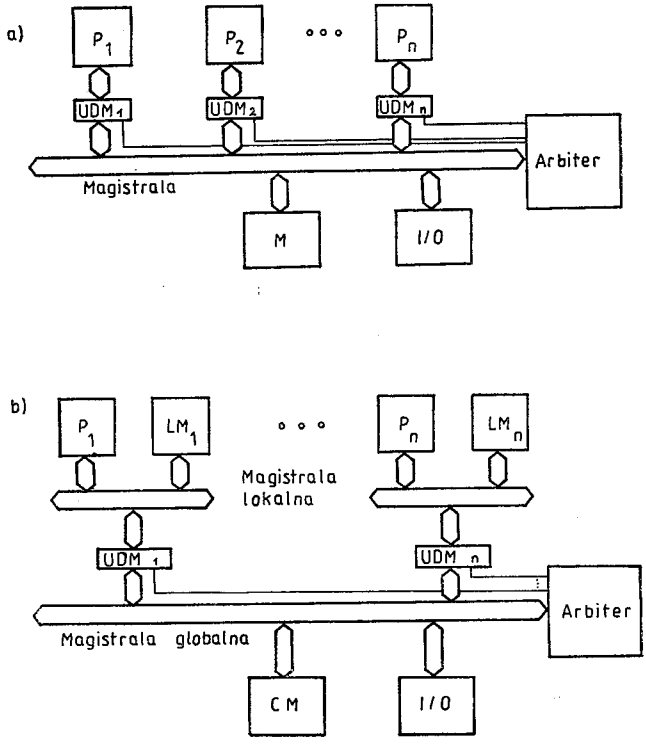


Rys. 2. Przebieg funkcji określającej wielkość mocy obliczeniowej w zależności od liczby procesorów

Fakt, że poszczególne procesory uzyskują dostęp do zasobów systemu poprzez jeden wspólny zespół połączeniowy powoduje, że w systemach wieloprocesorowych występuje oddziaływanie (interferencja) poszczególnych procesorów na siebie. Oddziaływanie poszczególnych procesorów na siebie powoduje, że przyrost mocy obliczeniowej (szybkości obliczeń, throughput) nie jest w całym zakresie funkcją rosnącą liczby procesorów. Przykładowy przebieg funkcji określającej wielkość mocy obliczeniowej w zależności od liczby procesorów przedstawiony został na rys. 2. W porównaniu do idealnej charakterystyki liniowej, charakterystyka rzeczywista osiąga dla pewnej liczby procesorów nasycenie, a następnie wraz z dalszym wzrostem liczby procesorów spadek. Efekt nasycenia i degradacji charakterystyk określających moc obliczeniową systemu jest wynikiem rywalizacji procesorów o dostęp do zasobów systemu, a także wynikiem wzrostu liczby wymienianych informacji sterujących między procesorami niezbędnych dla zapewnienia komunikacji międzyprocesorowej.

1. SYSTEMY WIELOPROCESOROWE PRACUJĄCE ZE WSPÓLNĄ MAGISTRALĄ

Jedną z najprostszych a zarazem najtańszą organizacją fizyczną systemu wieloprocesorowego jest organizacja systemu pracującego ze wspólną magistralą, w której dostęp do zasobów systemu odbywa się poprzez przydział magistrali procesorom żądającym tego przydziału na pewien moment czasu.



Rys. 3. Architektura systemu wieloprocesorowego ze wspólną magistralą a) bez magistrali lokalnych b) z magistralami i pamięciami lokalnymi

Najprostsza architektura systemu wieloprocesorowego ze wspólną magistralą przedstawiona jest na rys. 3a. W skład tego systemu wchodzi n procesorów $P_1, P_2, \ldots P_n$, jedna wspólna pamięć M oraz wspólny układ służący do zapewnienia współpracy z urządzeniami WE-WY. Dostęp do magistrali systemu jest załatwiony poprzez układy dostępu do magistrali (UDM), które sterowane są sygnałami wysyłanymi z bloku nadzorującego pracą całego systemu nazywanego arbitrem. Praca systemu ze wspólną magistralą odbywa się w następujący sposób: każdy z procesorów komunikuje się z pamięcią główną poprzez wspólną magistralę i to zarówno w celu pobrania rozkazu z pamięci jak i pobrania lub zapisania danych w tej pamięci. Komunikacja międzyprocesorowa odbywa się również poprzez pamięć główną, gdyż komunikujące się między sobą procesory umieszczają niezbędne informacje w odpowiednich komórkach pamięci. Częste odwoływanie się poszczególnych procesorów do pamięci powoduje, że występuje tutaj bardzo silna interferencja międzyprocesorowa tak, że charakterystyki systemów ze wspólną magistralą ulegają szybkiemu nasyceniu. Jak wykazują badania [13, 17] nasycenie systemów pracujących z podziałem czasu występuje już dla $n = 3 \div 5$. Z tego względu bardziej użyteczny staje się system, w którym każdy procesor posiada własną pamięć lokalną LM_i zawierającą najczęściej program realizowany przez procesor. Pamięć główna CM służy tylko do przechowywania wybranych bloków danych oraz tych informacji, które służą do komunikacji międzyprocesorowej. Dzięki temu, że w takich systemach intensywność odwołań do

magistrali głównej znacznie maleje (odwołanie tylko do pamięci głównej CM oraz układów sterujących WE-WY), systemy wykorzystujące pamięci lokalne są wystarczająco sprawne nawet dla liczby procesorów $n > 10$ [19]. Analiza systemów ze wspólną szyną pracujących na zasadzie przydziału procesorowi magistrali na pewien moment czasu (niezbędny np. do przesyłania wyników do lub z pamięci czy też urządzeń WE-WY) przeprowadzona została w pracach [17, 19, 20]. Rozważania w tych pracach dotyczyły systemów z pamięciami lokalnymi i prowadzone były przy założeniu markowowskich modeli pracy procesorów tzn. przy założeniu, że odstęp czasu pomiędzy odwołaniami do pamięci jest opisany zmienną losową o rozkładzie wykładniczym oraz czas trwania kontaktu z pamięcią jest również opisany zmienną losową o rozkładzie wykładniczym. Zaproponowany model matematyczny opisujący zachowanie się takiego systemu jest co prawda procesem Markowa, lecz posiada bardzo dużą liczbę stanów dla $n \geqslant 5$ [17]. Dlatego też znalezienie rozkładu stacjonarnego a tym samym określenie charakterystyk systemu nastręcza dużych trudności obliczeniowych. Autorzy proponują więc dwa podejścia prowadzące do uzyskania wyników:

— zastosowanie modeli przybliżonych [19], [18] — dla systemów wieloszynowych przy czym zaproponowany tam model może być zastosowany również dla systemów jednoszynowych),
— wykorzystanie symulacji [17, 20].

Z uzyskanych wyników widać, że dla małych intensywności odwołań do pamięci głównej, liczba procesorów pracujących w systemach wieloprocesorowych z pamięciami lokalnymi może wynosić nawet $10 \div 12$. W szczególności systemy takie mogą być bardzo użyteczne w warunkach gdy procesory spełniają rolę wyspecjalizowanych systemów obliczeniowych, zaś kontakt przez magistralę główną odbywa się tylko z urządzeniami WE-WY. W pracy [16] przedstawiony został model systemu wieloprocesorowego bez pamięci głównej pracującego ze wspólną magistralą, w którym procesory wyposażone są w pamięci lokalne i rozwiązują tylko ściśle określone zadania. Proces wejściowy dekomponowany zostaje na szereg zadań, które rozwiązywane są w sposób nakładkowy (pipelining) przez poszczególne procesory. Z istoty działania systemów pracujących w sposób nakładkowy wynika ich okresowość, ze stałą długością cyklu przetwarzania przez każdy procesor, co z kolei pozwala określić żądane charakterystyki systemu.

Przytoczone modele matematyczne opisujące zachowanie się systemów wieloprocesorowych ze wspólną szyną dotyczyły tylko sytuacji, gdy współpraca między procesorami odbywała się na poziomie zadań i praktycznie nie zależała od użytego procesora. W pracach [8, 9, 10, 11, 26, 27] zajęto się analizą systemów wieloprocesorowych zbudowanych z mikroprocesorów. Prace [8, 10] nie dotyczą co prawda systemów ze wspólną szyną, ale opisują specyfikę działania specjalizowanych systemów wieloprocesorowych wykorzystujących własności mikroprocesora MC 6809, jednakże pozostałe prace traktują o systemach jednoszynowych. We wszystkich tych pracach w trakcie analizy zakłada się, że komunikacja procesorów z pamięcią odbywa się w czasie trwania taktu zegarowego (lub nawet części [8]) i dostęp określonego procesora do magistrali następuje na czas trwania jednego taktu zegarowego. Zaproponowana metoda pracy systemów wieloprocesorowych wykorzystuje naturalny podział cyklu maszynowego mikroprocesora na cykle zegarowe oraz fakt, że tylko w jednym ściśle określonym cyklu zegarowym następuje odwołanie do pamięci.

Wynika stąd wniosek, że w trakcie jednego cyklu maszynowego można dokonać przydziału magistrali (pamięci) kilku różnym procesorom pracującym współbieżnie. Pierwszą pracą poświęconą analizie tak pracującego systemu była praca S. Hoenera, W. Roehdera [9], w której autorzy dokonali oceny charakterystyk systemu wielomikroprocesorowego pracującego na μ PZ-80, przy czym przydział procesora do szyny załatwiony był poprzez określenie priorytetów poszczególnym mikroprocesorom. Z kolei w pracach [26, 27] przedstawiona została analiza systemu wielomikroprocesorowego zrealizowanego na μP Intel 8080 (MCY 7880), w którym wszystkie procesory pracują jako równouprawnione. Co prawda w pracach [2, 27] zaproponowano synchroniczny przydział procesora do szyny, lecz istota pracy takiego systemu pozostaje podobna do ogólnie przyjętego sposobu pracy systemów wieloprocesorowych z pojedynczą magistralą.

Ważną rolę w zapewnieniu realizacji współpracy wszystkich procesorów w systemie wieloprocesorowym spełnia blok nazywany arbitrem systemu. Zadaniem bloku arbitra jest zapewnienie przydziału procesorów do zasobów systemu w taki sposób, ażeby nie pojawiły się sytuacje konfliktowe polegające na tym, że dwa lub więcej procesorów żądać będzie dostępu do tego samego zasobu. Rozstrzygnięcie o tym, który z procesorów uzyska dostęp do zasobu musi być znane w momencie, gdy procesor otrzyma informację o możliwości korzystania z wybranego zasobu. Każdy blok arbitra w systemie wieloprocesorowym może być realizowany zarówno środkami sprzętowymi jak i programowymi [21].

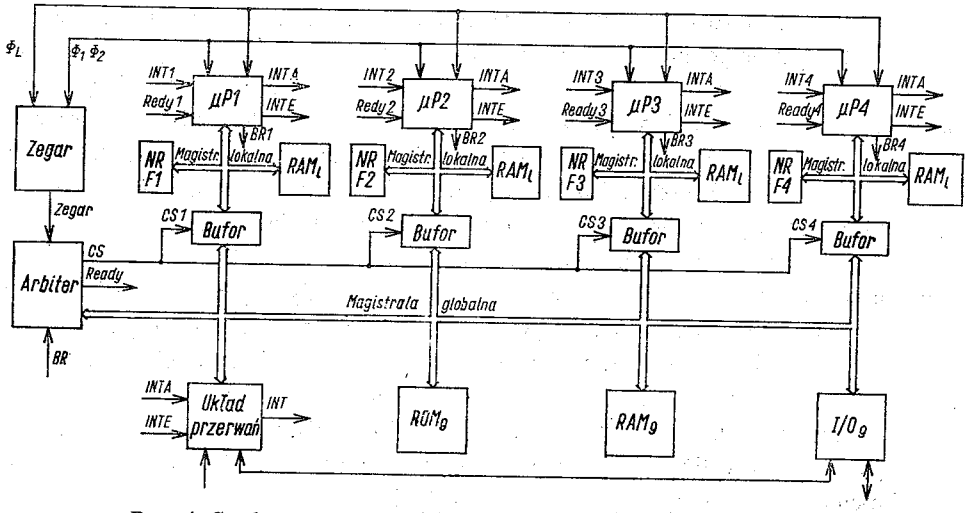
2. REALIZACJA SYSTEMU WIELOPROCESOROWEGO Z WYKORZYSTANIEM MIKROPROCESORÓW MCY 7880

Jednomagistralowy układ systemu wielomikroprocesorowego, który pracuje na zasadzie przydziału cyklu zegarowego poszczególnym mikroprocesorom zbudowany został na Wydziale Telekomunikacji i Elektrotechniki ATR w Bydgoszczy. System składa się z czterech mikroprocesorów MCY 7880, z których każdy jest równouprawniony. Pracuje on z układem arbitra zdecentralizowanego, przy czym w odróżnieniu od modelu zaproponowanego w pracy [27] wykorzystuje się asynchroniczny sposób przydziału magistrali procesorowi. Schemat blokowy systemu przedstawiono na rys. 4.

W jego skład wchodzą następujące moduły:

— cztery niezależne układy mikroprocesorowe identyfikowane między sobą, które mogą być wyposażone w pamięci lokalne RAM, przyłączone do magistrali lokalnej każdego z mikroprocesorów,
— moduł wspólnego zegara generującego wysokonapięciowe sygnały zegarowe Φ_1 i Φ_2 wymagane przez mikroprocesory MCY 7880,
— moduł arbitra sterującego dostępem mikroprocesorów do magistrali globalnej, w sytuacji gdy dwa lub więcej mikroprocesorów żąda dostępu do magistrali,
— moduł obsługi zgłoszeń pracujący na zasadzie przerwań,
— pamięć globalna programu ROM_g,
— pamięć globalna danych RAM_g,
— blok wejścia-wyjścia I/O_g.

Program zawarty jest w pamięci stałej ROM_g wspólnej dla wszystkich mikroprocesorów. Aby w tym samym czasie mogły się bezkonfliktowo wykonywać programy, poszczególne mikroprocesory wyposażono w numery identyfikacyjne. Umożliwia to realizację metody indeksowego adresowania pamięci globalnej przez wspólne i jednakowe dla



Rys. 4. Struktura systemu wieloprocesorowego ze wspólną magistralą

wszystkich procesorów podprogramy obsługi poszczególnych zadań. W rezultacie uzyskano cztery niezależne programowo obszary globalnej pamięci RAM_g. Szczegółowy opis systemu można znaleźć w pracy [3].

Założenia jakie przyjęto do realizacji systemu są następujące:

— istnieje pewien podzbiór źródeł, które generują zgłoszenie do obsługi przez procesor,

— każdy niezajęty procesor może podjąć obsługę dowolnego zgłoszenia,

— sumaryczna intensywność zgłoszeń pojawiających się do obsługi jest za duża na to, aby zgłoszenia te mogły być obsłużone przez pojedynczy mikroprocesor.

Przykładem takiego systemu spełniającego powyższe założenia może być np. sterownik centrali elektronicznej, gdzie zgłoszenia są inicjowane przez abonentów, zaś procesy (programy) stawiane do obsługi przez system mogą być różne, choć są związane z obsługą abonentów. Rozwiązaniem tak określonego zadania może być np. system składający się z kilku autonomicznie pracujących procesorów, z których każdy dysponuje własną biblioteką podprogramów służących do realizacji zadań. Zasadniczym problemem przy takim podejściu jest problem polegający na zapewnieniu współpracy wszystkich sterowanych przez procesor urządzeń, które spełniają rolę urządzeń WE-WY. Problem współpracy urządzeń zewnętrznych znika wtedy, gdy wszystkie te urządzenia komunikują się poprzez wspólną magistralę. Ten ostatni czynnik stał się więc podstawowym argumentem przemawiającym za tym, aby przyjętym rozwiązaniem stał się system wielomikroprocesorowy.

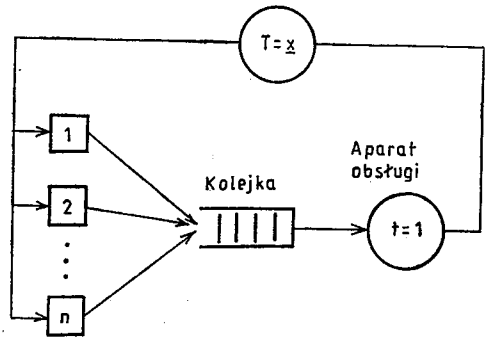
3. ANALIZA PRACY SYSTEMU WIELOMIKROPROCESOROWEGO PRACUJĄCEGO Z PODZIAŁEM CYKLU MASZYNOWEGO

Zgodnie z uwagami podanymi w poprzedniej części pracy, w tej części dokonana zostanie matematyczna analiza pracy systemu wielomikroprocesorowego. W tym celu sprecyzowany zostanie model matematyczny opisujący zachowanie się całego systemu a następnie dokonana zostanie ocena charakterystyk systemu. Do analizy przyjmować będziemy system, którego konfiguracja pokazana została na rys. 3a. Zgodnie z założeniami przyjętymi w poprzedniej części pracy system pracuje w następujący sposób: po pojawieniu się do obsługi przez system wieloprocesorowy zgłoszenia, nadzorczy blok sterujący powołuje do pracy jeden z niezajętych mikroprocesorów. W każdej chwili czasowej w systemie pracować może od 0 do N procesorów. Zgłoszenie pojawiające się w sytuacji, gdy pracuje N procesorów nie mogą być natychmiast obsłużone.

Analiza pracy systemu wielomikroprocesorowego dokonana będzie w sytuacji gdy w systemie pracuje $0 \leqslant n \leqslant N$ procesorów. Ponieważ każdy z pracujących procesorów realizuje pewien proces, więc może on w określonych momentach czasowych żądać dostępu do szyny. Układ arbitra zapewnia dostęp do magistrali głównej systemu tylko jednemu mikroprocesorowi zmuszając pozostałe (żądające dostępu do szyny) do oczekiwania w kolejce na moment przyzwolenia dostępu do magistrali. Modelem matematycznym, który wiernie przybliża zachowanie się rozważanego systemu jest model dyskretnego systemu kolejkowego opisany następującymi założeniami:

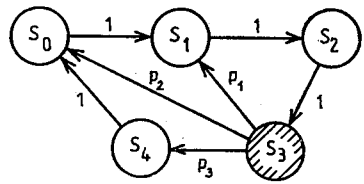
1. liczba źródeł jest skończona i równa n,
2. obsługa każdego zgłoszenia jest synchroniczna i trwa jedną jednostkę czasu,
3. w każdej chwili do obsługi pojawić się może k zgłoszeń, $0 \leqslant k \leqslant n$,
4. w przypadku, gdy do obsługi pojawiło się więcej niż jedno zgłoszenie lub gdy zgłoszenia pojawiają się w sytuacji, gdy w poprzednich chwilach znajdowały się nieobsłużone zgłoszenia, wówczas układ sterujący ustawia pojawiające się zgłoszenia w kolejkę przy czym kolejność zgłoszeń ustawianych w kolejce jest losowana, zaś pierwsze zgłoszenie z kolejki jest pobierane do obsługi,
5. po dokonaniu obsługi, zgłoszenie przebywa w systemie pewien losowy odcinek czasu, a następnie zgłasza się ponownie do obsługi.

Zauważyć należy, że w podanym systemie kolejkowym źródłami zgłoszeń są mikroprocesory żądające dostępu do magistrali, zaś aparatem obsługi jest sama magistrala, gdyż niezależnie od tego, czy określony mikroprocesor żąda dostępu do pamięci, czy też do układu sterowania I/O, zawsze będzie kontaktować się on poprzez magistralę. Z powyższych założeń wynika, że mamy do czynienia z modelem sieci systemów obsługi, w której ponadto pojawiające się zgłoszenia nie są zgłoszeniami pojedynczymi. Model przedstawionego systemu kolejkowego pokazany został na rys. 5. Z założeń dotyczących modelu wynika, że ze względu na czas T przebywania w systemie, który opisany jest zmienną losową X, takiej sieci systemów obsługi nie można w prosty sposób traktować jako systemu markowowskiego. W przypadku gdy zgłoszenia nie przebywają dodatkowo w systemie, ale bezpośrednio po wyjściu z aparatu obsługi wracają do źródła, wówczas taki system kolejkowy daje się łatwo opisać przy pomocy dyskretnego łańcucha Markowa [24, 26]. Rzeczywista praca mikroprocesora Intel 8080 zawiera jeden cykl zegarowy, w którym następuje odwo-



Rys. 5. Model systemu kolejkowego

łanie do pamięci oraz 2, 3 lub 4 dodatkowe cykle zagarowe, w których wykonywane są pewne mikrooperacje wewnątrz procesora. Wynika to po prostu z faktu, że wszystkie cykle maszynowe mikroprocesora Intel 8080 (MCY 7880) mają długość (liczoną w cyklach zegarowych) 3, 4 lub 5. Jeżeli dokona się analizy całej listy rozkazów mikroprocesora 8080 lub analizy programu, który ma być wykonywany, wówczas łatwo będzie można określić rozkład prawdopodobieństwa długości cykli maszynowych. Przyjmijmy, że z analizy programu określone zostały prawdopodobieństwa p_1, p_2, p_3 tego, że cykl maszynowy mikroprocesora wynosi 3, 4 lub 5 odpowiednio. Uwzględniając wówczas ten fakt możemy przyjąć, że praca jednego mikroprocesora opisana jest dyskretnym łańcuchem Markowa,



Rys. 6. Graf opisujący sposób pracy mikroprocesora MCY 7880

którego graf pokazano na rys. 6 zaś macierz przejść P wynikająca z tego grafu jest następująca:

$$P = \begin{array}{c} \\ S_0 \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array} \begin{array}{c} \begin{array}{ccccc} S_0 & S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ p_2 & p_1 & 0 & 0 & p_3 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Ponieważ odczyt i zapis do pamięci jak i urządzeń WE-WY odbywa się w 3 takcie zegarowym stąd można znaleźć średnie prawdopodobieństwo P_3 przebywanie w stanie S_3, którego odwrotność jest ponadto równa średniej długości cyklu maszynowego mikro-

procesora Intel 8080. Z przedstawionego grafu łańcucha Markowa łatwo wyliczyć wielkość P_3, która jest równa:

$$p_3 = \frac{1}{3p_1+4p_2+5p_3}. \tag{1}$$

Z grafu przedstawiającego sposób zachowania mikroprocesora Intel 8080 łatwo określić również zmienną losową X, która opisuje czas przebywania zgłoszenia w systemie. Zmienna X jest dyskretną zmienną losową przyjmującą wartości 2, 3, 4 z prawdopodobieństwami p_1, p_2, p_3, odpowiednio.

Dla celów analizy rozważanego systemu zdefiniowany zostanie obecnie rozszerzony łańcuch Markowa różny od tego, który został określony w pracy [26]. Pozwoli on opisać w wierny sposób zachowanie się badanego systemu wielomikroprocesorowego, w którym pracuje n procesorów. Każdemu stanowi A_i zdefiniowanego łańcucha odpowiada wektor:

$$A \overset{\text{df}}{\equiv} \langle k_0, k_1, k_2, k_3, k_4 \rangle, \tag{2}$$

przy czym:

$$\sum_{j=0}^{4} k_j = n, \tag{3}$$

zaś

k_j — określa liczbę procesorów znajdujących się w stanie S_j (zgodnie z łańcuchem Markowa opisującym zachowanie się μP MCY 7880).

Zakładając, że stan A_i jest określony w postaci (2) prawdopodobieństwa przejścia ze stanu A_i do A_j możemy określić w następujący sposób:

1. dla $k_3 \neq 0$

$$\left.\begin{aligned}
&\text{a) } P(A_i, A_j) = p_1 \Leftrightarrow A_j = \langle k_4, k_0+1, k_3+k_2-1, 0\rangle \\
&\text{b) } P(A_i, A_j) = p_2 \Leftrightarrow A_j = \langle k_4+1, k_0, k_1, k_3+k_2-1, 0\rangle \\
&\text{c) } P(A_i, A_j) = p_3 \Leftrightarrow A_j = \langle k_4, k_0, k_1, k_3+k_2-1, 1\rangle \\
&\text{d) } P(A_i, A_j) = 0, \text{ dla pozostałych } A_j.
\end{aligned}\right\} \tag{4}$$

2. dla $k_3 = 0$

$$\left.\begin{aligned}
&\text{a) } P(A_i, A_j) = 1 \Leftrightarrow A_j = \langle k_4, k_0, k_1, k_3, 0\rangle \\
&\text{b) } P(A_i, A_j) = 0, \text{ dla pozostałych } A_j.
\end{aligned}\right\} \tag{5}$$

Należy zauważyć, że dla tak określonego łańcucha liczba stanów jest znacznie większa od liczby stanów łańcucha określonego w pracy [26], jednakże zaproponowany model systemu wielomikroprocesorowego pozwala uwzględnić zarówno czas T jak i fakt, że zgłoszenia mogą być zgłoszeniami wielokrotnymi (do obsługi mogą pojawiać się jednocześnie nawet 3 zgłoszenia). Liczbę istotnych stanów łańcucha Markowa w zależności od liczby procesorów określa poniższa tabelka:

n	2	3	4	5	6	7	8	9	10
$\{\overline{A_i}\}$	14	28	40	34	34	34	34	34	34

(6)

Dla zdefiniowanego w postaci (2), (4), (5) łańcucha Markowa możliwe jest znalezienie stacjonarnych rozkładów prawdopodobieństw przebywanie w poszczególnych stanach

łańcucha stosując znane metody wyznaczenia prawdopodobieństw. Po to jednak, aby określić żądane charakterystyki rozważanego systemu, potrzebna jest znajomość rozkładu prawdopodobieństwa tego, że w kolejce do obsługi oczekuje s zgłoszeń ($s = 0, 1, 2, \ldots, n$). Prawdopodobieństwo to wyliczyć można stosując następującą zasadę:

$$P_s = \sum_{A_{is}} P(A_{is}), \tag{7}$$

przy czym:

$$A_{is} \overset{\text{df}}{\equiv} \langle k_0, k_1, k_2, s, k_4 \rangle. \tag{8}$$

Znaleziony rozkład prawdopodobieństw P_s liczby zgłoszeń oczekujących na obsługę pozwala oszacować charakterystyki analizowanego systemu. W tym celu skorzystać trzeba z zależności podanych w pracy [26] z tym, że w miejsce P_i podstawić trzeba prawdopodobieństwo P_s.

Dla celów analizy wyznaczone zostały charakterystyki systemu wielomikroprocesorowego dla 3 przypadków różniących się rozkładami p_1, p_2, p_3. Dla przykładowych obliczeń przyjęto następujące rozkłady:

I. $p_1 = 0.4, \quad p_2 = 0.3, \quad p_3 = 0.3$

II. $p_1 = 0.7, \quad p_2 = 0{,}2, \quad p_3 = 0.1$

III. $p_1 = 0.66, \quad p_2 = 0.2, \quad p_3 = 0.14.$

Uzyskane wyniki obliczeń, prawdopodobieństw P_s dla każdego z 3 powyższych przypadków przedstawione zostały w tabelach 1, 2 i 3. Na podstawie uzyskanych prawdopodo-

Tabela 1

Prawdopodobieństwa P_s przebywania w stanach dla rozkładu
I. $p_1 = 0.4$, $p_2 = 0.3$, $p_3 = 0.3$

P_s \ n	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$
P_0	0.501978	0.281882	0.097826	0.00	0.00	0.00
P_1	0.440318	0.528741	0.46761	0.18	0,00	0,00
P_2	0.057707	0.179415	0.387608	0.54	0.18	0.00
P_3	—	0.009962	0.046956	0.28	0,54	0.18
P_4	—	—	0.0000	0.00	0.28	0.54
P_5	—	—	—	0.00	0.00	0.28
P_6	—	—	—	—	0.00	0.00
P_7	—	—	—	—	—	0.00

Tabela 2

Prawdopodobieństwa P_s przebywania w stanach dla rozkładu
II. $p_1 = 0.7$, $p_2 = 0.2$, $p_3 = 0.1$

P_s \ n	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$
P_0	0.428113	0.178570	0.025208	0.00	0.00	0.00
P_1	0.516301	0.614977	0.316724	0.03	0.00	0.00
P_2	0.055585	0.198138	0.630430	0.34	0.03	0.00
P_3	—	0.008315	0.027647	0.63	0.34	0.03
P_4	—	—	0.0000	0.00	0.63	0.34
P_5	—	—	—	0.00	0.00	0.63
P	—	—	—	—	0.00	0.00
P_7	—	—	—	—	—	0.00

Tabela 3

Prawdopodobieństwa P_s przebywania w stanach dla rozkładu
III. $p_1 = 66$, $p_2 = 0.2$, $p_3 = 0.14$

P_s \ n	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$
P_0	0.441803	0.199380	0.037107	0.0000	0.0000	0.0000
P_1	0.500721	0.595095	0.349231	0.0476	0.0000	0.0000
P_2	0.057476	0.197210	0.578194	0.3848	0.0476	0.0000
P_3	—	0.008315	0.035468	0.5676	0.3848	0.0476
P_4	—	—	0.000000	0.0000	0.5676	0.3848
P_5	—	—	—	0.0000	0.0000	0.5676
P_6	—	—	—	—	0.0000	0.0000
P_7	—	—	—	—	—	0.0000

bieństw P_s obliczone zostało prawdopodobieństwo P_c konfliktu według następującej zasady:

$$P_c = \sum_{s=2}^{m} P_s. \tag{9}$$

Odpowiednio dla każdego z trzech przypadków I, II, III otrzymano wyniki, które

Tabela 4

Prawdopodobieństwa P_c konfliktu dla przykładowych 3 rozkładów

P_c \ n	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$
I	0.057707	0.189377	0.434664	0.8200	1	1
II	0.055585	0.206453	0.658077	0.9700	1	1
III	0.057476	0.205525	0.613662	0.9524	1	1

przedstawione zostały w tabeli 4. Z kolei obliczone zostały również średnie długości cyklu maszynowego dla każdej liczby n procesów według następującej zależności [26]:

$$T_{śr(n)} = 3p_1+4p_2+5p_3+\sum_{s=2}^{n} P_s(s-1). \tag{10}$$

Tabela 5

Średnia długość cyklu maszynowego jako funkcja liczby procesorów (dla przykładowych 3 rozkładów)

$T_{śr}$ \ n	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$
I	3.957	4.0993	4.3814	5.000	6.000	7.000
II	3.455	3.6147	4.0857	5.0000	6.000	7.000
III	3.537	3.6938	4.1392	5.0000	6.000	7.000

Średnie długości cyklu maszynowego jako funkcji liczby procesów n są dla przypadków I, II, III przedstawione w tabeli 5. Mając obliczone średnie długości cyklu maszynowego obliczone zostały efektywności wykorzystania jednego procesora według następującej zależności:

$$E = \frac{3p_1+4p_2+5p_3}{T_{śr(n)}}. \tag{11}$$

Wyniki obliczeń umieszczone są w tabeli 6.

Tabela 6

Efektywność wykorzystania jednego procesora pracującego w systemie wielokprocesorowym

E \ n	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$
I	0.9856	0.9514	0.8901	0.7800	0.6500	0.5571
II	0.9841	0.9406	0.8322	0.6800	0.5667	0.4857
III	0.9839	0.9421	0.8407	0.6960	0.5800	0.4971

Tabela 7

Współczynnik multiplikacji mocy obliczeniowej

E \ n	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$
I	1.9712	2.8542	3.5604	3.9	3.9	3.9
II	1.9682	2.8218	3.3288	3.4	3.4	3.4
III	1.9678	2.8262	3.3628	3.48	3.48	3.48

Współczynnik multiplikacji mocy obliczeniowej (*speed up*) można było łatwo obliczyć korzystając z zależności:

$$W_m = n \times E. \tag{12}$$

Dla przykładowych obliczeń otrzymano wyniki przedstawione w tabeli 7. Z wyników uzyskanych w trakcie obliczeń nasuwa się następujący wniosek:

— system wielomikroprocesorowy zbudowany z mikroprocesorów MCY 7880 zwiększa swoje parametry obliczeniowe, w szczególności współczynnik multiplikacji mocy obliczeniowej, dla $n \leqslant 5$.

Z ekonomicznego punktu widzenia wynika, że najbardziej optymalne są systemy składające się z 3 lub 4 mikroprocesorów, gdyż zapewniają one współczynnik wykorzystania każdego z mikroprocesorów (efektywność mikroprocesora) większy od 0.8.

4. WYNIKI SYMULACYJNE

Po to aby stwierdzić na ile bliski rzeczywistości jest przyjęty model, dokonano oceny pracy systemu wielomikroprocesorowego z wykorzystaniem symulacji. Przyjęty model symulacyjny wiernie odtwarza warunki pracy jednomagistralowego systemu wieloprocesorowego przedstawionego na rys. 3 oraz umożliwia badanie zarówno struktury systemu wieloprocesorowego bez magistrali lokalnych (rys. 3a) jak i struktury systemu wieloprocesorowego z magistralami i pamięciami lokalnymi (rys. 3b).

Pojawiające się zgłoszenie podejmowane jest przez mikroprocesor będący w stanie oczekiwania przy czym obsługa zgłoszeń jest realizowana zgodnie z programem umieszczonym w pamięci głównej CM (rys. 3) na zasadzie obsługi przerwań. Wszystkie mikroprocesory wykonują instrukcję zgodnie z ich naturalnym podziałem na poszczególne cykle maszynowe, w których odbywa się odwołanie do pamięci lub układów wejścia-wyjścia, zaś z kolei każdy cykl maszynowy składa się z kilku cykli zegarowych, przy czym w jednym z cykli zegarowych następuje odwołanie mikroprocesora do pamięci (cykl pamięciowy). Na czas jednego cyklu (taktu) zegarowego otwierają się bufory szyny adresowej, danych oraz szyny sterującej w układzie dostępu do magistrali UDM (rys. 3). Pozwala to na odczyt informacji z pamięci głównej lub zapis informacji do pamięci głównej lub układów wejścia-wyjścia. W następnym cyklu zegarowym następuje (o ile to jest potrzebne) otwarcie

buforów szyn dla kolejnego mikroprocesora. Tak jak to wynika z poprzednich części pracy w zbiorze wszystkich cykli maszynowych mikroprocesora MCY 7880 istnieją cykle składające się z trzech, czterech lub pięciu cykli zegarowych, zaś dostęp do magistrali dla mikroprocesora realizowany jest w trzecim takcie zegarowym T_3 każdego z cykli maszynowych. W pozostałych taktach zegarowych realizowane są mikrooperacje w stanach T_1, T_2, T_4 lub T_5, zaś wewnętrzne szyny adresowe danych oraz sterujące mikroprocesora są odcięte od magistrali głównej. Jeśli linia „*Ready*" mikroprocesora jest w stanie niskim i wykonał się stan T_2, wówczas następuje wstrzymanie pracy mikroprocesora w stanie T_w.

Właściwa praca systemów wielomikroprocesorowych ze wspólną magistralą wymaga stosowania układu arbitra sterującego dostępem mikroprocesorów do wspólnej magistrali (wspólnych zasobów), w sytuacji gdy dwa lub więcej mikroprocesory żądają dostępu do magistrali. W zależności od liczby przyłączonych mikroprocesorów oraz intensywności ich obsługi, tworzyć się może kolejka mikroprocesorów żądających dostępu do wspólnej magistrali, którą arbiter obsługuje według regulaminu *„pierwszy wchodzi/pierwszy wychodzi"* z czasem obsługi równym okresowi cyklu zegarowego mikroprocesora.

Ważniejszymi zmiennymi występującymi w programie są:

PR [0:14] — tablica rozkładu prawdopodobieństw występowania poszczególnych typów rozkazów PR_i w programie

TP [1:N] — wygenerowana na podstawie rozkładu występowania poszczególnych typów rozkazów (lub cykli maszynowych) tablica programu, N — maksymalna długość programu w cyklach maszynowych

TR [1:14, 0:5] — tablica typów rozkazów, każdy wiersz odpowiada jednemu z czternastu grup rozkazowych

Rozkazy typu ROM	4 5 4 3 4 3 3 4 3 3 3 4 3 3 3 3 10
Rozkazy typu RAM	4 3 4 3 3 5 3 3 4 3 3 3 4 3 3 3 3 5 3 3 3 3 4 3 3 3 5

Podkreślone cykle maszynowe dotyczą odwołań do pamięci typu RAM.

pr — zmienna „*procesor*" zawierająca informację o tym, które mikroprocesory oczekują na dostęp do magistrali. Bity 0, 1, ..., $n-1$ zmiennej pr odpowiadają mikroprocesorom 1, 2, ..., n. Wystąpienie „1" w którymś z tych bitów oznacza, że odpowiadający mu mikroprocesor oczekuje na dostęp do magistrali. Po zrealizowaniu dostępu wartość „1" jest zastępowana wartością „0"

ar — zmienna „*arbiter*", co najwyżej jeden z bitów 0, 1, ..., $n-1$ zmiennej *ar* może przyjmować wartość „1". Wystąpienie „1" w którymś z tych bitów oznacza, że odpowiadający mu mikroprocesor otrzyma dostęp do magistrali

inr — zmienna określająca długość kolejki

TC [1:n], T [1:n] — tablice cykli zegarowych, opisują stan w jakim znajdują się mikroprocesory, ich wartości zmieniają się wraz z realizowanymi cyklami zegarowymi. Wartości TC_i zmieniają się o 1 w każdym realizowanym cyklu zegarowym w zakresie od:

1 — dla ROM
−5 — dla RAM
11 — dla RAM lokalnego

Wartość $TC_i = 3$ lub -3 odpowiada żądaniu dostępu do magistrali przez i-ty mikroprocesor. Wartości T_i zmieniają się odwrotnie o „−1" począwszy od 3, 4, 5 lub 10 tzn. od ilości cykli zegarowych w danym cyklu maszynowym. $T_i = 0$ oznacza wykonanie danego cyklu maszynowego i przystąpienie do wykonania kolejnego w programie cyklu maszynowego przez jeden z mikroprocesorów.

TN [1:n] — tablica numerów, zawiera numery aktualnie wykonywanych cykli maszynowych.

Ponadto w programie symulacji występuje szereg liczników, w których nalicza się ilość wystąpień cykli maszynowych o określonej długości, ilość odwołań do magistrali, stanu oczekiwania na dostęp do magistrali, itd.

Algorytm symulacji jest następujący:

1. Wczytanie parametrów początkowych
   1.1. Wczytanie tablicy rozkazów TR
   1.2. Wczytanie tablicy rozkładu prawdopodobieństw wystąpienia poszczególnych rozkazów w programie oraz prawdopodobieństwa odwołania się do RAM lokalnego.
2. Generacja programu
   2.1. Losowanie rozkazu
   2.2. Jeśli długość programu nie została przekroczona wówczas następuje wpisanie rozkazu do tablicy programu TP i powtórzenie p. 2.1.
   2.3. Z utworzonego programu TP losowanie cykli maszynowych, w których wystąpi odwołanie do RAM lokalnego
   2.4. Losowanie w tablicy programu TP miejsca początków programu (strat) dla poszczególnych mikroprocesorów.
3. Nadanie wartości początkowych
   3.1. Zerowanie liczników
   3.2. Określenie wartości TC, T w zależności od początkowych cykli maszynowych (p. 2.4.) oraz zwiększenie wartości odpowiednich liczników.

Tabela 8

Prawdopodobieństwa P_s przebywania w stanach (symulacja po cyklach maszynowych)

P_i \ n	2		3		4		5	
	S	M	S	M	S	M	S	M
P_0	0.44225	0.44180	0.19889	0.19938	0.03674	0.03711	0.00004	0.00000
P_1	0.49859	0.50072	0.59654	0.59510	0.35176	0.34923	0.04671	0.04760
P_2	0.05916	0.05748	0.19678	0.1972	0.57505	0.57819	0.38618	0.38480
P_3	—	—	0.00779	0.00831	0.03652	0.03547	0.56701	0.56760
P_4	—	—	—	—	0.00003	0.00000	0.00002	0.00000
P_5	—	—	—	—	—	—	0.00002	0.00000

Tabela 9

Prawdopodobieństwo P_s przebywania w stanie (symulacja po cyklach rozkazowych)

P_i \ n	2		3		4		5	
	S	M	S	M	S	M	S	M
P_0	0.50740	0.50661	0.28775	0.28639	0.09819	0.09760	0.000054	0.00000
P_1	0.43481	0.43735	0.53061	0.53393	0.50292	0.49986	0.17422	0.17250
P_2	0.05779	0.05604	0.17277	0.17098	0.35348	0.36056	0.59120	0.59500
P_3	—	—	0.00907	0.0870	0.04548	0.04198	0.23392	0.23250
P_4	—	—	—	—	0.00003	0.0000	0.00012	0.00000
P_5	—	—	—	—	—	—	0.00000	0.00000

bywania w poszczególnych stanach. Wyznaczono również dokładny rozkład długości cykli maszynowych jaki występuje w realizowanym programie symulacyjnym. Obliczone z symulacji prawdopodobieństwa p_1, p_2, p_3 przyjęto następnie jako dane do analizy systemu z wykorzystaniem zaproponowanego modelu. Uzyskano następujące wyniki (S — wyniki uzyskane z symulacji, M — wyniki analityczne uzyskane dla przyjętego modelu), które przedstawione są w tabelach 8 i 9:

— symulacja po cyklach maszynowych (tabela 8):
$p_1 = 0.66 \quad p_2 = 0.20 \quad p_3 = 0.14$

— symulacja po cyklach rozkazowych (tabela 9):
$p_1 = 0.31 \quad p_2 = 0.44 \quad p_3 = 0.25$

4. Wykonanie symulacji
   4.1. j otrzymuje wartość *inr* mikroprocesorów aktualnie oczekujących na dostęp do pamięci
   4.2. Zwiększenie zawartości licznika dla j odwołań do magistrali
   4.3. Jeśli $j = 0$, $ar = 0$ i skok do 4.7.
   4.4. Z „*kolejki*" pobierany jest numer $j1$ pierwszego mikroprocesora żądającego dostępu do pamięci, jego odwołanie zostanie teraz zrealizowane, pozostałym mikroprocesorom zwiększa się czas oczekiwania na dostęp do pamięci
   4.5. $ar = 2f(j1-1)$
   4.6. $pr = pr-ar$
   w pr zostaje usunięta „1" odpowiadająca $j1$ mikroprocesorowi, numer $j1$ mikroprocesora zostaje usunięty z kolejki
   4.7. Zmiana stanu TC, T dla mikroprocesorów nie oczekujących na dostęp do magistrali $T_i = T_i - 1$
   4.7.1. Jeśli $T_i = 0$, tzn. cały cykl maszynowy został zrealizowany, mikroprocesor przechodzi do wykonania kolejnego cyklu maszynowego (jak w p. 3.2.)
   4.7.2. Jeśli $T_i \neq 0$ to $TC_i = TC_i + 1$
   4.7.3. Jeśli po zmianie T_i $TC_i = 3$ lub $TC_i = -3$ wówczas do odpowiedniego bitu pr wprowadzona będzie „1". Dla mikroprocesorów, które w tym cyklu zegarowym po raz pierwszy żądały dostępu do pamięci ($TC_i = 3$ lub $TC_i = -3$) losowana jest kolejność przydziału magistrali i w takiej kolejności są dopisywane do kolejki mikroprocesory oczekujące na dostęp do pamięci
   4.8. Skok do p. 4.1.

Wykonanie p. 4. może zostać przerwane na specjalne żądanie np. gdy przynajmniej jeden z mikroprocesorów wykona cały program.

Symulację realizowano dla dwóch przypadków:

— dla przypadku, w którym wygenerowano zgodnie z określonym rozkładem ciąg cykli maszynowych,
— dla przypadku, w którym wygenerowano zgodnie z określonym rozkładem ciąg cykli rozkazowych, z których każdy składa się z kilku cykli maszynowych.

W pierwszym przypadku tablica programu tworzona była na podstawie wypełnienia jej odpowiednimi cyklami maszynowymi o określonej długości przy czym kolejne cykle były statystycznie niezależne. W rzeczywistości jednak realizacja programu na mikroprocesorze wymaga wykonania ciągu rozkazów, w których w określonej kolejności pojawia się sekwencja poszczególnych cykli maszynowych. Lista rozkazów mikroprocesora MCY 7880 zawiera w sobie 10 różnych grup rozkazów, a zatem można postawić pytanie, na ile deterministyczny czynnik decydujący o tym, iż po pewnym cyklu maszynowym wystąpi w cyklu rozkazowym inny z góry znany cykl maszynowy, spowoduje różnicę w uzyskanych wynikach symulowanego systemu i analizowanego modelu.

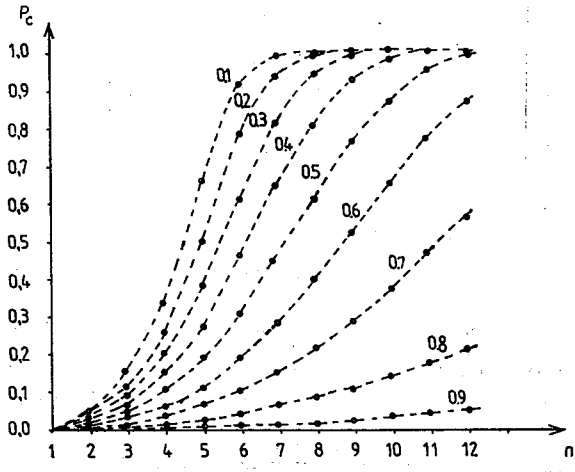
Porównania wyników uzyskanych w trakcie symulacji i testowania modelu dokonano w następujący sposób:

dla wygenerowanego ciągu cykli maszynowych lub rozkazowych przeprowadzono symulację zgodnie z założeniami dotyczącymi sposobu pracy modelu zaproponowanego systemu wielomikroprocesorowego, wyznaczając z symulacji prawdopodobieństwa prze-

Zwiększenie liczby procesorów powyżej 5-ciu nie wnosi nic nowego, gdyż uzyskane wyniki symulacyjne są bardzo zbliżone do tych, które otrzymano na drodze analitycznej dla analizowanego modelu.

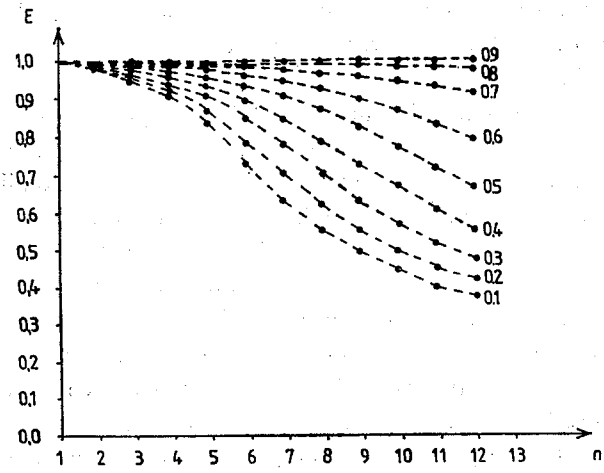
Porównanie wyników symulacji i obliczeń modelowych pozwala stwierdzić, że przyjęty model matematyczny wiernie odtwarza sposób pracy systemu wielomikroprocesorowego. Ponadto, co jest bardzo ważne, analiza pracy takiego systemu wielomikroprocesorowego może być przeprowadzana tylko w warunkach znajomości rozkładów długości cykli maszynowych, bez konieczności dokonywania głębszej analizy występujących w cyklach rozkazowych zależności statystycznych między cyklami maszynowymi.

Ponieważ systemy wielomikroprocesorowe pracujące ze wspólną magistralą mogą być realizowane tylko dla małej liczby procesorów, w trakcie symulacji analizowanych systemów, zdecydowano się na przeprowadzenie dodatkowej symulacji systemów wielomikroprocesorowych wyposażonych w pamięci lokalne (rys. 3b). Systemy takie, co prawda bardziej złożone i kosztowne, posiadają wszystkie zalety systemów ze wspólną magistralą i zapewniają lepsze parametry użytkowe. Analiza symulacyjna systemów z pamięciami lokalnymi dokonana została przede wszystkim pod kątem efektywności pracy mikroprocesora, współczynnika przyspieszenia obliczeń oraz prawdopodobieństwa konfliktu. Symulację przeprowadzono przy założeniu, że generowane były z określonym prawdopodobieństwem cykle maszynowe mikroprocesorów MCY 7880, oraz, że z żądanym prawdopodobieństwem wyznaczone były w trakcie symulacji cykle z odwołaniem do magistrali lokalnej. O tym, że zdecydowano się tylko na dokonanie analizy symulacyjnej tych systemów zdecydował fakt, iż łańcuch Markowa opisujący działanie modelu takiego systemu posiadał bardzo dużą liczbę stanów (powyżej 200 stanów), co utrudniało wyznaczenie stacjonarnych prawdopodobieństw przebywania w poszczególnych stanach łańcucha.

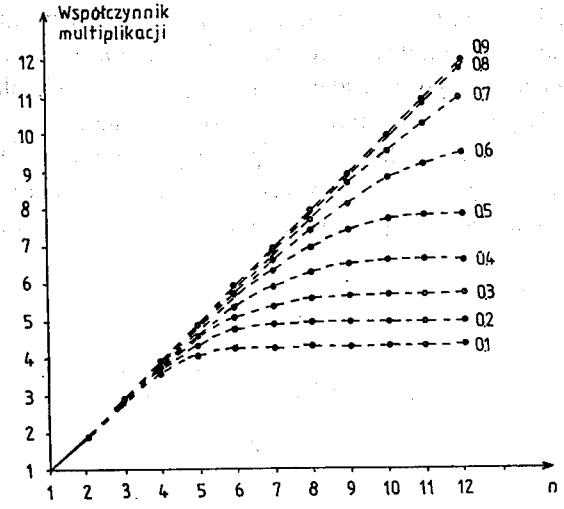


Rys. 7. Przebieg określający wartość prawdopodobieństwa konfliktu jako funkcja liczby procesorów

Uzyskane wyniki przedstawione zostały na rys. 7, 8, 9. Na rys. 7 przedstawiony został wykres prawdopodobieństwa konfliktu jako funkcji liczby procesorów, przy założeniu różnych prawdopodobieństw odwołania się do pamięci lokalnej. Odrzucone zostały dwa



Rys. 8. Przebieg funkcji określającej wielkość efektywności wykorzystania mikroprocesora w zależnośc od liczby procesorów



Rys. 9. Przebiegi funkcji współczynnika multiplikacji obliczeń w zależności od liczby procesorów

skrajne przypadki: przypadek przedstawiający pracę bez pamięci lokalnych tzn. gdy prawdopodobieństwa odwołania do magistrali lokalnej było równe 0 oraz przypadek pracy systemu bez pamięci głównej, wtedy gdy prawdopodobieństwo odwołania do magistrali lokalnej było równe 1. Z kolei na rys. 8 przedstawione zostały przebiegi efektywności wykorzystania mikroprocesora jako funkcji liczby procesorów w systemie, zaś na rys. 9 przebiegi współczynnika multiplikacji obliczeń. Z uzyskanych wyników widać, że dla prawdopodobieństw odwołania do pamięci lokalnej równych i większych od 0.5 znacznej poprawie ulegają wszystkie rozważane parametry systemu. W szczególności szybko rośnie współczynnik multiplikacji obliczeń, który dla prawdopodobieństw odwołań do pamięci lokalnych zawartych między 0.7 a 0.9 rośnie prawie liniowo w zakresie do $n = 12$. Sposób

pracy systemu wielomikroprocesorowego, w którym prawdopodobieństwo odwołania do magistrali głównej jest niewielkie, rzędu $0.1 \div 0.2$ ma miejsce wtedy, gdy w pamięci lokalnej umieszczone są wszystkie programy oraz większość danych, z których korzysta się w trakcie realizacji procesu obliczeniowego, zaś poszczególne mikroprocesory kontaktują się poprzez magistralę główną tylko z urządzeniami WE-WY oraz pobierają niektóre dane z pamięci głównej. Sposób taki jest jak widać bardzo opłacalny, gdyż współczynnik wykorzystania każdego z procesorów jest bliski 1.

5. PODSUMOWANIE I WNIOSKI

W pracy omówiony został system wielomikroprocesorowy zrealizowany na mikroprocesorze MCY 7880. Po omówieniu ogólnych problemów związanych z realizacją dowolnych systemów wieloprocesorowych, bliżej zajęto się omówieniem sposobu pracy systemu ze wspólną magistralą. Z kolei zaproponowano rozwiązanie systemu wieloprocesorowego, w którym wykorzystano specyfikę pracy mikroprocesora MCY 7880. Następnie podano model matematyczny opisujący pracę zaproponowanego systemu. Dla wybranych rozkładów prawdopodobieństw przedstawione zostały wyniki obliczeń. Ostatnia część pracy poświęcona została porównaniu uzyskanych wyników analitycznych z wynikami otrzymanymi na podstawie symulacji zaproponowanego systemu. Z porównania wyników uzyskanych metodą symulacyjną i analityczną wysunąć można wniosek, że przyjęty model matematyczny odtwarza rzeczywisty sposób pracy rozważanego systemu.

W końcu przedstawiono również wyniki symulacyjne systemu wielomikroprocesorowego pracującego z pamięciami lokalnymi.

Z uzyskanych wyników widać, że opłacalna liczba procesorów pracujących w systemach wielomikroprocesorowych bez pamięci lokalnych jest równa 4, zaś w systemach pracujących z pamięciami lokalnymi liczba mikroprocesorów może wynosić nawet kilkanaście przy założeniu, że prawdopodobieństwo odwołania do magistrali głównej jest niewielkie, rzędu $0.1 \div 0.3$. Efektywność wykorzystania mikroprocesora jest wówczas nie mniejsza niż 0.8. Dodatkowo należy tutaj dodać, że próby przeprowadzone na rzeczywistym modelu systemu potwierdzają wyniki uzyskane zarówno na drodze analitycznej jak i symulacyjnej.

BIBLIOGRAFIA

1. D. P. Bhandarkar: *Some Performance Issues in Multiprocessor System Design.* IEEE Trans. on Computers, vol C-26, May 1977
2. R. Dirska, A. Jurkiewicz, A. Zabłudowski: *Arbiter systemu wielomikroprocesorowego z podziałem czasu na szynach.* Materiały IV Krajowej Konferencji Naukowo-Technicznej „Zastosowanie mikroprocesorów w automatyce i pomiarach", Warszawa 13.09.1984, 162—167
3. R. Dirska, A. Jurkiewicz, A. Zabłudowski: *Komputer wielomikroprocesorowy.* prace PIE, *Elementy półprzewodnikowe i układy scalone.* 1986, z. 4
4. P. H. Enslow: *Multiprocessor organization — A Survey.* Computing Surveys, 1977, vol. 9, No 1, pp. 103—129

5. P. M. Enslow: *Systemy cyfrowe wieloprocesorowe.* Warszawa: WNT, 1978
6. E. T. Fathi, M. Krieger: *Multiple Microprocessor Systems: What, Why, end When.* Computer 1983, vol. 16, No 3, pp. 23—32
7. K. T. Fung, M. C. Torng: *On the Analysis of Memory Conflicts and Bus Contention in Multiple Microprocessor System.* IEEE Trans. on Computers. 1979, śvol. C-28, No 1
8. A. Gago, J. Biscarri: *Low — Cost Multiprocessor System.* Electronics Letters. November, 1981, vol. 17, No 24, pp. 924—925
9. S. Hoener, W. Roehder: *Efficiency of a Multi Microprocessor System with The Shared Busses.* Euromicro, 1977, pp. 35—42
10. Y. Hoffner: *Construction of an Inexpensive and Flexible Multiprocessor System* Mircoprocessosr and Microsystems 1983, vol. 7, No 2, pp. 111—116
11. V. Hoffner, M. F. Smith: *Communication Between Two Microprocessors Though Common Memory.* Microprocessors and Microsystems, 1982, vol. 6, No 6, pp. 303—308
12. G. H. Hoogendoorn: *A General Model for Memory Interference in Multiprocessors.* IEEE Trans. on Computers 1977, vol. C-26 No 10 Oct., pp. 998—1005
13. A. B. Kovaleski: *High — Speed Bus Arbiter for Multiprocessors.* IEE Procceding March, 1983, vol. 130, PtE, No 2, pp. 49—56
14. J. Kriz: *A Queueing Analysis of a Symmetric Multiprocessor with Shared Memories and Biusses.* IEE Procedings 1983, vol. 130, PtE., No 3, pp. 83—89
15. T. Lang, M. Valero, I. Alegre: *Bandwidth of Crossbar and Multiple — Bus Connections for Multiprocessors.* IEEE Trans. on Computers, 1982, vol C-31, No 12, pp. 1227—1234
16. P. Markenscoff: *A Deterministic Model for Evaluating the Performace of a Multiple Processor System with a Shared Bus.* IEEE Trans. on Computers, 1983, vol. C-33, No 3, pp. 281—285
17. M. A. Marsan, F. Gregoretti: *Memory Interference Models for a Multi — Microprocessor Systems with a Shared Bus and a Single External Common Memory.* Microprocessing and Microprogramming, 1981, vol. 7, pp. 124—133
18. M. A. Marsan, M. Gerla: *Markov Models for Multiple Bus Multiprocessor Systems,* IEEE Trans. on Computers, 1982, vol. C-31, No 3, pp. 239—248
19. M. A. Marsan, G. Balbo, G. Conte: *Compatarive Performance Analysis of Single Bus Multiprocessor Architectures.* IEEE Trans. on Computers, 1982, vol. C-31, No 12, pp. 1179—1191
20. M. A. Marsan, G. Balbo, G. Conte, F. Gregoretti: *Modelling Bus Contention and Memory Interference in a Multiprocessor System.* IEEE Trans. on Computers, 1983, vol. C-32, No 1, pp. 60—72
21. J. C. C. Nelson, M. K. Refei: *Design of a Hardware Arbiter for Multimicroprocessor Systems.* Microprocessors and Microsystems, 1984, vol. 8, No 1, pp. 21—24
22. B. P. Rau: *Interleaved Memory Bandwidth in a Model of a Multiprocessor.* IEEE Trans. on Computers, 1979, vol. C-28, No 9, pp. 678—681
23. A. S. Sethi, N. Deo: *Interference in Multiprocessor Systems with Localized Memory Access Probabilites.* IEEE Trans. on Computers, 1979, vol. C-28, No 2, pp. 157—163
24. K. O. Siomolas, B. A. Bowen: *Performance of Cross — Bar Multiprocessor Systems.* IEEE Trans. on Computers, 1983, vol. C-32, No 7, pp. 689—695
25. D. W. L. Yen, J. H. Patel, E. S. Davidson: *Memory Interference in Synchronous Multiprocessor Systems.* IEEE Trans. on Computers, 1982, vol. C-31, No 11, pp. 116—1121
26. A. Zabłudowski, A. Jurkiewicz: *Analysis of Multi — Microprocessor System with Time — Shared Bus.* Microprocessing and Microprogramming, 1984, vol. 13, pp. 171—177
27. A. Zabłudowski, R. Dirska, A. Jurkiewicz: *Prosty system wielomikroprocessorowy,* Materiały Konferencji Naukowo-Technicznej „Mikrokomputery w Automatyce i Technice Systemów" Wrocław, 18—21, wrzesień 1984, s. 84—91

A. ZABŁUDOWSKI, R. DIRSKA, A. JURKIEWICZ

MULTIMICROPROCESSOR SYSTEM WITH TIME—SHARED BUS

Summary

The analysis of a multimicroprocessor system built — up with the microprocessor chips Intel 8080 is given. One of the purposes of this paper was the presentation of the problems arising in multimicroprocessor systems. The logical and physical structure of multiprocessor systems is described. A multiprocessor system with time — shared bus and, particulary, the arbiter unit are more closely discussed. The mathematical model of a multimicroprocessor system and the results obtained both by computation and simulation are presented in the two last parts of the paper.

A. ZABŁUDOWSKI, R. DIRSKA, A. JURKIEWICZ

SYSTÈME MULTIMICROPROCESSEUR MONOBUS

Résumé

Dans cet article un système multimicroprocesseur monobus, construit de microprocesseurs MCY 7880, est exposé. La conception générale du système est présentée et son modèle mathématique est donné. Le modèle décrit l'action du système réel et permet de désigner ses caractéristiques indispensables.

A. ZABŁUDOWSKI, R. DIRSKA, A. JURKIEWICZ

DAS EINMAGISTRALEN-VIELMIKROPROZESSORENSYSTEM

Zusammenfassung

In diesem Artikel wird ein aus den Mikroprozessoren MCY 7880 gebautes Einmagistralen-Vielmikroprozessorensystem dargestellt. Die allgemeine Konzeption für die Lösung des Systems wird geschildert sowie sein mathematisches Modell gebracht. Das Modell beschreibt die Wirkungsweise eines realen Systems, und gestattet, seine notwendigen charakteristischen Größen zu bestimmen.

А. ЗАБЛУДОВСКИ, Р. ДИРСКА, А. ЮРКЕВИЧ

ОДНОМАГИСТРАЛЬНАЯ МНОГОМИКРОПРОЦЕССОРНАЯ СИСТЕМА

Резюме

Представлена многомикропроцессорная система построенная при использовании микропроцессоров МСУ —7880 — общая концепция компоновки, действительная схема и математическая модель описывающая действия этой схемы, позволяющая определить необходимые характеристики системы. Труд подразделен на пять частей. В первой дефинированы многопроцессорная система, представлена логическая организация системы, приведена и обсуждена характеристика расчетной мощности системы. Во второй части обсуждены системы с общей магистралью. Третья часть описывает действительную реализацию системы использующей естественное подразделение машинного цикла микропроцессора на часовые циклы. Четвертая и пятая часть посвящены анализу предложенного решения многомикропроцессорной системы. Предложенная математическая модельявляется цепью Маркова в которой состояния представлены пятьюкомпонентным вектором определяющим количества процессоров находящихся в отдельных часовых циклах микропроцессора МСУ—7880. Сравнены результаты полученные аналитическим и симуляционным методами с результатами полученными симуляционным методом работающих с одной магистралью и главной памятью, добавочными магистралями и местными памятьями.

SPIS TREŚCI

CONTENTS — TABLE DES MATERIES — INHALT

СОДЕРЖАНИЕ

WYTYCZNE DLA AUTORÓW

Komitet Redakcyjny prosi o przestrzeganie następujących wytycznych przy przygotowywaniu maszynopisów artykułów nadsyłanych do opublikowania.

1. *Tematyka i charakter artykułów.* Redakcja przyjmuje do druku prace przeglądowe, komplikacyjne i monograficzne, wchodzące w zakres szeroko pojętej elektroniki, które powinny jednak zawierać własny wkład twórczy Autora polegający na: oryginalnym ujęciu zagadnienia, własnej klasyfikacji, krytycznej ocenie (teorii lub metod), wyciągnięciu wniosków co do celowości takiego lub innego działania, prognostyce itp. Autorów obowiązuje jak najdalej posunięta zwięzłość.

2. *Wymagania podstawowe.* Artykuły należy nadsyłać w maszynopisie, w dwóch egzemplarzach, w zasadzie w języku polskim; dopuszczalne są jednak również artykuły w językach: angielskim, francuskim, niemieckim i rosyjskim. Maszynopis powinien być napisany jednostronnie przez czarną taśmę, na maszynie do pisania z niezabrudzonymi i nieuszkodzonymi znakami. Dopuszcza się odręczne czytelne uzupełnianie tekstu atramentem lub długopisem kolorem czarnym lub ciemnoniebieskim znaków specjalnych oraz znaków w językach, których alfabetów nie ma na maszynach do pisania, np. znaków matematycznych, chemicznych, liter greckich. Maszynopis powinien być napisany na papierze do maszyny do pisania koloru białego, formatu A4; numeracja ciągła na wszystkich stronach.

3. *Sposób pisania tekstu.* Tekst w maszynopisie powinien być pisany bez używania wyróżnień, a w szczególności nie dopuszcza się spacjowania (rozstrzelenia), podkreślania i pisania tekstów dużymi literami, z wyjątkiem wyrazów, które umownie pisze się dużymi literami (np. FORTRAN). Proponowane wyróżnienia Autor może zaznaczyć w maszynopisie (zwykłym ołówkiem) za pomocą przyjętych znaków adiustacyjnych (podkreślenie linią przerywaną — spacjowanie, podkreślenie linią ciągłą — pogrubienie, podkreślenie wężykiem — kursywa itp.). Na jednej stronie maszynopisu powinno być 30 wierszy po około 60 znaków łącznie z odstępami. Marginesy każdej strony powinny mieć następujące wymiary: górny — ok. 25 mm, lewy — ok. 35 mm. Tekst maszynopisu powinien być napisany z podwójnym odstępem między wierszami; tytuły i podtytuły małymi literami. Akapity należy rozpoczynać z wcięciem równym trzem uderzeniom maszyny do pisania.

4. *Sposób pisania tablic.* Tablice powinny być napisane w układzie zbliżonym do układu zecerskiego. Tytuły rubryk pionowych i poziomych powinny być napisane małymi literami z podwójnym odstępem między wierszami. Przypisy (notki) dotyczące tablic należy pisać bezpośrednio pod tablicą. Tablice należy numerować kolejno liczbami arabskimi; u góry każdej tablicy podać tytuł. Tablice umieścić na końcu maszynopisu.

5. *Sposób pisania wzorów matematycznych.* Rozmieszczenie znaków, cyfr, liter i odstępów powinno być zbliżone do rozmieszczenia elementów druku. Wskaźniki i wykładniki potęg powinny być napisane wyraźnie i być prawidłowo obniżone lub podwyższone w stosunku do linii wiersza podstawowego. Znaki nad literami i cyframi: strzałki, linie, kropki, daszki itp. powinny być napisane dokładnie nad tymi elementami, do których się odnoszą. Numery wzorów należy umieszczać z prawej strony.

6. *Przygotowanie materiału ilustracyjnego.* Rysunki, wykresy i fotografie należy wykonywać zgodnie z obowiązującymi normami Polskiego Komitetu Normalizacji, Miar i Jakości (oznaczonymi literą E i numeracją od 01200 do 01245; przydatna może się okazać książka K. Michela, T. Sapińskiego: „Rysunek techniczny elektryczny"), na oddzielnych arkuszach, z podaniem kolejnych numerów rysunków. W maszynopisie artykułu na marginesie, obok właściwego tekstu, należy podać jedynie odnośny numer rysunku, a na oddzielnym arkuszu wykaz podpisów pod rysunki. Wszystkie rysunki, wykresy i fotografie należy nazywać w tekście rysunkami (skrót: rys.). U samego dołu rysunku (a przy fotografiach na odwrocie) należy wpisać czytelnie numer rysunku, tytuł pracy i nazwisko autora. Ostateczne wykonanie rysunków obowiązuje Redakcję.

7. *Streszczenia.* Do każdej nadsyłanej pracy należy dołączyć krótkie streszczenie (analizę) w języku polskim (w 5 egz.) oraz streszczenie (w 2 egz.) w językach obcych: angielskim, francukim,

niemieckim i rosyjskim. W razie niemożności przygotowania streszczeń w językach obcych Autor powinien podać przynajmniej terminy obcojęzyczne niezbędne do wykonania tłumaczenia.

8. *Bibliografia.* Na końcu maszynopisu należy podać w przyjętej przez Autora kolejności (np. chronologicznej, alfabetycznej itp.) wykaz publikacji, na które Autor w tekście się powołuje, lub które uważa za słuszne wymienić z innych powodów. W każdej pozycji wykazu należy podać w następującej kolejności: pierwsze litery imion, nazwisko autora, po czym po przecinku pełny tytuł dzieła lub artykułu; dalej, w przypadku książki — wydawcę, miejsce wydania i rok, a w przypadku artykułu — tytuł czasopisma, numer zeszytu, rok wydania i ewent. numer strony. Pozycje wykazu powinny być ponumerowane.

9. *Informacje dodatkowe.*

— Niezastosowanie się Autora do podanych wyżej wytycznych pociągnie za sobą konieczność potrącenia z honorarium autorskiego kosztów związanych z doprowadzeniem dostarczonych materiałów do postaci wymaganej przez Redakcję.

— Autorowi przysługuje bezpłatnie 25 egz. odbitek pracy. Dodatkowe egzemplarze Autor może zamówić w Redakcji na własny koszt.

— **Autora obowiązuje korekta autorska, którą należy zwracać w ciągu 3 dni** pod adresem Redakcji.

— Redakcja prosi Autorów o podawanie miejsca pracy i adresu prywatnego, a także o powiadamianie o zmianie adresu.

NOTATKI

NOTATKI

NOTATKI

NOTATKI

Cena zł 250,—

Rozprawy Elektrotechniczne

Kwartalnik

Prenumeratę na kraj przyjmują i informacji o cenach udzielają urzędy pocztowe i doręczyciele na wsi oraz Oddziały RSW „Prasa-Książka-Ruch" w miastach.

Prenumeratę ze zleceniem wysyłki za granicę przyjmuje RSW „Prasa-Książka-Ruch", Centrala Kolportażu Prasy i Wydawnictw, ul. Towarowa 28, 00-958 Warszawa, konto PBK XIII Oddział w Warszawie Nr 370044-1195-139-11. Wysyłka za granicę pocztą zwykłą jest droższa od prenumeraty krajowej o 50% dla zleceniodawców indywidualnych i o 100% dla zlecających instytucji i zakładów pracy.

Terminy przyjmowania prenumerat na kraj i za granicę:
— do dnia 10 listopada na I półrocze roku następnego i na cały rok następny,
— do dnia 1 czerwca na II półrocze roku bieżącego.

Bieżące i archiwalne numery można nabyć lub zamówić we Wzorcowni Ośrodka Rozpowszechniania Wydawnictw Naukowych PAN, Pałac Kultury i Nauki, 00-901 Warszawa.

Subscription orders for all the magazines published in Poland available through the local press distributors or directly through the Foreign Trade Enterprise ARS POLONA, 00-068 Warszawa, Krakowskie Przedmieście 7, Poland.

Our bankers:

BANK HANDLOWY WARSZAWA S.A.

Rozpr. Elektrot. T. 34, z. 3, 705—1040, Warszawa 1988

Indeks 37483

POLSKA AKADEMIA NAUK
KOMITET ELEKTRONIKI I TELEKOMUNIKACJI

PLISSN 0035-9386

ROZPRAWY ELEKTROTECHNICZNE

TOM XXXIV — ZESZYT 4

KWARTALNIK

PAŃSTWOWE WYDAWNICTWO NAUKOWE
WARSZAWA 1988

POLSKA AKADEMIA NAUK
KOMITET ELEKTRONIKI I TELEKOMUNIKACJI

ROZPRAWY ELEKTROTECHNICZNE

TOM XXXIV — ZESZYT 4

KWARTALNIK

PAŃSTWOWE WYDAWNICTWO NAUKOWE
WARSZAWA 1988

ADRES REDAKCJI
00-665 Warszawa, ul. Nowowiejska 15/19
Politechnika, Instytut Telekomunikacji, Gmach Elektroniki, pok. 401

Dyżury Redakcji: środy i piątki, godz. 11—13
tel. 28 89 81 lub 210 07-737

Telefony domowe: redaktora naczelnego — 12 17 65
zastępcy redaktora naczelnego — 26 83 41
sekretarza odpowiedzialnego — 25 29 18

PAŃSTWOWE WYDAWNICTWO NAUKOWE
Warszawa, Miodowa 10

Nakład 510 (420+90)	Oddano do składania 21.X.1988 r.
Ark. wyd. 23,5 Ark. druk. 17,5+wkł.	Podpisano do druku w lipcu 1989 r.
Papier offset. kl. III 70 g, 70×100	Druk ukończono w sierpniu 1989 r.
Zamówienie nr 6854/12/88. A-77	Cena zł 250,—

Drukarnia im. Rewolucji Październikowej, Warszawa

ROZPRAWY ELEKTROTECHNICZNE, 1988, 34, z. 4, ss. 1043–1050

621.316

Zastosowanie minimaksowej procedury optymalizacyjnej do identyfikacji warunków brzegowych

JAN SIKORA

Instytut Elektrotechniki Teoretycznej i Miernictwa Elektrycznego, Politechnika Warszawska

RADOSŁAW PYTLAK

Instytut Automatyki, Politechnika Warszawska

Otrzymano 1987.03.13

Autoryzowano do druku 1987.12.04

W pracy przedstawiono zastosowanie minimaksowej minimalizacji funkcji celu do identyfikacji warunków brzegowych. Uzyskane wyniki porównano z wynikami uzyskanymi przy pomocy minimalizacji średniokwadratowej (LSP).

1. WSTĘP

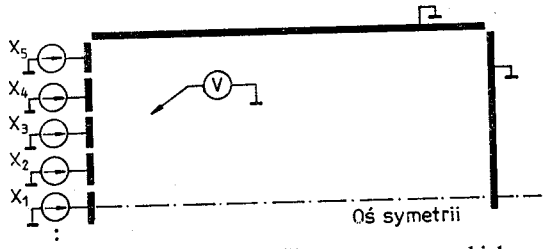
Znaczenie pojęcia wrażliwości w teorii układów znacznie wzrosło od czasów H. W. Bodego. Ważnym etapem w rozwoju metod wrażliwościowych są prace M. Bychowskiego, który wykazał, że wrażliwość można wyznaczyć pośrednio, bez różniczkowania, metodami analizy sieci.

W ostatnich latach daje się zauważyć zastosowanie metod wrażliwościowych w syntezie układów aktywnych *RC*. Jako pierwsi I. Ishizaki i H. Watanabe [3] zaproponowali wykorzystanie szeregów Taylora i wrażliwości w projektowaniu układów. Znacznie mniejsze zastosowanie wrażliwości ma miejsce w dziedzinie projektowania układów o parametrach rozłożonych. Ścisły związek zagadnień polowych i obwodowych, który wykazał Zienkiewicz [7], umożliwia zastosowanie opracowanych już metod w teorii pola. Przykładem wzrastającego zainteresowania tymi metodami mogą być prace [2, 6]. W niniejszej pracy zaproponowano algorytm minimaksowej optymalizacji z liniowymi ograniczeniami do identyfikacji warunków brzegowych w oparciu o metodę wrażliwościową.

2. MODEL MATEMATYCZNY IDENTYFIKACJI WARUNKÓW BRZEGOWYCH

Rozpatrzmy problem identyfikacji warunków brzegowych w obszarze przedstawionym na Rys. 1.

Zmianę funkcji stanu (potencjał) $\varphi = \{\varphi_1, \ldots, \varphi_i, \ldots, \varphi_m\}^T$ w otoczeniu nominalnych wartości parametrów układu można przedstawić w postaci szeregu Taylora



Rys. 1. Obszar w którym identyfikowane są warunki brzegowe

$$\varphi_i - \varphi_{0i} = \sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial x_j} (x_j - x_{0j}) + R_n \tag{1}$$

gdzie pochodne cząstkowe rzędu wyższego niż pierwszy wchodzą w skład reszty R_n. Dokonując linearyzacji wyrażenia (1) przez pominięcie pochodnych cząstkowych wyższego rzędu i zapisując wyrażenie (1) w postaci macierzowej dla obszaru z rys. 1 otrzymamy

$$\begin{bmatrix} \frac{\partial \varphi_1}{\partial x_1} & \cdots & \frac{\partial \varphi_1}{\partial x_n} \\ \cdots & & \cdots \\ \frac{\partial \varphi_m}{\partial x_1} & \cdots & \frac{\partial \varphi_m}{\partial x_n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \varphi_1 \\ \vdots \\ \varphi_m \end{bmatrix} \tag{2}$$

gdzie w przypadku ogólnym $m \geqslant n$. Macierz współczynników układu (2) jest macierzą wrażliwości (macierzą Jacobiego) i znajomość jej jest konieczna w każdej procedurze optymalizacyjnej opartej na metodach gradientowych.

Z powyższych rozważań wynika że należy znaleźć rozwiązanie układu równań (2) o postaci

$$y = Ex \tag{3}$$

gdzie przez y oznaczono wektor funkcji stanu w wybranych punktach obszaru, natomiast E jest macierzą $m \times n$, $m \geqslant n$ oraz rząd $E < n$.

Ponieważ rozwiązanie układu (3) nie musi istnieć, jesteśmy zainteresowani w znalezieniu takiego wektora x, dla którego $||y - Ex||$ będzie najmniejsza, tzn.

$$\tilde{x} \in \operatorname{Argmin} ||y - Ex|| \tag{4}$$

Rozwiązanie tego zadania może być niejednoznaczne. Zawężamy więc zbiór rozwiązań poprzez nałożenie na wektor x ograniczeń o postaci

$$Ax \leqslant b \tag{5}$$

Postępowanie nasze uzasadnione jest również tym, że w istocie poszukujemy minimum lokalnego na zbiorze pożądanych wartości x, wywołujących efekt bliski y. W rezultacie mamy następujące zadanie

$$\min_{Ax \leqslant b} ||y - Ex||. \tag{6}$$

Aby określić do końca zadanie optymalizacji, musimy jeszcze sprecyzować normę z jaką mamy do czynienia w (6). Na podstawie rozważań zawartych w [1] wybrano normę Czebyszewa (l_∞), która obok normy l_1 jest najbardziej odpowiednia do rozwiązywania zadania aproksymacji [1].

Jeżeli

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}, \qquad E = \begin{bmatrix} E_1 \\ \vdots \\ E_m \end{bmatrix},$$

gdzie E_i są wierszami o wymiarze n, to wówczas nasze zadanie przedstawia się następująco

$$\min_{x \in R^n} \max_{1 \leqslant i \leqslant m} |y_i - \langle E_i, x \rangle| \tag{7}$$

przy ograniczeniach: $Ax \leqslant b$ gdzie $\langle \cdot, \cdot \rangle$ jest iloczynem skalarnym, tzn. poszukujemy minimum funkcjonału $F_l(x) = \max\limits_{1 \leqslant i \leqslant m} |y_i - \langle E_i, x \rangle|$ przy ograniczeniach liniowych.

Do rozwiązania zadania (7) wykorzystano algorytm, przedstawiony w pracy [4, 5] dla zadań minimaksowych o postaci

$$\min_{x \in R^n} \max_{1 \leqslant i \leqslant m} f_i(x) \tag{8}$$

$$A^1 x \leqslant b^1, \qquad A^2 x = b^2$$

gdzie funkcje f_i są różniczkowalne w sposób ciągły. Obecnie przedstawimy modyfikację tego algorytmu dla zadania

$$\min_{x \in R^n} \max_{1 \leqslant i \leqslant m} |f_i(x)| = \min_{x \in R^n} f(x) \tag{9}$$

gdzie jak poprzednio $F(x) = \max\limits_{1 \leqslant i \leqslant m} |f_i(x)|$ przy ograniczeniach $Ax \leqslant b$ oraz przy tych samych założeniach odnośnie funkcji f_i. Przypuśćmy, że $\varepsilon > 0$, $L > 0$, $\lambda, \alpha \in (0, 1)$ oraz, że na $k-1$ iteracji algorytmu obliczono wektor x^{k-1}. Wówczas na k-tej iteracji algorytmu należy wykonać następujące operacje:

Krok 1: określ zbiór indeksów ξ — aktywnych: $M_\varepsilon(x^{k-1}) = \{1 \leqslant i \leqslant m: |f_i(x)^{k-1}| \geqslant F(x^{k-1}) - \varepsilon\}$ w zależności od tego, czy $F(x^{k-1}) > \xi$ względnie $F(x^{k-1}) \leqslant \xi$ rozwiąż jedno z dwóch zadań:

A. jeżeli $F(x^{k-1}) > \xi$

$$\min_{\substack{(\sigma, \delta x) \in R^{n+1} \\ \|\delta x\|_\infty \leqslant L}} \sigma$$

$$\tilde{f}_i(x^{k-1}) + \langle \nabla \tilde{f}_i(x^{k-1}), \delta x \rangle \leqslant \sigma, \qquad i \in M_\xi(x^{k-1}) \tag{10}$$

$$A(x^{k-1} + \delta x) \leqslant b$$

przy czym $\tilde{f}_i(\cdot) = +f_i(\cdot)$, gdy $f_i(x^{k-1}) > 0$ oraz $\tilde{f}_i(\cdot) = -f_i(\cdot)$, gdy $f_i(x^{k-1}) < 0$

B. jeżeli $F(x^{k-1}) \leqslant \xi$

$$\min_{(\sigma, \delta x) \in R^{n+1}} \sigma$$

$$\begin{aligned} f_i(x^{k-1})+\langle \nabla f_i(x^{k-1}), \delta x\rangle &\leqslant \sigma \\ -f_i(x^{k-1})-\langle \nabla f_i(x^{k-1}), \delta x\rangle &\leqslant \sigma, \quad i \in M_\varepsilon(x^{k-1}) \\ A(x^{k-1}+\delta x) &\leqslant b \end{aligned} \tag{11}$$

Konieczność uwzględnienia dwóch zadań programowania liniowego wynika z faktu występowania wartości bezwzględnej w definicji funkcjonału $F(x)$. Jeżeli wartość funkcjonału F w punkcie x^{k-1} jest mniejsza lub równa ξ, to przy określeniu nowego kierunku należy uwzględnić zarówno funkcje $f_i(x^{k-1})$ jak i $-f_i$ dla $i \in M_\xi(x^{k-1})$.
Niech rozwiązaniem zadania (10), względnie (11), będzie wektor $(\sigma^k, \delta x^k)$.

Krok 2: określ $\delta^k = \sigma^k - f(x^{k-1})$, a następnie znajdź najmniejszą liczbę naturalną $k = 0, 1, 2, \ldots$ taką, że

$$F(x^{k-1}+\lambda^k \cdot \delta x^k) - F(x^{k-1}) \leqslant \alpha \delta^k \cdot \lambda^k \tag{12}$$

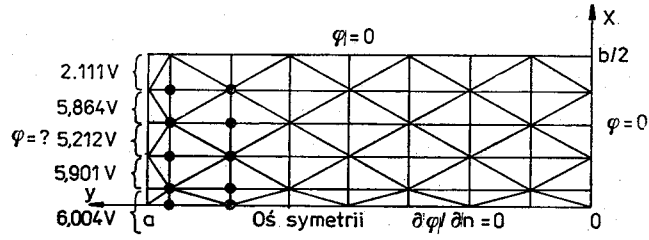
podstaw $x^k = x^{k-1} + \lambda^{k_{\min}} \delta x^k$, zwiększ k o jeden.
Analogicznie jak w [4] można pokazać, że opisany algorytm jest globalnie zbieżny, tzn. każdy punkt skupienia ciągu $\{x^k\}_0^\infty$ generowanego przez algorytm spełnia warunki konieczne optymalności. Dobierając parametr ε można zapewnić, że powyższy algorytm z powodzeniem rozwiązuje zadanie (4) również w przypadku, gdy rząd $E < n$.

3. PRZYKŁAD OBLICZENIOWY

Opracowany algorytm zastosowano w dwóch przykładach. Przykłady te różnią się między sobą modelem matematycznym zastosowanym do obliczenia macierzy wrażliwości. Pierwszy z nich oparty jest na metodzie elementów skończonych, a więc macierz wrażliwości liczona jest metodami numerycznymi, drugi zaś na metodzie rozdzielenia zmiennych, a więc macierz wrażliwości liczona jest w sposób analityczny.
W obu przypadkach wyniki porównano z metodą minimalizacji średniokwadratowej.

3.1. PRZYKŁAD PIERWSZY

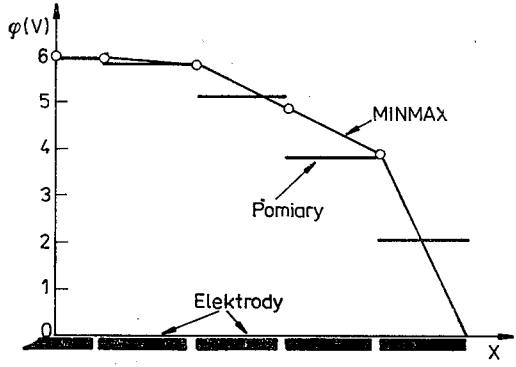
Rozważany obszar przedstawiony na Rys. 1 pokryto siatką elementów skończonych w ten sposób aby węzły sieci pokryły się z punktami pomiarowymi bazy danych (wektor $\{\varphi_1, \ldots, \varphi_m\}^T$) z równania (2).
Sieć elementów skończonych obszaru z zaznaczonymi punktami pomiarowymi oraz brzegiem, na którym poszukiwany jest warunek brzegowy, przedstawiono na Rys. 2. Dodatkowo na Rys. 2. zaznaczono wartości identyfikowanego potencjału brzegowego, który został zmierzony już po dokonaniu identyfikacji.
Macierz wrażliwości wyznaczono metodą perturbacji numerycznej [6]. Wyniki identyfikacji przedstawiono w tabeli 1.
Elektroda reprezentująca brzeg, na którym identyfikowany jest potencjał, składa się z odcinków, na których panuje stały potencjał (patrz Rys. 2). Dlatego też w celu lepszej interpre-



Rys. 2. Dyskretyzacja obszaru, warunki brzegowe i punkty pomiarowe

Tabela 1

Wartości potencjału w węzłach brzegowych	x_1	x_2	x_3	x_4	x_5
Minimalizacja minimaksowa	6,137	6,016	5,904	4,949	4,037
Minimalizacja średniokwadratowa	5,688	6,041	5,868	4,873	3,830



Rys. 3. Porównanie wyników minimalizacji minimaksowej z pomiarami: przykład 1

tacji uzyskanych wyników porównanie eksperymentu numerycznego z pomiarami przedstawiono na Rys. 3.

3.2. PRZYKŁAD DRUGI

W przykładzie tym do aproksymacji poszukiwanego potencjału brzegowego zastosowano aproksymację opartą na metodzie rozdzielenia zmiennych. Wyrażenie na wartość potencjału wewnątrz obszaru (po uwzględnieniu zadanych warunków brzegowych na części brzegu) ma postać

$$\varphi(x, y) = \sum_{n=1}^{\infty} A_n \sinh \frac{(2n-1)\pi y}{b} \sin \frac{(2n-1)\pi x}{b} \tag{13}$$

gdzie b — szerokość obszaru (Rys. 2).

Jak wynika z równania (7) nieznany potencjał brzegowy poszukiwany będzie pośrednio poprzez szukanie współczynników rozkładu na szereg Fouriera. Aby rozmiary zadania były porównywalne z rozmiarami zadania w przykładzie pierwszym, ograniczymy się tylko do pięciu ($n = 5$) wyrazów szeregu Fouriera. Ze względu na symetrię zadania poszukiwane będą jedynie nieparzyste wyrazy rozwinięcia w szereg Fouriera.

Macierz wrażliwości dla bazy pomiarowej $\varphi = (\varphi_1, \ldots, \varphi_{m=10})^T$, w tym przypadku można wyrazić analitycznie

$$\begin{bmatrix} \sinh\frac{1\pi y_1}{b}\sin\frac{1\pi x_1}{b}, & \ldots, & \sinh\frac{(2n-1)\pi y_1}{b}\sin\frac{(2n-1)\pi x_1}{b} \\ \vdots & & \vdots \\ \sinh\frac{1\pi y_m}{b}\sin\frac{1\pi x_m}{b}, & \ldots, & \sinh\frac{(2n-1)\pi y_m}{b}\sin\frac{(2n-1)\pi x_m}{b} \end{bmatrix} \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix} = \begin{bmatrix} \varphi_1(x_1, y_1) \\ \vdots \\ \varphi_m(x_m, y_m) \end{bmatrix} \tag{14}$$

gdzie x_i, y_i — współrzędne punktów pomiarowych.

Ze względu na arytmetykę komputera MERA-400 koniecznym okazało się przeskalowanie macierzy wrażliwości przez wprowadzenie nowej zmiennej

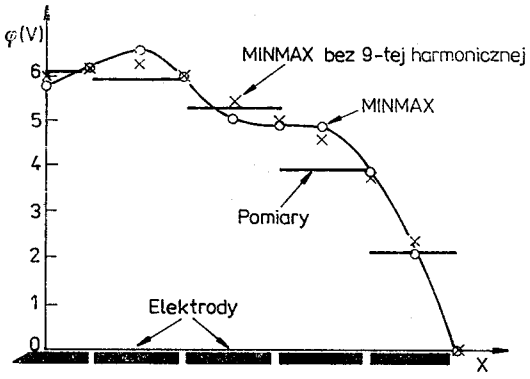
$$A'_n = A_n \sinh\frac{(2n-1)\pi a}{b} \tag{15}$$

Wyniki identyfikacji przedstawiono w tabeli 2.

Tabela 2

Wartości współczynników rozwinięcia w szereg Fouriera	A'_1	A'_2	A'_3	A'_4	A'_5
Minimalizacja minimaksowa	6,9556	1,0737	0,4309	0,3503	−0,3478
Minimalizacja średniokwadratowa	6,8107	0,9784	0,3097	0,3490	0,2E-11
Wartości potencjału w wybranych zgodnie z przykładem 1 punktach brzegowych	x_1	x_2	x_3	x_4	x_5
Minimalizacja minimaksowa	5,61	6,08	5,95	4,91	3,96
Minimalizacja średniokwadratowa	5,79	5,94	5,93	4,78	3,71

Porównanie wyników numerycznych z pomiarami napięcia na odcinkach elektrody przedstawiono na Rys. 4.



Rys. 4. Porównanie wyników minimalizacji minimaksowej z pomiarami: przykład 2

ZAKOŃCZENIE

Przedstawiony w pracy algorytm minimalizacji minimaksowej do identyfikacji warunków brzegowych w przedstawionych przykładach daje zachęcające rezultaty.

Na podstawie przedstawionych rezultatów wydaje się zbyt wczesną próba oceny, który z zastosowanych algorytmów jest lepszy. Zadania identyfikacji czy też syntezy są na tyle trudne, że z praktycznego punktu widzenia, dobrze jest rozwiązać to samo zadanie dwoma różnymi metodami, dokonując minimalizacji funkcji celu w różnych normach. Zaproponowany algorytm można bez trudu przystosować do rozwiązywania zadań optymalizacji w normie l_1.

BIBLIOGRAFIA

1. R. Fletcher: *Practical methods of optimization.* J. Wiley and Sons, 1981
2. Z. Gurdal, R. T. Haftka: *Sensitivity derives for static test loading boundary conditions.* AIAA Journal, Jan. 1985, vol. 23, No 1
3. I. Ishizaki, H. Watanabe: *An iterative method for network design as a non-linear programing probleme.* Nippon Electric Company Ltd. Japan, Monograph DEB 485, 1964
4. R. Pytlak: *Minimaksowe zadanie sterowania optymalnego z czasem dyskretnym: własności rozwiązań i algorytmy.* Rozprawa doktorska. Wydział Elektroniki PW, 1986
5. R. Pytlak, K. Malinowski: *Numerical Methods for the Minimax Optimal in Lecture Notes on Control and Information Sciences.* Springer-Verlag, 1986, Vol. 83, (ed. J. L. Lions and A. Bensoussan)
6. J. Sikora: *Analiza wrażliwościowa a metoda wymuszeń jednostkowych: perspektywy jej zastosowania do zagadnień odwrotnych.* IX SPETO, Gliwice—Wisła, 23—26 IV 1986
7. O. C. Zienkiewicz: *Metoda elementów skończonych.* Warszawa: Arkady, 1972

J. SIKORA, R. PYTLAK

APPLICATION OF MIN-MAX OPTIMISATION PROCEDURE TO IDENTIFICATION OF BOUNDARY CONDITIONS

Summary

Application of MIN-MAX procedure to identification of the boundary conditions is presented. Obtained results are compared with the results achieved by means of the least squares procedure.

J. SIKORA, R. PYTLAK

APPLICATION DE LA PROCEDURE MINIMAXIMALE OPTIMALE À L'IDENTIFICATION DES CONDITIONS AUX LIMITES

Résumé

Ce travail présente l'application de la minimisation minimaximale à l'identification des conditions aux limites. Les résultats ainsi obtenus ont été comparés avec ceux de la minimisation moyenne quadratique.

J. SIKORA, R. PYTLAK

ANWENDUNG DER MINIMAX-OPTIMIERUNG ZWECKS IDENTIFIZIERUNG DER RANDBEDINGUNGEN

Zusammenfassung

In diesem Aufsatz wurde die Identifizierung der Randbedingungen dargestellt. Die Optimierung, der Zielfunktion wurde mittels der MINIMAX durchgeführt. Die Ergebnisse wurden mit der mittels der Least Squares (LSP) gewonnenen Optimierung verglichen.

Я. СИКОРА, Р. ПЫТЛЯК

ПРИМЕНЕНИЕ МИНИМАКСНОЙ ОПТИМИЗАЦИОННОЙ ПРОЦЕДУРЫ ДЛЯ ИДЕНТИФИКАЦИИ КРАЕВЫХ УСЛОВИЙ

Резюме

Представлено применение минимаксной минимизации функции цели идентификации краевых условий. Полученные результаты сравнены с результатами полученными с помощью среднеквадратичной минимизации.

ROZPRAWY ELEKTROTECHNICZNE, 1988, 34, z. 4, ss. 1051–1063

621.372.8

Zastosowanie metody elementów skończonych do analizy pól elektromagnetycznych w bezstratnych falowodach jednorodnych

KAROL ANISEROWICZ

Wydział Elektryczny, Politechnika Białostocka

Otrzymano 1987.09.17

Autoryzowano do druku 1988.01.05

Przedstawiono analizę pól elektromagnetycznych w bezstratnych falowodach jednorodnych z użyciem metody elementów skończonych. Do obliczeń zastosowano ośmiowęzłowe elementy izoparametryczne drugiego rzędu. Powstałe macierzowe zagadnienie własne rozwiązano metodą iteracji podprzestrzennych. Zamieszczono przykładowe rysunki rozkładów pól.

1. WSTĘP

Rozwój współczesnej techniki mikrofalowej powoduje wzrost zainteresowań falowodami o złożonych kształtach przekroju poprzecznego, umożliwiających uzyskanie specjalnych właściwości (np. niskiej częstotliwości granicznej rodzaju podstawowego pola elektromagnetycznego, odpowiednio szerokiego pasma częstotliwości występowania tylko tego rodzaju). Prace badawcze i projektowe nad nimi często opierały się głównie na danych eksperymentalnych. Istotnym wsparciem dalszego rozwoju było zastosowanie numerycznych metod obliczeniowych, a szczególnie metody elementów skończonych (MES).

Podstawy teoretyczne i algorytmy obliczeniowe zastosowań MES do analizy falowodów zostały opublikowane m.in. w pracach [1, 3, 15], jednakże dalej prowadzone są badania dotyczące obrania optymalnego sposobu rozwiązywania rozpatrywanego zagadnienia brzegowego. Dotyczy to szczególnie znalezienia odpowiedniej metody rozwiązywania powstałego w trakcie tej analizy macierzowego zagadnienia własnego, a także wyboru właściwego rodzaju elementów skończonych. Opracowany pakiet programów powinien umożliwić dokładne i szybkie obliczanie oraz przedstawianie graficzne rozkładów pól w falowodach o dowolnych kształtach przekroju poprzecznego. Optymalizacja obliczeń jest bardzo istotna, gdyż liczba operacji komputerowych występujących przy poszukiwaniu wartości i wektorów własnych macierzy jest znacznie większa, niż przy rozwiązywaniu układu równań tego samego rzędu.

Głównym problemem przy analizie danego typu falowodu jest określenie częstotliwości granicznej i rozkładu pola elektromagnetycznego rodzaju podstawowego. Istotne jest przy tym wyznaczenie pasma częstotliwości, w którym może istnieć tylko rodzaj pod-

stawowy. Niekiedy potrzebne jest obliczenie częstotliwości granicznych i rozkładów pól kilku następnych rodzajów. Równoczesne wyznaczenie kilku pierwszych par własnych (wartości własnych i wektorów własnych) macierzowego problemu własnego $[K]\{\Phi\} = \lambda[M]\{\Phi\}$ jest najbardziej efektywne przy użyciu metody iteracji podprzestrzennych [4], która nie była jeszcze znana autorom wyżej wymienionych prac.

Na podstawie przeglądu literatury dotyczącej MES zdecydowano zastosować krzywoliniowe, ośmiowęzłowe elementy izoparametryczne drugiego rzędu.

Dalsze rozważania dotyczą falowodów cylindrycznych o dowolnym, stałym przekroju poprzecznym. Zakłada się liniowość i bezstratność ośrodka wypełniającego ich wnętrze, doskonałą przewodność elektryczną ścianek, a także brak źródeł pola, ładunków i prądów wymuszonych. Ścianki są równoległe do osi 0z globalnego układu współrzędnych kartezjańskich. Założono sinusoidalną zależność pól od czasu.

2. MODEL MATEMATYCZNY ROZKŁADU POLA W FALOWODZIE

Rozkład pola elektromagnetycznego w falowodach bezstratnych otrzymuje się po rozwiązaniu równania Helmholtza

$$\nabla^2\Phi + \omega^2\varepsilon\mu\Phi = 0, \tag{1}$$

przy czym Φ spełnia jednorodne warunki brzegowe Dirichleta (dla fal typu E) lub Neumanna (dla fal typu H) [5, 8, 11]. Jeśli poszukujemy rozkładów pól typu E, to szukana funkcja Φ może być interpretowana jako elektryczny potencjał Hertza lub składowa wzdłużna E_z natężenia pola elektrycznego, zaś dla pól typu H — jako magnetyczny potencjał Hertza lub składowa wzdłużna H_z natężenia pola magnetycznego. Rozwiązania równania (1) poszukuje się w postaci

$$\Phi(x, y, z) = u(x, y)e^{-j\beta_z z}, \tag{2}$$

gdzie β_z jest stałą fazową falowodu w kierunku osi 0z. Pozwala to sprowadzić równanie (1) do zagadnienia dwuwymiarowego

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \lambda u = 0, \tag{3}$$

gdzie

$$\lambda = \omega^2\varepsilon\mu - \beta_z^2. \tag{4}$$

Funkcja $u(x, y)$ realizuje ekstremum funkcjonału

$$F(u) = \frac{1}{2}\int_{\Omega}\int \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 - \lambda u^2\right] \mathrm{d}x\,\mathrm{d}y, \tag{5}$$

przy czym obszar płaski Ω oznacza powierzchnię przekroju poprzecznego falowodu. Obszar Ω dzielony jest na podobszary Ω^e (elementy skończone). Na rys. 1 przedstawiono przyjęte do obliczeń ośmiowęzłowe elementy izoparametryczne drugiego rzędu.

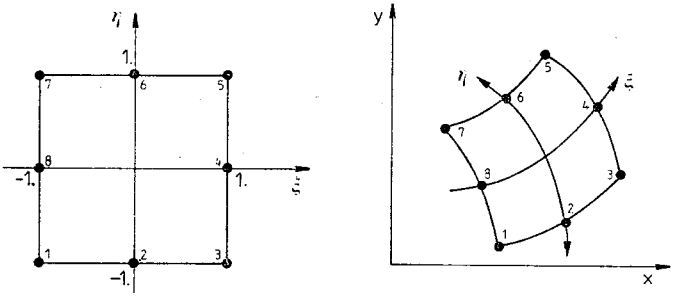
Funkcje kształtu zastosowanych elementów są określone następująco [7, 18]:

$$N_j(\xi,\eta) = \frac{1}{4}(1+\xi\xi_j)(1+\eta\eta_j)(\xi\xi_j+\eta\eta_j-1) \quad \text{dla } \xi_j = \pm 1,\ \eta_j = \pm 1,$$

$$N_j(\xi,\eta) = \frac{1}{2}(1-\xi^2)(1+\eta\eta_j) \quad \text{dla } \xi_j = 0,\ \eta_j = \pm 1, \tag{6}$$

$$N_j(\xi,\eta) = \frac{1}{2}(1-\eta^2)(1+\xi\xi_j) \quad \text{dla } \xi_j = \pm 1,\ \eta_j = 0,$$

gdzie ξ_j, η_j są współrzędnymi lokalnymi węzła j ($j = 1, \ldots, 8$).



Rys. 1. Ośmiowęzłowy element izoparametryczny drugiego rzędu w lokalnym (a) i globalnym (b) układzie współrzędnych

Elementy te umożliwiają dobrą aproksymację rozkładu obliczanej wielkości oraz dobre zobrazowanie krzywizny brzegu analizowanego obszaru (w falowodach — zwykle odcinki linii prostych lub krzywe stożkowe). Ponadto, przy opracowaniu graficznym wyników, stosunkowo prosty algorytm wykreślania krzywoliniowych warstwic rozkładu pola wewnątrz podobszarów Ω^e można ułożyć jedynie dla wielomianów aproksymacyjnych drugiego stopnia.

Wewnątrz elementu skończonego rozkład szukanej wielkości przedstawia się za pomocą liniowej kombinacji funkcji kształtu

$$u^e(\xi,\eta) = \sum_j N_j u_j, \tag{7}$$

gdzie: u_j — wartości funkcji u w węzłach j elementu e.

Siatka dyskretyzacji obszaru Ω zawiera n węzłów. Wartości węzłowe u_i ($i = 1, \ldots, n$) obliczane są z warunku koniecznego istnienia ekstremum funkcjonału, jakim jest zerowanie się pochodnych $\frac{\partial F}{\partial u_i}$. Oznaczając składnik funkcjonału pochodzący od elementu e przez F^e, otrzymuje się

$$\frac{\partial F^e}{\partial u_i} = \sum k_{ij}u_j - \lambda \sum_j m_{ij}u_j, \quad i = 1, \ldots, n, \tag{8}$$

gdzie

$$k_{ij} = \iint_{\Omega^e} \left(\frac{\partial N_i}{\partial x}\frac{\partial N_j}{\partial x} + \frac{\partial N_i}{\partial y}\frac{\partial N_j}{\partial y} \right) \mathrm{d}x\,\mathrm{d}y, \tag{9}$$

$$m_{ij} = \iint_{\Omega^e} N_i N_j \mathrm{d}x\,\mathrm{d}y. \tag{10}$$

Ze sformułowania izoparametrycznego wynikają zależności geometryczne podobne do (7), pozwalające całki (9) i (10) wyrazić we współrzędnych lokalnych [18]

$$x = \sum_j N_j x_j, \quad y = \sum_j N_j y_j. \tag{11}$$

We wzorach tych (x_j, y_j) są współrzędnymi globalnymi węzła j.

Po przekształceniach otrzymuje się

$$k_{ij} = \int_{-1}^{1}\int_{-1}^{1}\left[\left(\frac{\partial N_i}{\partial \xi}\frac{\partial y}{\partial \eta} - \frac{\partial N_i}{\partial \eta}\frac{\partial y}{\partial \xi}\right)\left(\frac{\partial N_j}{\partial \xi}\frac{\partial y}{\partial \eta} - \frac{\partial N_j}{\partial \eta}\frac{\partial y}{\partial \xi}\right) + \right.$$
$$\left. + \left(-\frac{\partial N_i}{\partial \xi}\frac{\partial x}{\partial \eta} + \frac{\partial N_i}{\partial \eta}\frac{\partial x}{\partial \xi}\right)\left(-\frac{\partial N_j}{\partial \xi}\frac{\partial x}{\partial \eta} + \frac{\partial N_j}{\partial \eta}\frac{\partial x}{\partial \xi}\right)\right]\frac{\mathrm{d}\xi\,\mathrm{d}\eta}{\det[\boldsymbol{J}]}, \tag{12}$$

$$m_{ij} = \int_{-1}^{1}\int_{-1}^{1} N_i N_j \det[\boldsymbol{J}]\,\mathrm{d}\xi\,\mathrm{d}\eta, \tag{13}$$

przy czym

$$\frac{\partial x}{\partial \xi} = \sum_j \frac{\partial N_j}{\partial \xi} x_j, \quad \frac{\partial x}{\partial \eta} = \sum_j \frac{\partial N_j}{\partial \eta} x_j,$$
$$\frac{\partial y}{\partial \xi} = \sum_j \frac{\partial N_j}{\partial \xi} y_j, \quad \frac{\partial y}{\partial \eta} = \sum_j \frac{\partial N_j}{\partial \eta} y_j, \tag{14}$$

$$[\boldsymbol{J}] = \begin{bmatrix} \dfrac{\partial x}{\partial \xi} & \dfrac{\partial y}{\partial \xi} \\ \dfrac{\partial x}{\partial \eta} & \dfrac{\partial y}{\partial \eta} \end{bmatrix}. \tag{15}$$

Całkowanie (12), (13) jest przeprowadzane numerycznie, z wykorzystaniem kwadratur Gaussa-Legendre'a [4, 7, 14, 18].

Po zsumowaniu wpływu wszystkich elementów według zależności

$$\frac{\partial F}{\partial u_i} = \frac{\partial\left(\sum_e F^e\right)}{\partial u_i} = \sum_e \frac{\partial F^e}{\partial u_i} = 0, \quad i = 1, \ldots, n, \tag{16}$$

zostaje zbudowany układ równań

$$[\boldsymbol{K}]\{\Phi\} = \lambda[\boldsymbol{M}]\{\Phi\}. \tag{17}$$

Warunki brzegowe Dirichleta uwzględnia się przez znaną modyfikację macierzy $[\boldsymbol{K}]$ i $[\boldsymbol{M}]$, opisaną np. w pracach [7, 14. 18]. Warunki brzegowe Neumanna spełniane są automatycznie.

3. ALGEBRAICZNE ZAGADNIENIE WARTOŚCI WŁASNYCH

Równanie macierzowe (17) stanowi tzw. uogólnione (rozszerzone) algebraiczne zagadnienie własne. Istnieje wiele metod rozwiązywania tych zagadnień. Rozwiązanie problemu własnego jest równoważne z obliczeniem pierwiastków wielomianu charakterystycznego

$$g(\lambda) = \det([\boldsymbol{K}] - \lambda[\boldsymbol{M}]) = 0, \tag{18}$$

którego stopień jest równy rzędowi n macierzy $[\boldsymbol{K}]$ i $[\boldsymbol{M}]$. Nie istnieją ogólne reguły ścisłego wyznaczania pierwiastków wielomianu, gdy jego stopień jest większy od 4, zatem metody rozwiązania problemu (17) są wyłącznie iteracyjne. Algorytmy podstawowych metod można podzielić na następujące grupy [4]:

— metody iteracji wektorowych (iteracja prosta, odwrotna);
— metody przekształceń ortogonalnych (Jacobiego, Givensa, Householdera);
— metody rozwiązywania wielomianu charakterystycznego (jawne, uwikłane);
— metody oparte na właściwościach ciągów Sturma.

Duże znaczenie w wielu metodach ma wykorzystanie właściwości ilorazu Rayleigha. Jest też duża grupa technik kombinowanych, łączących powyższe algorytmy. Różnorodność ich wynika z faktu, że nie istnieje dotąd metoda, która zawsze prowadziłaby do efektywnego rozwiązania. Wybór metody zależy od specyfiki danego problemu, jak np. od rozmiarów i struktur macierzy, liczby poszukiwanych par własnych, uwarunkowania numerycznego zagadnienia.

Do wyznaczania pojedynczych par własnych najdogodniejsze są metody iteracji wektorowych. Analiza pełnego widma macierzy jest wykonywana metodami przekształceń ortogonalnych. Najwięcej zagadnień wiąże się z obliczaniem kilku par własnych wewnątrz określonego pasma widma. Metody oparte na analizie wielomianu charakterystycznego lub właściwościach ciągów Sturma są efektywne jedynie dla macierzy o niedużych wymiarach.

Rozważane w niniejszej pracy macierze są średniej wielkości i mogą mieć szerokie pasmo elementów niezerowych. Poszukiwanych jest zwykle kilka najmniejszych wartości własnych i odpowiadających im wektorów własnych. Aktualnie optymalną metodą rozwiązywania takiego zagadnienia jest metoda iteracji podprzestrzennych [4], należąca do grupy metod kombinowanych. Jest ona rozwinięciem analizy Rayleigha-Ritza, opisanej np. w pracach [4, 10, 17]. Zakłada się w niej, że macierze $[\boldsymbol{K}]$ i $[\boldsymbol{M}]$ są symetryczne i dodatnio określone. Podstawowy algorytm metody iteracji podprzestrzennych wyznaczania p pierwszych par własnych $(\lambda_i, \{\Phi_i\})$, $i = 1, \ldots, p$, składa się z trzech podstawowych etapów:

1. Wyznaczenie zbioru q wektorów początkowych

$$[\Phi_0] = [\{\varphi_1^{(0)}\}, \{\varphi_2^{(0)}\}, \ldots, \{\varphi_p^{(0)}\}, \ldots, \{\varphi_q^{(0)}\}], \tag{19}$$

przyjmując wskaźnik iteracji $k = 0$ oraz $\lambda_1^{(0)} = \lambda_2^{(0)} = \ldots = \lambda_p^{(0)} = 1$, przy czym $q > p$ w celu polepszenia zbieżności obliczeń.

2. Zastosowanie równoczesnej wektorowej iteracji odwrotnej dla q wektorów oraz

analizy Rayleigha-Ritza, w celu uzyskania p najlepszych aproksymacji wartości i wektorów własnych. Składają się na to następujące operacje:

— obliczenie wstępnej aproksymacji q wektorów własnych $[\bar{\Phi}_{k+1}]$ z zależności

$$[K][\bar{\Phi}_{k+1}] = [M][\Phi_k]; \tag{20}$$

— obliczenie macierzy tzw. zredukowanego zagadnienia własnego (o wymiarach $q \times q$)

$$[\tilde{K}_{k+1}] = [\bar{\Phi}_{k+1}]^T[K][\bar{\Phi}_{k+1}], \tag{21}$$

$$[\tilde{M}_{k+1}] = [\bar{\Phi}_{k+1}]^T[M][\bar{\Phi}_{k+1}]; \tag{22}$$

— rozwiązanie zagadnienia własnego w q-wymiarowej podprzestrzeni, z wykorzystaniem np. metody Jacobiego

$$[\tilde{K}_{k+1}][Q_{k+1}] = [\tilde{M}_{k+1}][Q_{k+1}][\Lambda_{k+1}]; \tag{23}$$

— obliczenie aproksymacji wektorów własnych

$$[\Phi_{k+1}] = [\bar{\Phi}_{k+1}][Q_{k+1}]; \tag{24}$$

— sprawdzenie zbieżności obliczeń według kryterium

$$\left| \frac{\lambda_i^{(k+1)} - \lambda_i^{(k)}}{\lambda_i^{(k+1)}} \right| \leqslant \varepsilon, \quad i = 1, 2, \ldots, p, \tag{25}$$

gdzie ε jest zakładaną dokładnością wyniku (zwykle $\varepsilon = 10^{-6}$). Jeśli nie osiągnięto zakładanej dokładności, to powtarza się obliczenia (20)—(25).

3. Po stwierdzeniu zbieżności w poprzednim etapie — określenie błędu wyznaczenia wektorów własnych oraz sprawdzenie, czy rzeczywiście obliczono poszukiwane pary własne. Błąd wyznaczenia i-tego wektora własnego obliczany jest jako iloraz

$$\sigma_i = \frac{\|[K]\{\varphi_i^{(l)}\} - \lambda_i^{(l)}[M]\{\varphi_i^{(l)}\}\|_2}{\|[K]\{\varphi_i^{(l)}\}\|_2}, \tag{26}$$

gdzie (l) oznacza indeks ostatniej iteracji, zaś $\|\varphi\|_2$ — normę euklidesową wektora $\{\varphi\}$ w n-wymiarowej przestrzeni rzeczywistej

$$\|\varphi\|_2 = \sqrt{\sum_{j=1}^{n} \varphi_j^2}. \tag{27}$$

Zazwyczaj, jeśli $\varepsilon = 10^{-2s}$, to $\sigma_i \leqslant 10^{-s}$. Zwykle wystarcza przyjąć $s = 3$. Należy zaznaczyć, że wymienione wskaźniki błędów dotyczą tylko rozwiązania macierzowego zagadnienia własnego, a nie rozkładu pola w falowodzie.

Kontrola zbieżności wyników do poszukiwanych par własnych jest przeprowadzana na podstawie szerszej analizy, dotyczącej właściwości ciągów Sturma [4, 17]. Wynika z niej, że po rozkładzie macierzy $[K]-\mu[M]$ na czynniki $[L][D][L]^T$ liczba ujemnych elementów w $[D]$ jest równa liczbie wartości własnych mniejszych od μ. Zatem, przy przyjęciu $\mu = 1{,}01\lambda_p^{(l)}$ (wartość sprawdzona praktycznie [2, 4]) powinno być p ujemnych elementów w macierzy $[D]$. Jeśli ich liczba jest większa, to należy zmienić wektory startowe lub zwiększyć ich liczbę i powtórzyć obliczenia.

Efektywność obliczeń metodą iteracji podprzestrzennych zależy w istotny sposób od obranego układu wektorów początkowych $[\Phi_0]$. Niewłaściwe ich przyjęcie może spowodować powolną zbieżność obliczeń, zbieżność kilku wektorów startowych do tego samego wektora własnego lub opuszczenie niektórych wartości własnych z interesującego nas fragmentu widma. Zbieżność do różnych par własnych można zapewnić przez wybór ortogonalnego układu początkowego. Przy poszukiwaniu sposobu zapewnienia zbieżności do pożądanych par własnych pomocna jest analogia do interpretacji mechanicznej macierzy $[K]$ jako macierzy „sztywności", zaś macierzy $[M]$ jako macierzy „masy" [4, 14, 18]. Otóż, analizując drgania własne konstrukcji podpartej (zamocowanej), należy spodziewać się maksymalnych amplitud w tych miejscach, gdzie skupiona jest stosunkowo duża masa przy jednocześnie małej sztywności. Wskaźnikami tych punktów są maksymalne wartości ilorazów m_{ii}/k_{ii} elementów z głównych przekątnych macierzy $[K]$ i $[M]$. Wektory startowe powinny zatem „pobudzać" wskazane stopnie swobody. Jest to reguła zaproponowana w pracy [4]. Jest to możliwie optymalny sposób wyboru wektorów $[\Phi_0]$ w przypadku fal typu E (warunki brzegowe Dirichleta odpowiadają zamocowaniu konstrukcji). Natomiast podczas obliczeń testowych zauważono, że przy warunkach brzegowych Neumanna (dla pól typu H) można podać inną regułę startową, prowadzącą do zmniejszenia liczby iteracji. Warunki brzegowe Neumanna odpowiadają analizie konstrukcji drgającej swobodnie. Maksymalne amplitudy drgań występują wówczas na brzegach nie zamocowanych, gdzie występują najmniejsze wartości ilorazu m_{ii}/k_{ii}.

Podsumowując, automatycznie wybierany w programie układ wektorów początkowych jest następujący:

— dla fal typu E wektor $\{\varphi_1^{(0)}\}$ jest równy głównej przekątnej macierzy $[M]$, zaś $\{\varphi_2^{(0)}\}$, ..., $\{\varphi_q^{(0)}\}$ są wektorami jednostkowymi, z wartościami równymi 1 dla tych współrzędnych, które odpowiadają największym ilorazom m_{ii}/k_{ii} (w kolejności malejącej);

— dla fal typu H wektor $\{\varphi_1^{(0)}\}$ jest bez zmian, zaś $\{\varphi_2^{(0)}\}, \ldots, \{\varphi_q^{(0)}\}$ stanowią wektory jednostkowe, z wartościami współrzędnych równymi 1 w miejscach odpowiadających najmniejszym ilorazom m_{ii}/k_{ii} (w kolejności rosnącej).

Autorzy pracy [4] sugerują, iż optymalną liczbą tak obranych wektorów startowych jest

$$q = \min(2p, p+8). \tag{28}$$

Zapewne jest to słuczne dla analizowanych przez nich wielkich macierzy, natomiast dla niezbyt dużych problemów własnych (jak w niniejszej pracy) najkrótsze czasy obliczeń uzyskiwano przy wyborze q raczej w pobliżu środka przedziału (28) — przy $p < 6$. Przy opisanym wyborze wektorów początkowych liczba iteracji podprzestrzennych zazwyczaj nie przekracza dziesięciu (przeważnie 5—7 cykli).

W przypadku jednorodnych warunków brzegowych Neumanna (pola typu H) macierz $[K]$ jest dodatnio półokreślona ($\lambda_1 = 0$). W celu zapewnienia dodatniej określoności tej macierzy należy zastosować przesunięcie widma problemu własnego (17) w kierunku wartości dodatnich:

$$([K]+\mu[M])\{\Phi\} = (\lambda+\mu)[M]\{\Phi\}. \tag{29}$$

Wartości własne tak zmodyfikowanego problemu są powiększone o μ, zaś wektory własne są nie zmienione. Optymalna wartość μ znajduje się w pobliżu najmniejszej dodatniej wartości własnej danego zagadnienia [4, 14]. Przy obliczaniu kilku pierwszych rodzajów fal typów E i H jako pierwsze wygodnie jest wyznaczyć rodzaje E, gdyż nie wymaga się tu przesunięcia widma (29). Następnie obliczane są rodzaje H, przy czym proponuje się przyjęcie $\mu = \lambda_{E1}/2$, gdzie λ_{E1} jest najmniejszą wartością własną dla fal typu E. Przesunięcia widma dla fal typu H nie trzeba wykonywać wówczas, gdy wykorzystuje się istniejące linie symetrii w celu nadania zerowych wartości u_i w określonych węzłach.

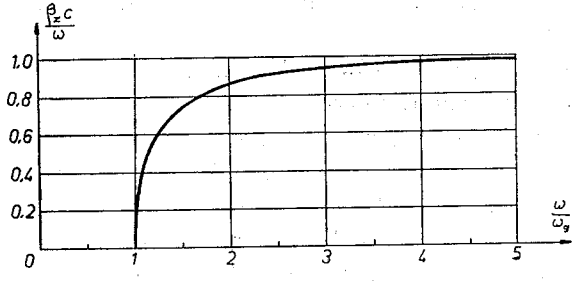
4. OBLICZANIE ROZKŁADÓW PÓL W FALOWODACH

Kompletną informację o rozkładzie pola (2) zawiera rozwiązanie równania (3) oraz równanie dyspersyjne (4), które po przekształceniu przybiera postać

$$\frac{\beta_z c}{\omega} = \sqrt{1-\left(\frac{\omega_g}{\omega}\right)^2}, \tag{30}$$

gdzie:

$$c = \frac{1}{\sqrt{\varepsilon\mu}}, \quad \omega_g = \lambda c^2.$$

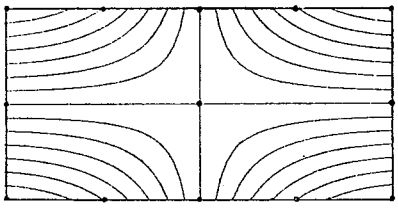


Rys. 2. Krzywa dyspersji falowodu jednorodnego

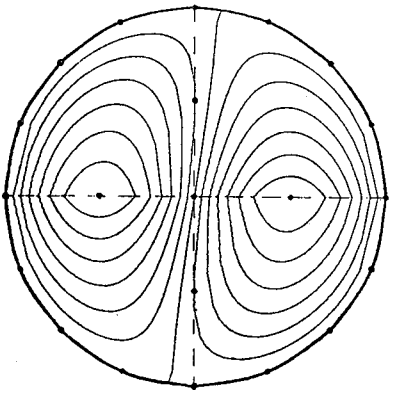
Równanie to można przedstawić graficznie za pomocą krzywej dyspersji, która w postaci znormalizowanej (rys. 2) jest słuszna dla dowolnego falowodu jednorodnego.

Obliczenia numeryczne przeprowadzono na komputerze ODRA 1305. Podczas uruchamiania programów dokonywano obliczeń kontrolnych dla przykładów możliwych do rozwiązania metodą rozdzielenia zmiennych (falowód prostokątny o stosunku boków 2:1 oraz falowód kołowy). Wskaźnikiem dokładności uzyskanych wyników może być iloraz $\varrho = \frac{\omega_p}{\omega_g}$, gdzie ω_p jest przybliżeniem pulsacji granicznej ω_g, uzyskanym metodą elementów skończonych. Przy podziale przekroju falowodu prostokątnego zaledwie na dwa elementy uzyskano przykładowo dla rodzaju H_{10} $\varrho = 1{,}00375$, zaś dla H_{11} — $\varrho = 1{,}0836$. Przy siatce sześcioelementowej uzyskano $\varrho = 1{,}00082$ dla rodzaju H_{10} oraz

$\varrho = 1{,}00353$ dla H_{11}. Obliczenia pulsacji granicznych falowodu kołowego przy podziale przekroju na cztery elementy są wykonywane z następującymi wskaźnikami błędów: $\varrho = 1{,}0042$ dla rodzaju podstawowego H_{11}, $\varrho = 1{,}0113$ dla E_{01}, $\varrho = 1{,}0187$ dla E_{11}, $\varrho = 1{,}0908$ dla H_{01}, $\varrho = 1{,}179$ dla E_{21}. Dokonywanie obszerniejszych zestawień świadczących o zbieżności wyników nie jest w niniejszej pracy celowe. Na rys. 3 i 4 przedstawiono przykładowe rozkłady $u(x, y)$ dla przykładów testowych. Na rysunkach tych zaczernione punkty oznaczają położenia węzłów siatek dyskretyzacji, a linie przerywane na rys. 4 — granice elementów skończonych. Należy podkreślić bardzo dobre przedstawienie krzywizny brzegu przez obrane elementy.



Rys. 3. Rozkład $u(x, y)$ dla rodzaju H_{11} w falowodzie prostokątnym przy podziale przekroju na dwa elementy skończone



Rys. 4. Rozkład $u(x, y)$ dla rodzaju E_{11} w falowodzie kołowym, przy podziale przekroju na cztery elementy skończone

Spotykane oznaczenia rodzajów fal ograniczają się zwykle do najczęściej stosowanych typów falowodów: prostokątnego i kołowego. W celu prostego przedstawienia wyników podanych na rysunkach 5—8 przyjęto następujące, jednoindeksowe oznaczenia:

H_n — dla fal typu H;

E_n — dla fal typu E.

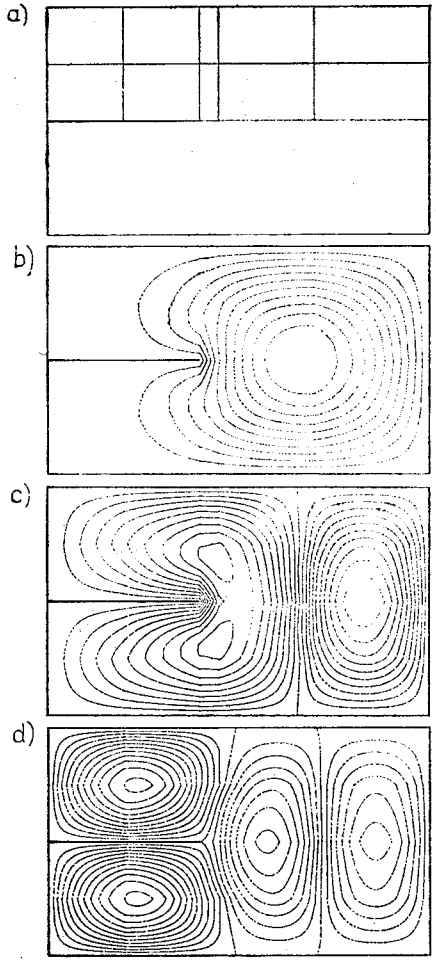
Indeks n oznacza kolejny numer wartości własnej danego typu, przy czym numer 1 odpowiada najmniejszej wartości własnej. Podstawowy rodzaj fali ma wszędzie nazwę H_1.

PODSUMOWANIE

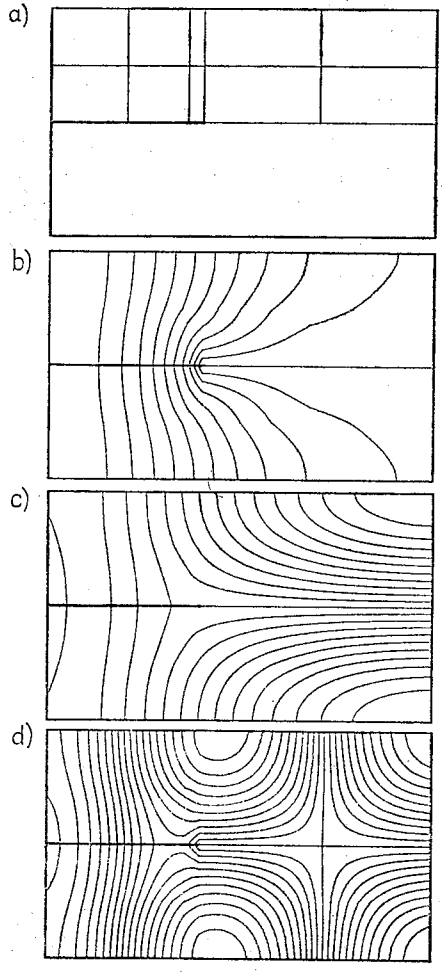
W analizie falowodów jednorodnych można równoważnie posługiwać się potencjałami wektorowymi Hertza lub składowymi wzdłużnymi natężeń pól E_z i H_z, traktując je jako rozwiązania równania (1) przy odpowiednich warunkach brzegowych.

Przy obliczaniu pól w falowodach o różnych kształtach przydatny jest zaproponowany wyżej sposób oznaczania rodzajów fal.

Zastosowane izoparametryczne elementy skończone drugiego rzędu zapewniają dużą



Rys. 5. Rozkłady rodzajów E w falowodzie prostokątnym z przegrodą: (a) — siatka dyskretyzacji, (b) — E_1, (c) — E_2, (d) — E_3

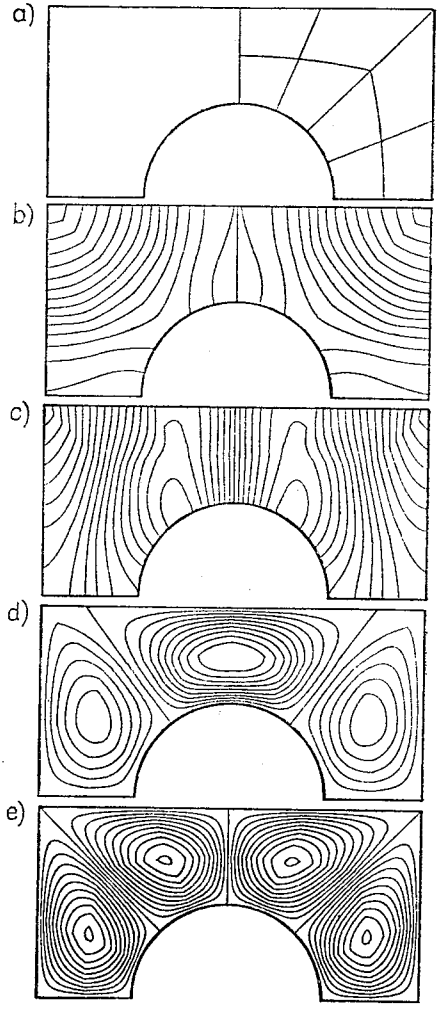


Rys. 6. Rozkłady rodzajów H w falowodzie prostokątnym z przegrodą: (a) — siatka dyskretyzacji, (b) — H_1, (c) — H_2, (d) — H_3

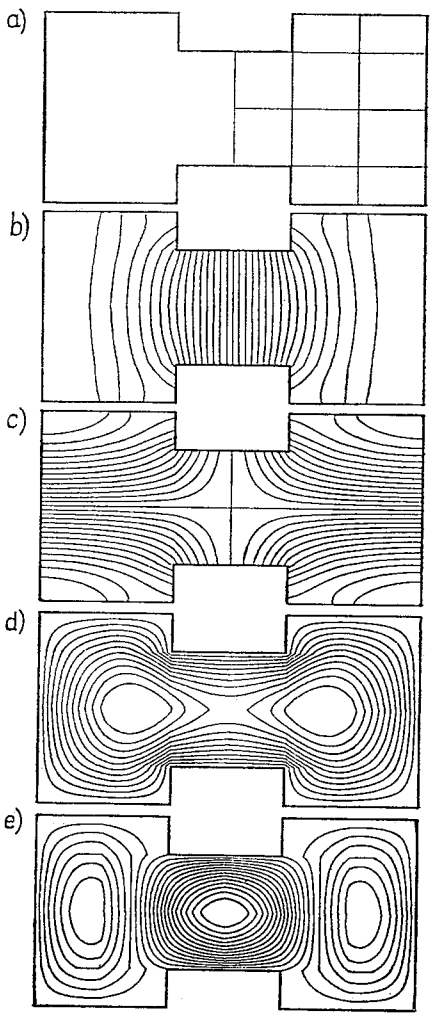
dokładność i szybką zbieżność obliczeń przy zagęszczaniu siatki podziału, nie stwarzając jednocześnie dużych trudności przy opracowaniu graficznym. Stosowanie jednolitego rodzaju elementów od początku analizy (generacja siatki) do jej końca (opracowanie graficzne) znacznie ułatwia pracę użytkownika programów numerycznych.

Metoda iteracji podprzestrzennych jest aktualnie optymalna do obliczeń pól w falowodach jednorodnych. W tablicy 1 przedstawiono porównanie jej z często stosowanymi metodami transformacyjnymi [4, 6, 12, 16, 17]. W tablicy tej n oznacza rząd macierzy $[K]$ i $[M]$, zaś m — szerokość półpasma elementów niezerowych.

Opracowana reguła automatycznego wyboru wektorów startowych zapewnia dużą efektywność i szybką zbieżność obliczeń iteracyjnych do poszukiwanych par własnych. Przy obliczaniu rozkładów pól typu H z wykorzystaniem przesunięcia widma (29),



Rys. 7. Rozkłady pól w falowodzie prostokątnym z grzbietem półkolistym: (a) — siatka dyskretyzacji, (b) — H_5, (c) — H_4, (d) — E_3, (e) — E_4



Rys. 8. Rozkłady pól w falowodzie dwugrzbietowym: (a) — siatka dyskretyzacji, (b) — H_1, (c) — H_4, (d) — E_2, (e) — E_3

zbieżność obliczeń jest wolniejsza, niż dla pól typu E. Należy więc, gdy jest to możliwe, wykorzystywać istniejące linie symetrii w celu wprowadzenia warunków brzegowych Dirichelta również dla pól typu H. Równocześnie może to prowadzić do zmniejszenia wymiarów macierzy $[\boldsymbol{K}]$ i $[\boldsymbol{M}]$, a więc w efekcie do wyraźnego skrócenia czasu pracy komputera.

Wyznaczone wartości własne są przybliżeniami z nadmiarem rzeczywistych wartości własnych, co wynika z właściwości zastosowanej analizy Rayleigha-Ritza [4, 10].

Zbudowane programy mogą mieć zastosowanie przy różnych pracach wspomagających analizę i projektowanie falowodów o specjalnych właściwościach i rzadziej spotykanych kształtach.

Tablica 1

Porównanie metod rozwiązywania zagadnienia własnego
$[K]\{\Phi\} = \lambda[M]\{\Phi\}$

Metoda obliczeń	Oszacowanie maksymalnej liczby operacji arytmetycznych	Zarezerwowana pamięć komputera
metoda Jacobiego (n wartości i wektorów własnych)	$18n^3+0(n^2)$	$2n^2$
metoda Householdera-QR (n wartości i p wektorów własnych)	$2n^3+pn^2+0(n^2)$	n^2+6n
metoda iteracji podprzestrzennych (przy założeniu, że wykonano 10 kroków iteracyjnych)	$nm^2+nm(4+2p)+4np+$ $+20nq\left(2m+q+\frac{3}{2}\right)+0(q^3)$	$2nm+2n(q+1)$

BIBLIOGRAFIA

1. S. Ahmed; *Finite-element method for waveguide problems*, Electronics Letters, 1968, Vol. 4, No. 18, pp. 387—389
2. K. Aniserowicz: *Analiza pola elektromagnetycznego w falowodach metodą elementów skończonych wyższego rzędu.* Rozprawa doktorska, Szczecin 1987
3. P. H. Arlett, A. K. Bahrani, O. C. Zienkiewicz; *Application of finite elements to the solution of Helmholtz equation.* Proc. IEE, 1968 No. 12
4. K. J. Bathe, E. L. Wilson: *Numerical methods in finite element analysis.* Prentice-Hall, 1976
5. R. E. Collin: *Prowadzenie fal elektromagnetycznych.* Warszawa: WNT, 1966
6. M. Dryja, J. i M. Jankowscy: *Przegląd metod i algorytmów numerycznych, Część 2.* Warszawa: WNT, 1982
7. E. Hinton, D. R. J. Owen: *Finite element programming.* London: Academic Press, 1977
8. R. Litwin: *Teoria pola elektromagnetycznego.* Warszawa: WNT, 1973
9. R. Litwin, M. Suski: *Technika mikrofalowa.* Warszawa: WNT, 1972
10. S. G. Michlin, C. L. Smolicki: *Metody przybliżone rozwiązywania równań różniczkowych i całkowych.* Warszawa: PWN, 1970
11. T. Morawski, W. Gwarek: *Teoria pola elektromagnetycznego.* Warszawa: WNT, 1985
12. A. Ralston: *Wstęp do analizy numerycznej*, Warszawa: PWN, 1983
13. K. R. Richter i in.: *Ein Beitrag zur Berechnung elektromagnetischer Felder mittels numerischer Methoden*, AEÜ, 1983, Band 37, Heft 5/6, ss. 174—182
14. H. R. Schwarz: *Methode der finiten Element.* Stuttgart: Teubner, 1980
15. P. Silvester: *A general high-order finite-element waveguide analysis program.* IEEE Trans. MTT-17, 1969, No. 4, pp. 204—210
16. J. Szmelter: *Metody komputerowe w mechanice.* Warszawa: PWN, 1980
17. J. H. Wilkinson: *The algebraic eigenvalue problem.* Clarendon Press, Oxford 1965
18. O. C. Zienkiewicz: *Metoda elementów skończonych.* Warszawa: Arkady, 1972

K. ANISEROWICZ

APPLICATION OF FINITE ELEMENT METHOD TO THE ANALYSIS OF ELECTROMAGNETIC FIELDS IN LOSSLESS HOMOGENEOUS WAVEGUIDES

Summary

A finite element analysis of electromagnetic fields in lossless homogeneous waveguides is described. Eight-node isoparametric elements of the second order are used. The matrix eigenproblem is solved by the subspace iteration method. Some examples of field plots are presented.

K. ANISEROWICZ

APPLICATION DE LA METHODE DES ELEMENTS FINIS À L'ANALYSE DES CHAMPS ÉLECTROMAGNÉTIQUES DANS LES GUIDES D'ONDES HOMOGÈNES SANS DISSIPATION D'ÉNERGIE

Résumé

On a présenté l'analyse des champs éelectromagnétiques dans les guides d'ondes homogènes sans dissipation d'énergie, en se servent de la méthode d'éléments finis. Dans les calculs on s'est servi d'éléments isoparamétriques à huit noeuds du second ordre. Le propre problème matriciel, en déduit, a été résolu par méthode d'itérations sous espace. On a inclu, à titre d'exemple, des dessins de la répartition des champs.

K. ANISEROWICZ

ANWENDUNG DER METHODE FINITER ELEMENTE FÜR DIE ANALYSE VON ELEKTROMAGNETISCHEN FELDERN IN HOMOGENEN VERLUSTLOSEN HOHLLEITERN

Zusammenfassung

Im Beitrag wurde eine Analyse der elektromagnetischen Felder in homogenen verlustlosen Hohlleitern unter Anwendung der Methode finiter Elemente dargestellt. Für die Berechnungen wurden 8-knotige isoparametrische Elemente 2. Ordnung angewandt. Die aufgetretenen Matrixeigenwertprobleme wurden mittels simultaner Vektoriteration gelöst. Als Beispiele wurden Zeichnungen der Feldverteilung beigefügt.

К. АНИСЕРОВИЧ

ПРИМЕНЕНИЕ МЕТОДА КОНЕЧНЫХ ЭЛЕМЕНТОВ К АНАЛИЗУ ЭЛЕКТРОМАГНИТНЫХ ПОЛЕЙ В ГОМОГЕННЫХ ВОЛНОВОДАХ БЕЗ ПОТЕРЬ

Резюме

Представлен анализ электромагнитных полей в гомогенных волноводах без потерь при применении метода конечных элементов. Расчеты проведены при использовании восьмимодовых изопараметрических элементов второго порядка. Матричная проблема собственных значений решена методом подпространных итераций. Приведены примеры диаграмм распределений полей.

ROZPRAWY ELEKTROTECHNICZNE, 1988, 34, z. 4, ss. 1065–1076

621.372.8:621.315.61

Zastosowanie metody elementów skończonych do analizy pól elektromagnetycznych w bezstratnych falowodach z niejednorodnym wypełnieniem dielektrycznym

KAROL ANISEROWICZ

Wydział Elektryczny, Politechnika Białostocka

Otrzymano 1987.09.17
Autoryzowano do druku 1988.01.05

Przedstawiono analizę pól elektromagnetycznych w bezstratnych falowodach z niejednorodnym wypełnieniem dielektrycznym z użyciem metody elementów skończonych. Do obliczeń zastosowano ośmiowęzłowe elementy izoparametryczne drugiego rzędu. Powstałe macierzowe zagadnienie własne rozwiązano metodą iteracji podprzestrzennych i wektorowej iteracji odwrotnej. Zamieszczono przykładowe rysunki rozkładów pól.

1. WSTĘP

Uzyskiwanie specjalnych właściwości falowodów jest możliwe przez zastosowanie złożonych kształtów przekroju poprzecznego lub przez wypełnienie niejednorodnym dielektrykiem. Falowodowe rodzaje pól mogą występować również w liniach przesyłowych przeznaczonych dla fal TEM i są wówczas niepożądane. Często założenie rozchodzenia się fal TEM może być dużym uproszczeniem zjawiska, jak np. w niejednorodnych liniach paskowych. Analityczne metody obliczania pól elektromagnetycznych w przypadku niejednorodnej prowadnicy falowej są skuteczne jedynie w najprostszych przypadkach [5]. Zastosowanie metod numerycznych, a szczególnie metody elementów skończonych (MES) może być w bardziej skomplikowanych zagadnieniach jedynym skutecznym sposobem teoretycznego wsparcia badań eksperymentalnych.

Niniejszy artykuł jest kontynuacją pracy [3], rozrzerzając analizę na falowody o niejednorodnościach poprzecznych. Mimo opisania teorii zastosowań MES do analizy takich falowodów w kilku pracach, np. [1, 6], nie jest jeszcze rozstrzygnięty problem znalezienia optymalnej metody rozwiązywania powstałego w trakcie obliczeń macierzowego zagadnienia własnego oraz wyboru właściwego rodzaju elementów skończonych. Zastosowano ośmiowęzłowe izoparametryczne elementy skończone drugiego rzędu, a rozwiązania macierzowego zagadnienia własnego dokonano metodą iteracji podprzestrzennych i wektorowej iteracji odwrotnej [4].

2. MODEL MATEMATYCZNY ZAGADNIENIA BRZEGOWEGO

Syntetyczne ujęcie ogólnych właściwości fal prowadzonych w falowodach jednorodnych jest możliwe za pomocą potencjałów Hertza [5, 10, 13]. W przypadku niejednorodności wypełnienia dielektrycznego posługiwanie się nimi napotyka na trudności. Związane jest to z faktem, że pojawiają się wówczas fale typu EH, a potencjały wektorowe Hertza mogą mieć składowe nie tylko w kierunku osi 0z. Co prawda, np. w pracy [5] używa się potencjałów Hertza dla nieskomplikowanych niejednorodności w falowodzie prostokątnym, wyróżniając rodzaje elektryczne (PPE) i magnetyczne (PPH) w podłużnej płaszczyźnie przekroju, ale podejście to nie ma znamion ogólności i nie daje się zastosować dla dowolnego kształtu falowodu. Dogodniej jest posługiwać się składowymi wzdłużnymi natężeń pól elektromagnetycznych, które poszukujemy w postaci

$$E_z(x, y, z) = E_{zt}(x, y)e^{-j\beta_z z} \quad \text{oraz} \quad H_z(x, y, z) = H_{zt}e^{-j\beta_z z}. \tag{1}$$

Korzystając z tych zależności oraz z równań Maxwella w ośrodku bezźródłowym można wyprowadzić następujące wzory we współrzędnych kartezjańskich [6]:

$$\frac{\partial E_z}{\partial y} = -j\beta_z E_y - j\omega\mu H_x \tag{2}$$

$$\frac{\partial E_z}{\partial x} = -j\beta_z E_x + j\omega\mu H_y \tag{3}$$

$$j\omega\mu H_z = \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} \tag{4}$$

$$\frac{\partial H_z}{\partial y} = -j\beta_z H_y + j\omega\varepsilon E_x \tag{5}$$

$$\frac{\partial H_z}{\partial x} = -j\beta_z H_x - j\omega\varepsilon E_y \tag{6}$$

$$j\omega\varepsilon E_z = \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}. \tag{7}$$

Jeśli znane są składowe E_z i H_z, to można stąd wyznaczyć pozostałe składowe pól.

W celu uproszczenia zapisu przyjęto dalej, że symbole E_z i H_z będą oznaczać części poprzeczne E_{zt} i H_{zt} odpowiednich składowych pól. Spełniają one układ równań Helmholtza, które wygodnie jest zapisać w postaci

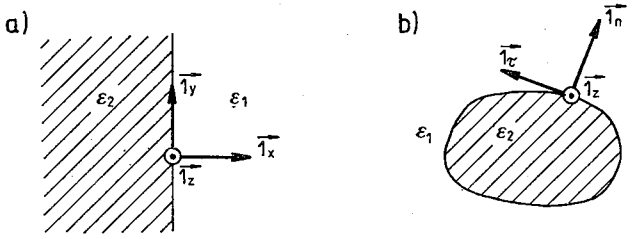
$$-\varepsilon\left(\frac{1}{k^2}\nabla_t^2 + 1\right)E_z = 0 \tag{8}$$

$$-\mu\left(\frac{1}{k^2}\nabla_t^2 + 1\right)H_z = 0, \tag{9}$$

gdzie:

$$\nabla_t^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}, \quad k^2 = \omega^2(\varepsilon\mu - \delta^2), \quad \delta = \frac{\beta_z}{\omega}.$$

Na ściankach falowodu E_z spełnia jednorodny warunek brzegowy Dirichleta $E_z = 0$, zaś H_z — warunek brzegowy Neumanna $\frac{\partial H_z}{\partial n} = 0$. Do dalszych obliczeń konieczne jest określenie związku między składowymi pól na granicy dwóch ośrodków. Można łatwo uogólnić zależności między składowymi w globalnym układzie kartezjańskim a lokalnym układem kierunków ortogonalnych (rys. 1).



Rys. 1. Wersory swobodne globalnego układu współrzędnych $0xyz$ (a) oraz lokalny układ kierunków ortogonalnych na powierzchni walcowej o dowolnym przekroju poprzecznym (b)

Niech τ oraz z będą ortogonalnymi kierunkami, stycznymi do powierzchni granicznej dwóch ośrodków, zaś n — kierunkiem prostopadłym do niej. Wykorzystując analogie $x-n$ oraz $y-\tau$ (z rysunku 1) oraz równości (3) i (5), można napisać [7]:

$$\frac{\partial E_z}{\partial n} = -\frac{\beta_z}{\omega\varepsilon}\frac{\partial H_z}{\partial \tau} + \frac{jk^2}{\omega\varepsilon} H_\tau. \tag{10}$$

Podobnie, uwzględniając (6) i (2), otrzymuje się

$$\frac{\partial H_z}{\partial n} = \frac{\beta_z}{\omega\mu}\frac{\partial E_z}{\partial \tau} - \frac{jk^2}{\omega\mu} E_\tau. \tag{11}$$

Funkcjonał energetyczny F, odpowiadający układowi równań (8), (9) jest sumą dwóch składników: elektrycznego F_E i magnetycznego F_H. Niech Ω oznacza rozważany obszar dwuwymiarowy (przekrój falowodu), zaś τ — jego brzeg. Wówczas

$$F_E = -\frac{1}{2}\iint_\Omega \varepsilon E_z\left(\frac{1}{k^2}\nabla_t^2 E_z + E_z\right)\mathrm{d}x\,\mathrm{d}y. \tag{12}$$

Wykorzystując tożsamość Greena

$$\iint_\Omega u\nabla^2 u\,\mathrm{d}x\,\mathrm{d}y = -\iint_\Omega (\nabla u)^2\mathrm{d}x\,\mathrm{d}y + \oint_\tau u\frac{\partial u}{\partial n}\,\mathrm{d}\tau \tag{13}$$

oraz równość (10) otrzymuje się

$$F_E = \frac{1}{2}\iint_\Omega \frac{\varepsilon}{k^2}(\nabla_t E_z)^2\,\mathrm{d}x\,\mathrm{d}y - \frac{1}{2}\iint_\Omega \varepsilon E_z^2\,\mathrm{d}x\,\mathrm{d}y + \frac{\beta_z}{2\omega}\oint_\tau \frac{1}{k^2}E_z\frac{\partial H_z}{\partial \tau}\,\mathrm{d}\tau - \frac{j}{2\omega}\oint_\tau E_z H_\tau\,\mathrm{d}\tau \tag{14}$$

Podobnie

$$F_H = \frac{1}{2}\iint_{\Omega} \frac{\mu}{k^2}(\nabla_t H_z)^2 \mathrm{d}x\,\mathrm{d}y - \frac{1}{2}\iint_{\Omega} \mu H_z^2 \mathrm{d}x\,\mathrm{d}y -$$

$$- \frac{\beta_z}{2\omega}\oint_{\tau} \frac{1}{k^2} H_z \frac{\partial E_z}{\partial \tau}\mathrm{d}\tau + \frac{j}{2\omega}\oint_{\tau} H_z E_\tau \mathrm{d}\tau \tag{15}$$

Ostatnie całki (z jednością urojoną) we wzorach (14), (15) dają w sumie zero w obszarze globalnym, gdyż wielkości podcałkowe E_z i E_τ znikają na ściankach przewodzących. Ponadto H_z i H_τ zerują się na liniach symetrii, co jest istotne przy obliczaniu pola we fragmencie obszaru Ω, z wykorzystaniem symetrii zagadnienia brzegowego. Funkcjonał dla zespołu wielu jednorodnych podobszarów Ω^e (elementów skończonych) jest sumą wkładów każdego z nich. Łatwo można zauważyć, że całki te są liczone wzdłuż każdego brzegu podobszaru dwukrotnie — w przeciwnych kierunkach, więc po zsumowaniu wzajemnie się zniosą.

Przedostatnie całki konturowe we wzorach (14), (15) po zsumowaniu dają zero w obszarach jednorodnych lub znikają przy częstotliwości granicznej danego rodzaju fal (gdy $\beta_z = 0$). Ich wpływ w obszarze globalnym jest niezerowy tylko na granicy dwóch ośrodków, przy $\beta_z > 0$. Całki te uwzględniają w modelu matematycznym związek między składowymi E_z i H_z. W celu uproszczenia algorytmu obliczeniowego przy całkowaniu numerycznym można je wyrazić za pomocą odpowiadających im całek powierzchniowych. Wykorzystując twierdzenie Stokesa

$$\oint_{\tau} \vec{P}\cdot \mathrm{d}\vec{\tau} = \iint_{\Omega} \nabla\times\vec{P}\cdot \mathrm{d}\vec{\Omega} \tag{16}$$

oraz tożsamość wektorową

$$\nabla\times(\varphi\vec{R}) = \varphi\nabla\times\vec{R} + (\nabla\varphi)\times\vec{R}, \tag{17}$$

otrzymuje się

$$\oint_{\tau} H_z \frac{\partial E_z}{\partial \tau}\mathrm{d}\tau = -\oint_{\tau} E_z \frac{\partial H_z}{\partial \tau}\mathrm{d}\tau = \iint_{\Omega} (\nabla H_z \times \nabla E_z)\cdot \vec{1}_z \mathrm{d}x\,\mathrm{d}y. \tag{18}$$

Zatem, we współrzędnych kartezjańskich całkowity funkcjonał przyjmuje postać

$$F = \frac{1}{2}\iint_{\Omega} \frac{\varepsilon}{k^2}\left[\left(\frac{\partial E_z}{\partial x}\right)^2 + \left(\frac{\partial E_z}{\partial y}\right)^2\right]\mathrm{d}x\,\mathrm{d}y + \frac{1}{2}\iint_{\Omega} \frac{\mu}{k^2}\left[\left(\frac{\partial H_z}{\partial x}\right)^2 + \left(\frac{\partial H_z}{\partial y}\right)^2\right]\mathrm{d}x\,\mathrm{d}y -$$

$$- \frac{1}{2}\iint_{\Omega} \varepsilon E_z^2 \mathrm{d}x\,\mathrm{d}y - \frac{1}{2}\iint_{\Omega} \mu H_z^2 \mathrm{d}x\,\mathrm{d}y + \frac{\beta_z}{\omega}\iint_{\Omega} \frac{1}{k^2}\left(\frac{\partial E_z}{\partial x}\frac{\partial H_z}{\partial y} - \frac{\partial E_z}{\partial y}\frac{\partial H_z}{\partial x}\right)\mathrm{d}x\,\mathrm{d}y. \tag{19}$$

Warunkiem koniecznym osiągnięcia ekstremum funkcjonału F jest

$$\frac{\partial F}{\partial E_z} = 0 \quad \text{i} \quad \frac{\partial F}{\partial H_z} = 0. \tag{20}$$

Obszar Ω przekroju poprzecznego falowodu dzielony jest na ośmiowęzłowe elementy

skończone drugiego rzędu, jak w pracy [3]. Każdemu węzłowi j ($j = 1, \ldots, 8$) elementu e przypisuje się dwie obliczane wielkości: e_j oraz h_j, będące odpowiednio wartościami amplitud natężenia pola elektrycznego E_z i magnetycznego H_z w punkcie j. Wewnątrz elementu skończonego rozkłady pól zakłada się w postaci

$$E_z^e(\xi, \eta) = \sum_j e_j N_j, \tag{21}$$

$$H_z^e(\xi, \eta) = \sum_j h_j N_j, \tag{22}$$

gdzie: ξ, η — współrzędne lokalne elementu e; N_j — funkcje kształtu.

Siatka dyskretyzacji obszaru Ω zawiera n węzłów. W przypadku obliczania rozkładów pól przy częstotliwościach granicznych węzłowi i ($i = 1, \ldots, n$) może być przyporządkowana tylko jedna z wielkości e_i lub h_i. Wartości węzłowe pól obliczane są z warunku (20). Oznaczając składnik funkcjonału F pochodzący od elementu e przez F^e po przekształceniach otrzymuje się

$$\frac{\partial F^e}{\partial \{\psi_i\}} = \frac{1}{\omega^2(\varepsilon\mu - \delta^2)} \sum_j \begin{bmatrix} \varepsilon s_{ij} & \delta d_{ij} \\ -\delta d_{ij} & \mu s_{ij} \end{bmatrix} \{\psi_j\} - \sum_j \begin{bmatrix} \varepsilon t_{ij} & 0 \\ 0 & \mu t_{ij} \end{bmatrix} \{\psi_j\}, \quad i = 1, 2, \ldots, n, \tag{23}$$

gdzie $\{\psi_i\} = \begin{Bmatrix} e_i \\ h_i \end{Bmatrix}$

oraz

$$s_{ij} = \iint_{\Omega^e} \left(\frac{\partial N_i}{\partial x} \frac{\partial N_j}{\partial x} + \frac{\partial N_i}{\partial y} \frac{\partial N_j}{\partial y} \right) \mathrm{d}x\,\mathrm{d}y, \tag{24}$$

$$t_{ij} = \iint_{\Omega^e} N_i N_j \,\mathrm{d}x\,\mathrm{d}y, \tag{25}$$

$$d_{ij} = \iint_{\Omega^e} \left(\frac{\partial N_i}{\partial x} \frac{\partial N_j}{\partial y} - \frac{\partial N_i}{\partial y} \frac{\partial N_j}{\partial x} \right) \mathrm{d}x\,\mathrm{d}y. \tag{26}$$

Transformację liniową, umożliwiającą przedstawienie całek (24)—(26) w znormalizowanych współrzędnych lokalnych, przedstawiają zależności [18]:

$$x = \sum_j N_j x_j, \quad y = \sum_j N_j y_j, \tag{27}$$

przy czym (x_j, y_j) są współrzędnymi globalnymi węzła j. We współrzędnych lokalnych całki (24)—(26) przyjmują postać

$$s_{ij} = \int_{-1}^{1} \int_{-1}^{1} \left[\left(\frac{\partial N_i}{\partial \xi} \frac{\partial y}{\partial \eta} - \frac{\partial N_i}{\partial \eta} \frac{\partial y}{\partial \xi} \right) \left(\frac{\partial N_j}{\partial \xi} \frac{\partial y}{\partial \eta} - \frac{\partial N_j}{\partial \eta} \frac{\partial y}{\partial \xi} \right) + \right.$$

$$\left. + \left(-\frac{\partial N_i}{\partial \xi} \frac{\partial x}{\partial \eta} + \frac{\partial N_i}{\partial \eta} \frac{\partial x}{\partial \xi} \right) \left(-\frac{\partial N_j}{\partial \xi} \frac{\partial x}{\partial \eta} + \frac{\partial N_j}{\partial \eta} \frac{\partial x}{\partial \xi} \right) \right] \frac{\mathrm{d}\xi\,\mathrm{d}\eta}{\det[J]}, \tag{28}$$

$$t_{ij} = \int_{-1}^{1}\int_{-1}^{1} N_i N_j \det[J]\,d\xi\,d\eta, \tag{29}$$

$$d_{ij} = \int_{-1}^{1}\int_{-1}^{1}\left[\left(\frac{\partial N_i}{\partial \xi}\frac{\partial y}{\partial \eta} - \frac{\partial N_i}{\partial \eta}\frac{\partial y}{\partial \xi}\right)\left(-\frac{\partial N_j}{\partial \xi}\frac{\partial x}{\partial \eta} + \frac{\partial N_j}{\partial \eta}\frac{\partial x}{\partial \xi}\right) - \left(-\frac{\partial N_i}{\partial \xi}\frac{\partial x}{\partial \eta} + \frac{\partial N_i}{\partial \eta}\frac{\partial x}{\partial \xi}\right)\left(\frac{\partial N_j}{\partial \xi}\frac{\partial y}{\partial \eta} - \frac{\partial N_j}{\partial \eta}\frac{\partial y}{\partial \xi}\right)\right]\frac{d\xi\,d\eta}{\det[J]}, \tag{30}$$

przy czym

$$[J] = \begin{bmatrix} \dfrac{\partial x}{\partial \xi} & \dfrac{\partial y}{\partial \xi} \\ \dfrac{\partial x}{\partial \eta} & \dfrac{\partial y}{\partial \eta} \end{bmatrix}. \tag{31}$$

Całki (28)—(30) oblicza się numerycznie, za pomocą kwadratur Gaussa—Legendre'a [4, 9, 16, 18].

Po zsumowaniu wpływu wszystkich elementów według zależności

$$\frac{\partial F}{\partial \{\psi_i\}} = \frac{\partial\left(\sum\limits_e F^e\right)}{\partial \{\psi_i\}} = \sum_e \frac{\partial F^e}{\partial \{\psi_i\}} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}, \quad i = 1, \ldots, n, \tag{32}$$

otrzymuje się układ równań

$$\frac{1}{\varepsilon\mu - \delta^2}\left[\begin{array}{c|c} \varepsilon[S] & \delta[D] \\ \hline -\delta[D] & \mu[S] \end{array}\right]\begin{Bmatrix} E_z \\ H_z \end{Bmatrix} = \omega^2\left[\begin{array}{c|c} \varepsilon[T] & [0] \\ \hline [0] & \mu[T] \end{array}\right]\begin{Bmatrix} E_z \\ H_z \end{Bmatrix}, \tag{33}$$

Ma on postać uogólnionego algebraicznego problemu wartości własnych

$$[K]\{\Phi\} = \lambda[M]\{\Phi\}, \tag{34}$$

gdzie rząd macierzy $[K]$ i $[M]$ wynosi $2n$, $\lambda = \omega^2$, zaś

$$\{\Phi\} = \begin{Bmatrix} e_1 \\ h_1 \\ e_2 \\ h_2 \\ \vdots \\ e_n \\ h_n \end{Bmatrix}. \tag{34a}$$

Po zbudowaniu układu równań uwzględnia się w nim warunki brzegowe przez znaną modyfikację macierzy $[K]$ i $[M]$ [9, 18].

3. OBLICZANIE ROZKŁADÓW PÓL W FALOWODACH NIEJEDNORODNYCH

Badanie rozkładów pól w falowodach o wypełnieniu niejednorodnym wymaga o wiele większej liczby obliczeń, niż w przypadku falowodów jednorodnych. Wynika to z dwóch przyczyn:

— dwukrotnie większych wymiarów macierzy $[K]$ i $[M]$, niż dla obszaru jednorodnego o tej samej liczbie węzłów siatki dyskretyzacji;
— zmienności rozkładu E_z i H_z w płaszczyźnie przekroju poprzecznego wraz ze zmianą ω oraz silnej zależności krzywej dyspersji od geometrii brzegów i stałych materiałowych ε, μ.

Zatem, aby otrzymać kompletną informację o zmienności rozkładu danego rodzaju pola w funkcji częstotliwości, należy wielkorotnie rozwiązać problem własny (33), zakładając różne wartości $\delta = \frac{1}{v_f}$ (gdzie v_f — prędkość fazowa propagacji fali elektromagnetycznej w falowodzie). Przy tym nie można sporządzić uniwersalnej krzywej dyspersji, jak w przypadku falowodów jednorodnych [3].

Załóżmy, że falowód wypełniony jest częściowo powietrzem, a częściowo materiałem o parametrach ε_1, μ_1. Z porównania zależności (33), (34) wynika, że macierz $[K]$ może być dodatnio określona tylko dla $0 \leqslant \delta < \sqrt{\varepsilon_0 \mu_0}$. W przypadku braku linii symetrii (gdzie $H_z = 0$), dodatnią określoność macierzy $[K]$ uzyskuje się stosując przesunięcie widma [3, 4].

Przypadek $\delta = \frac{1}{c} = \sqrt{\varepsilon_0 \mu_0}$ oznacza, że w równaniu Helmholtza $k = 0$, czyli rozkład pola w obszarze powietrznym jest opisany równaniem Laplace'a.

Dla $\frac{1}{c} < \delta < \sqrt{\varepsilon_1 \mu_1}$ z obszaru o parametrach ε_0, μ_0 otrzymuje się ujemne elementy macierzy $[K]$. Wiele wartości własnych jest ujemnych, przez co otrzymuje się uboczne, niefizyczne rozkłady pól. Pierwsza dodatnia wartość własna odpowiada rodzajowi podstawowemu. W praktyce należy się spodziewać, że liczba wartości ujemnych będzie stanowić około połowy wszystkich wartości własnych [6]. W zakresie $\delta \geqslant \sqrt{\varepsilon_1 \mu_1}$ fale elektromagnetyczne w falowodzie nie rozchodzą się.

Z powyższych rozważań wynika, że zakres efektywnego wykorzystania metody iteracji podprzestrzennych, wykorzystywnej w pracy [3] ograniczony jest do częstotliwości, przy których $0 \leqslant \delta < \frac{1}{c}$. Zastosowany algorytm obliczeń par własnych jest bowiem skuteczny tylko dla dodatnio określonych macierzy $[K]$ i $[M]$. Przy wyższych częstotliwościach należy stosować inną metodę obliczeń. Dla dużej liczby poszukiwanych par własnych $(\lambda_i, \{\Phi_i\})$ można posłużyć się metodą Householdera — QR [4, 17], natomiast pojedyncze pary własne najefektywniej oblicza się z wykorzystaniem wektorowej iteracji odwrotnej. W wersji podstawowej służy ona do obliczania najmniejszej wartości λ_1 i odpowiadającego jej wektora własnego $\{\Phi_1\}$.

Przy obliczeniach opisanych w niniejszej pracy stosowano następujący algorytm postępowania. Przyjmuje się pewien wektor startowy $\{\varphi_1\}$, po czym oblicza się poprawiony wektor

$$\{\psi_1\} = [M]\{\varphi_1\}. \tag{35}$$

Dla $k = 1, 2, \ldots$ oblicza się kolejno: $\{\bar{\varphi}_{k+1}\}$, $\{\bar{\psi}_{k+1}\}$, ϱ_{k+1}, $\{\psi_{k+1}\}$, korzystając z zależności

$$[K]\{\overline{\varphi}_{k+1}\} = \{\psi_k\}, \tag{36}$$

$$\{\overline{\varphi}_{k+1}\} = [M]\{\overline{\varphi}_{k+1}\}, \tag{37}$$

$$\varrho_{k+1} = \frac{\{\overline{\varphi}_{k+1}\}^T\{\psi_k\}}{\{\overline{\varphi}_{k+1}\}^T\{\overline{\psi}_{k+1}\}}, \tag{38}$$

$$\{\psi_{k+1}\} = \frac{\{\overline{\psi}_{k+1}\}}{\sqrt{\{\overline{\varphi}_{k+1}\}^T\{\psi_{k+1}\}}}. \tag{39}$$

Przy założeniu, że wektor $\{\varphi_1\}$ nie jest ortogonalny do $\{\Phi_1\}$, tzn. $\{\varphi_1\}^T[M]\{\Phi_1) \neq 0$, ciąg wektorów $\{\psi_{k+1}\}$ jest zbieżny do $[M]\{\Phi_1\}$, a ϱ_{k+1} — do λ_1 przy $k \to \infty$ [4]. Obliczenia iteracyjne (36)—(39) zostają przerwane, gdy kolejne przybliżenia poszukiwanej wartości własnej spełniają zależność

$$\left|\frac{\varrho_{k+1}-\varrho_k}{\varrho_{k+1}}\right| \leqslant \varepsilon, \tag{40}$$

gdzie ε jest zakładaną dokładnością obliczeń. Wektor $\{\Phi_1\}$ oblicza się z zależności

$$\{\Phi_1\} = \frac{\{\overline{\varphi}_{l+1}\}}{\sqrt{\{\overline{\varphi}_{l+1}\}^T\{\overline{\psi}_{l+1}\}}}, \tag{41}$$

gdzie l jest indeksem ostatniej iteracji. Zbieżność iteracji jest liniowa, ze wskaźnikiem λ_1/λ_2 (przy czym im mniejszy wskaźnik, tym lepsza zbieżność). Poprawienie wskaźnika zbieżności jest możliwe przez przesunięcie widma zagadnienia własnego

$$([K]-\mu[M])\{\Phi\} = (\lambda-\mu)[M]\{\Phi\}. \tag{42}$$

Wartościami własnymi tak zmodyfikowanego problemu są $\eta_i = \lambda_i - \mu$. Obliczenia są zbieżne do wartości własnej η_i najbliższej zeru i odpowiadającego jej wektora własnego $\{\Phi_i\}$ przy założeniu, że nie jest on ortogonalny do wektora startowego $\{\varphi_1\}$. Zatem przesunięcie widma może być również efektywnie wykorzystane do obliczenia par własnych, innych niż $(\lambda_1, \{\Phi_1\})$.

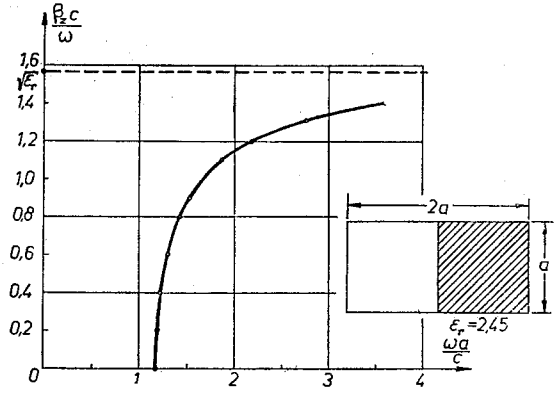
Metoda iteracji podprzestrzennych jest modyfikacją metody iteracji wektorowej odwrotnej, zatem rozważania na temat błędów wyznaczania wartości i wektorów własnych zawarte w pracy [3] są również słuszne w niniejszej pracy.

W celu zapewnienia szybkiej zbieżności obliczeń do poszukiwanej pary własnej $(\lambda_i, \{\Phi_i\})$ należy odpowiednio dobrać wartość μ oraz wektor startowy $\{\varphi_1\}$. Z obliczeń testowych wynika, że dobór μ zwykle nie jest krytyczny — po prostu μ powinno być bliskie poszukiwanej wartości własnej (im bliższe, tym szybsza zbieżność). Należy tu zaznaczyć, że μ nie może być równe (lub prawie równe) λ_i, w celu uniknięcia trudności numerycznych pojawiających się przy $\eta_i \approx 0$.

Wektor $\{\varphi_1\}$ wygodnie jest obierać jako wektor jednostkowy, o współrzędnej równej 1 w miejscu zapewniającym „pobudzenie" tego stopnia swobody, w którym przewiduje się maksymalną wartość amplitudy pola. Miejsce to można łatwo zidentyfikować po przeprowadzniu analizy metodą iteracji podprzestrzennych np. dla $\delta = 0$. Przy dobrym doborze μ liczba cykli iteracyjnych (36)—(39) jest równa 4... 5, (dla $\varepsilon = 10^{-6}$), zaś przy μ różniącym się o kilkadziesiąt procent od λ_i potrzeba kilkunastu iteracji.

4. PRZYKŁADY OBLICZENIOWE

Poprawność działania programów obliczeniowych sprawdzono na przykładach testowych, zawartych w pracach [1, 6]. Uzyskano zgodność wyników z danymi literaturowymi. Na rys. 2 przedstawiono krzywą dyspersji rodzaju podstawowego w testowym przykładzie falowodu prostokątnego.



Rys. 2. Krzywa dyspersji testowego przykładu falowodu prostokątnego

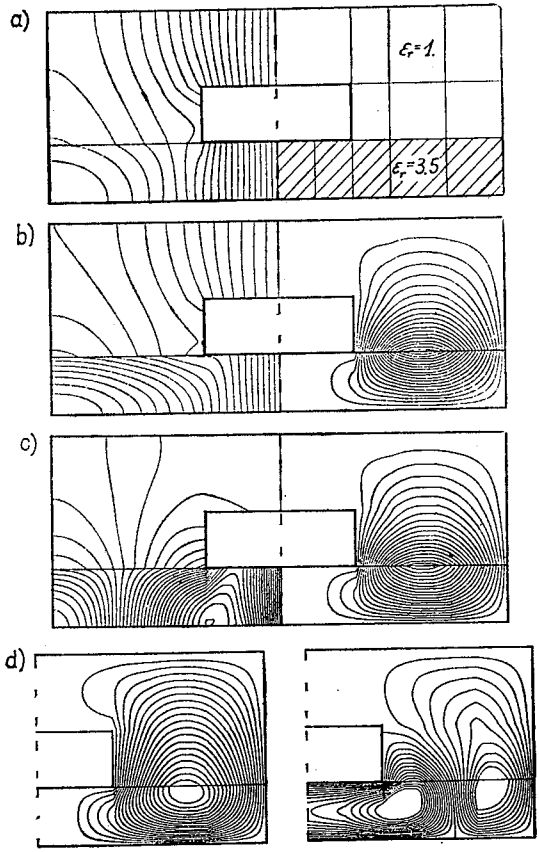
W celu prostego opisu wyników przedstawionych na rysunkach 3—4 przyjęto następujące, jednoindeksowe oznaczenia:
HE_n — dla fal nie posiadających składowej E_z przy $\omega = \omega_g$;
EH_n — dla fal nie posiadających składowej H_z przy $\omega = \omega_g$.
Indeks n oznacza kolejny numer wartości własnej danego typu, prz czym numer 1 odpowiada najmniejszej wartości własnej. Wyróżnikiem typu HE lub EH jest istnienie odpowiedniej składowej wzdłużnej pola przy częstotliwości granicznej. Rodzajem podstawowym w falowodach niejednorodnych jest HE_1.

Na rysunkach 3—4 przedstawiono wyniki obliczeń przykładów falowodów niejednorodnych, a na rys. 5 — krzywe dyspersji rodzajów podstawowych w tych falowodach.

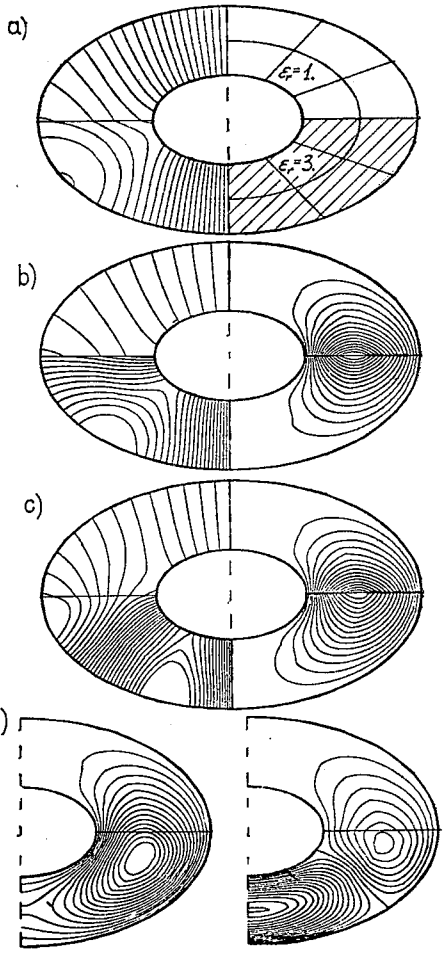
PODSUMOWANIE

W analizie falowodów niejednorodnych posługiwanie się składowymi wzdłużnymi natężeń pól E_z i H_z jest korzystniejsze, niż potencjałami wektorowymi Hertza.

Metoda iteracji podprzestrzennych może być efektywnie stosowana do analizy pól w pobliżu częstotliwości granicznych, dopóki macierz $[K]$ nie jest nieokreślona. Dla wyższych częstotliwości należy stosować inne metody (np. Householdera — QR [4]). Przy poszukiwaniu pojedynczych par własnych bardzo efektywna jest metoda wektorowej iteracji odwrotnej. Odpowiednie przesunięcie widma (42) zwiększa wyraźnie szybkość obliczeń komputerowych tą metodą. Dobór μ oraz wektora startowego $\{\varphi_1\}$ ma istotne znaczenie dla zbieżności obliczeń do określonej pary własnej. Obliczone wartości własne



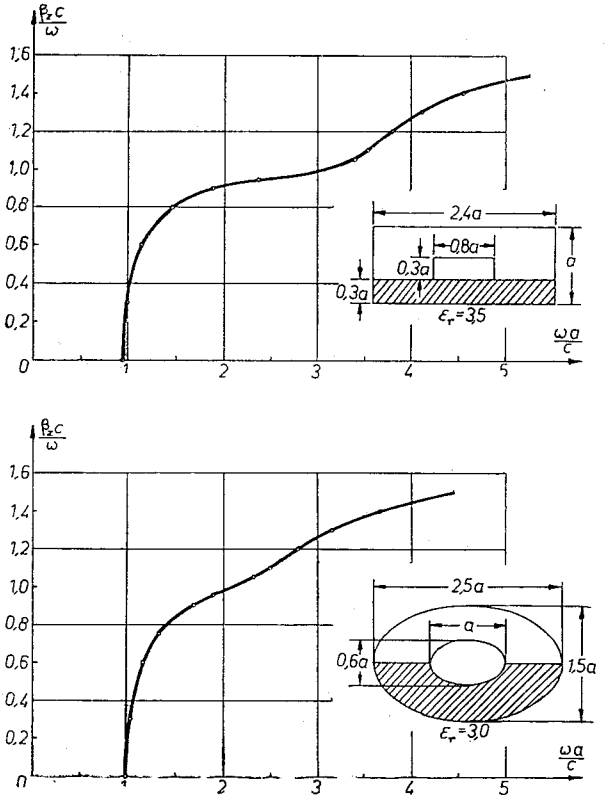
Rys. 3. Niejednorodna, ekranowana linia paskowa — z lewej strony rozkład składowej H_z, z prawej — E_z: (a) — HE_1 przy $\delta = 0$ oraz siatka dyskretyzacji, (b) — HE_1 przy $\delta c = 0{,}9$, (c) — HE_3 przy $\delta c = 0{,}9$, (d) — EH_1 przy $\delta = 0$ i EH_3 przy $\delta = 0$ (tylko składowe E_z, bo $H_z = 0$ przy $\delta = 0$)



Rys. 4. Niejednorodny, współosiowy falowód eliptyczny — z lewej strony rozkład składowej H_z, z prawej — E_z: (a) — HE_1 przy $\delta = 0$ oraz siatka dyskretyzacji, (b) — HE_1 przy $\delta c = 0{,}9$, (c) — HE_3 przy $\delta c = 0{,}9$, (d) — składowe E_z pól EH_1 i EH_3 przy $\delta = 0$

są przybliżeniami z nadmiarem rzeczywistych wartości własnych — podobnie jak w pracy [3]. Nakład obliczeń komputerowych przy analizie falowodów niejednorodnych jest znacznie większy, niż dla falowodów jednorodnych, toteż bardziej istotne staje się wykorzystanie istniejących linii symetrii w celu zmniejszenia liczby węzłów siatki dyskretyzacji.

Programy zbudowane na podstawie przedstawionej powyżej analizy mogą mieć zastosowanie przy analizie i projektowaniu niejednorodnych prowadnic falowych lub przy przewidywaniu możliwości wystąpienia niepożądanych, falowodowych rodzajów pól w liniach przesyłowych TEM.



Rys. 5. Krzywe dyspersji falowodów z rysunków 3 i 4

BIBLIOGRAFIA

1. S. Ahmed, P. Daly: *Finite-element methods for inhomogeneous waveguides*. Proc. IEE, 1969, Vol. 116, No. 10, pp. 1661—1664
2. K. Aniserowicz: *Analiza pola elektromagnetycznego w falowodach metodą elementów skończonych wyższego rzędu*, Rozprawa doktorska, Szczecin 1987
3. K. Aniserowicz: *Numeryczna analiza pól elektromagnetycznych w bezstratnych falowodach jednorodnych*, Rozprawy Elektrotechniczne, 1989 (w druku)
4. K. J. Bathe, E. L. Wilson: *Numerical methods in finite element analysis*. Prentice-Hall 1976
5. R. E. Collin: *Prowadzenie fal elektromagnetycznych*, Warszawa: WNT, 1966
6. Z. J. Csendes, P. Silvester: *Numerical solution of dielectric loaded waveguides: I — Finite-element analysis*, IEEE Trans. 1970, MTT-18, No. 12, pp. 1124—1131
7. P. Daly: *Finite-element coupling matrices. Electronics Letters*, 1969, Vol. 5, No. 24, pp. 613—615
8. M. Dryja, J. i M. Jankowscy: *Przegląd metod i algorytmów numerycznych*, Część 2, Warszawa: WNT, 1982
9. E. Hinton, D. R. J. Owen: *Finite element programming*. London: Academic Press, 1977
10. R. Litwin: *Teoria pola elektromagnetycznego*, Warszawa: WNT, 1973
11. R. Litwin, M. Suski: *Technika mikrofalowa*, Warszawa: WNT, 1972
12. S. G. Michlin, C. L. Smolicki: *Metody przybliżone rozwiązywania równań różniczkowych i całkowych*. Warszawa: PWN, 1970
13. T. Morawski, W. Gwarek, *Teoria pola elektromagnetycznego*, Warszawa: WNT, 1985

14. A. Ralston: *Wstęp do analizy numerycznej.* Warszawa: PWN, 1983
15. K. R. Richter i in.: *Ein Beitrag zur Berechnung elektromagnetischer Felder mittels numerischer Methoden.* AEÜ, 1983 Band 37, Heft 5/6, pp. 174—182
16. H. R. Schwarz: *Methode der finiten Elemente,* Stuttgart: Teubner, 1980
17. J. H. Wilkinson: *The algebraic eigenvalue problem.* Oxford: Clarendon Press, 1965
18. O. C. Zienkiewicz: *Metoda elementów skończonych.* Warszawa: Arkady, 1972

K. ANISEROWICZ

APPLICATION OF FINITE ELEMENT METHOD TO THE ANALYSIS OF ELECTROMAGNETIC FIELDS IN LOSSLESS WAVEGUIDES WITH INHOMOGENEOUS DIELECTRIC FILLING

Summary

A finite element analysis of electromagnetic fields in lossless waveguides with inhomogeneous dielectric filling is described. Eight-node isoparametric elements of the second order are used. The matrix eigenproblem is solved by the subspace iteration and inverse vector iteration methods. Some examples of field plots are presented.

K. ANISEROWICZ

APPLICATION DE LA MÉTHODE DES ELEMENTS FINIS À L'ANALYSE DES CHAMPS ÉLECTROMAGNÉTIQUES DANS LES GUIDES D'ONDES À REMPLISSAGE DIÉLECTRIQUE HÉTÉROGÈNE

Résumé

En se servant de la méthode des éléments finis on présente l'analyse des champs électromagnétiques sans pertes d'énergie dans les guides d'ondes à remplissage diélectrique hétérogène. Dans les calculs on s'est servi d'éléments isoparamétriques à huit noeuds du second ordre. Le problème propre matericiel en déduit a été résolu par méthode d'itération souse-espace et d'itération vectorielle inverse. On a inclu à titre d'exemple des dessins de la répartition des champs.

K. ANISEROWICZ

ANWENDUNG DER METHODE FINITER ELEMENTE FÜR DIE ANALYSE VON ELEKTROMAGNETISCHEN FELDERN IN INHOMOGENEN VERLUSTLOSEN HOHLLEITERN

Zusammenfassung

Im Beitrag wurde eine Analyse der elektromagnetischen Felder in verlustlosen Hohlleitern mit inhomogener dielektrischer Füllung unter Anwendung der Methode finiter Elemente dargestellt. Für die Berechnungen wurden 8-knotige isoparametrische Elemente 2. Ordnung angewandt. Die aufgetretenen Matrixeigenwertprobleme wurden mittels simultaner und der einfachen Vektoriteration gelöst. Als Beispiele wurden Zeichnungen der Feldverteilung beigefügt.

К. АНИСЕРОВИЧ

ПРИМЕНЕНИЕ МЕТОДА КОНЕЧНЫХ ЭЛЕМЕНТОВ К АНАЛИЗУ ЭЛЕКТРОМАГНИТНЫХ ПОЛЕЙ В ВОЛНОВОДАХ С ГЕТЕРОГЕННЫМ ДИЭЛЕКТРИЧЕСКИМ ЗАПОЛНЕНИЕМ БЕЗ ПОТЕРЬ

Резюме

Представлен анализ электромагнитных полей в волноводах с гетерогенным диэлектрическим заполнением без потерь при применении метода конечных элементов. Расчеты проведены при использовании восьмимодовых изопараметрических элементов второго порядка. Матричная проблема собственных значений решена методом подпространных итераций и методом обратной векторной итерации. Приведены примеры диаграмм распределений полей.

ROZPRAWY ELEKTROTECHNICZNE, 1988, 34, z. 4, ss. 1077 — 1086

621.3.087.92

Linearyzacja przetwarzania analogowo-cyfrowego

ROMAN ŚWIĄTEK, WIESŁAW TARCZYŃSKI

Instytut Elektrotechniki, Wyższa Szkoła Inżynierska w Opolu

Otrzymano 1987.06.11

Autoryzowano do druku 1988.03.02

W artykule opisane zostały sposoby poprawy liniowości przetwornika analogowo-cyfrowego przez zastosowanie dodatkowych układów cyfrowych. Korekta jest realizowana przez obróbkę cyfrową wartości wyjściowej przetwornika tak, aby uzyskać charakterystykę liniową dla danej rozdzielczości. Układami cyfrowymi są translatory kodu lub systemy mikroprocesorowe wyposażone w program obsługi przetwornika.

1. WSTĘP

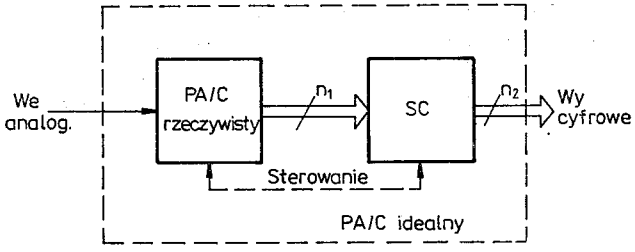
Wzrost zastosowań systemów mikroprocesorowych (μP) do celów metrologicznych i automatyki stwarza coraz większe zapotrzebowanie na układy przetworników analogowo-cyfrowych (PA/C) i cyfrowo-analogowych (PC/A). Przetworniki są jedynymi układami umożliwiającymi sprzężenie systemu μP z systemami analogowymi. Ponieważ zdecydowana większość obecnie konstruowanych systemów pomiarowych wykorzystuje system cyfrowy nie tylko do celów sterowania, ale także do przetwarzania, to jest oczywiste, że parametry przetwornika A/C decydują o metrologicznych własnościach tych systemów. Niezbędne zatem staje się udoskonalenie metod przetwarzania, dające w efekcie przede wszystkim skrócenie czasu przetwarzania oraz zwiększenie rozdzielczości.

Na świecie wiele firm produkuje różne typy przetworników A/C i C/A, co umożliwia wybór odpowiedniego układu do konkretnego zastosowania [1, 5]. W kraju brak jest układów przetworników, które łączyłyby w sobie w stopniu zadowalającym dwie najistotniejsze cechy, tj. szybkość i rozdzielczość. Produkowane układy hybrydowe są wprawdzie układami o względnie zadowalającej dokładności, ale za to są wolne, co uniemożliwia zastosowanie ich w wielu dziedzinach metrologicznych.

O parametrach przetwornika, zarówno dynamicznych jak i statycznych, decyduje konstrukcja układowa i technologia, według której został wykonany. Stawia to z kolei bardzo wysokie wymagania procesom technologicznym i pociąga za sobą konieczność budowy układów o bardzo rozbudowanej strukturze, a tym samym i kosztownych.

2. METODY KOREKCJI CHARAKTERYSTYKI PRZETWORNIKA

Udoskonalanie systemów cyfrowych, wzrost stopnia ich złożoności i możliwości funkcjonalnych przy jednoczesnej obniżce kosztów, skłania do zmiany w konstrukcji przetworników A/C. Możliwe staje się zatem przerzucenie odpowiedzialności za niektóre parametry przetwornika z jego części analogowej na rozbudowaną część cyfrową. Pozwala to z kolei na znaczne uproszczenie części analogowej pod względem technologicznym i układowym, kosztem parametrów statycznych (np. liniowości), które można skorygować w części cyfrowej tak, aby ostateczny sygnał cyfrowy odpowiadał przetworzeniu jak najbardziej zbliżonemu do idealnego. W takim przypadku staje się możliwe zastosowanie prostszych i zarazem szybszych oraz tańszych metod przetwarzania, np. metody z pojedynczym całkowaniem, która w klasycznych układach przetworników A/C nie mogła być powszechnie stosowana, głównie ze względu na znaczną nieliniowość i niestałość parametrów. Wynika stąd, że proces przetwarzania można podzielić na dwa zasadnicze etapy (rys. 1). W pierwszym — układ przetwornika A/C dokonuje konwersji,



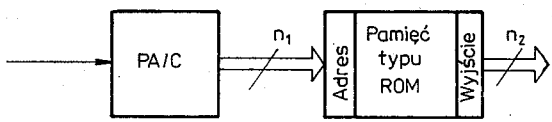
Rys. 1. Przetwornik A/C współpracujący z SC

w wyniku której otrzymuje się wielkość cyfrową odpowiadającą wejściowej wielkości analogowej, ale zniekształconą w wyniku niedoskonałości przetwornika. W drugim etapie ta wielkość cyfrowa zostaje przekształcona przez system cyfrowy (SC) i daje na wyjściu ostateczną wielkość cyfrową. Systemem cyfrowym może być w tym przypadku dowolny układ logiczny, który jest w stanie dokonać korekty wartości otrzymanej z przetwornika. Dołączenie do wyjścia przetwornika A/C systemu cyfrowego powoduje wydłużenie całkowitego czasu konwersji o czas potrzebny na przetwarzanie cyfrowe.

Spowoduje to zmniejszenie maksymalnej liczby przetworzeń w jednostce czasu przy zachowaniu czasu właściwej konwersji a/c. Nie ulegnie zatem zmianie wielkość dynamicznego błędu przetwarzania.

Korektę wyjściowej wielkości cyfrowej przetwornika A/C można przeprowadzić przez system cyfrowy w sposób układowy lub programowy. Przy korekcie układowej systemem cyfrowym jest układ logiczny o stałej konfiguracji, w którym podanie wielkości wejściowej powoduje wygenerowanie na wyjściu, z opóźnieniem wynikłym z czasu propagacji zastosowanych elementów, nowej skorygowanej wielkości cyfrowej. Tego typu układ pełni rolę translatora kodu i jest to metoda najszybsza w działaniu, lecz wymaga doboru układu indywidualnie do każdego przetwornika. Jednym ze sposobów realizacji tej metody może być zastosowanie jako translatora, pamięci typu ROM. Wyjściowa wielkość cyfrowa przetwornika A/C jest równocześnie adresem komórki pamięci, w której

znajduje się wartość skorygowana (rys. 2). Biorąc pod uwagę stopień złożoności układu, jest to metoda najprostsza w realizacji, ale musi być poprzedzona obliczeniami dającymi w wyniku zbiór wartości odpowiadających przetworzeniu idealnemu. Wartości te muszą być wpisane do pamięci. Należy zaznaczyć, że stosowanie translatora kodu wykonanego na bazie układów kombinacyjnych, ze względu na bardzo rozbudowaną strukturę, jest w tym przypadku mało przydatne i nie będzie brane pod uwagę w dalszych rozważaniach.



Rys. 2. Przetwornik A/C współpracujący z pamięcią stałą

Stosowanie metody programowej jest możliwe wtedy, gdy systemem cyfrowym jest system μP realizujący program obsługi przetwornika. Również i w tym przypadku wdrożenie metody musi być poprzedzone pomiarami identyfikującymi przetwornik, na bazie których zostaje opracowany program korekcji charakterystyki przetwarzania.

Opisane pobieżnie metody korekcji charakterystyk przetwornika A/C mogą być stosowane tylko dla niektórych parametrów statycznych przetwornika oraz dla układów o ściśle monotonicznych charakterystykach przetwarzania. Brak monotoniczności charakterystyki jest przyczyną gubienia słów kodowych, co uniemożliwia odtworzenie rzeczywistej wartości analogowej. Niemożliwe jest również, bez dodatkowych informacji o czynnikach zewnętrznych, dokonanie korekty wpływu temperatury oraz uwzględnienie procesu starzenia.

Wspólną cechą wszystkich metod korekcji jest konieczność wykonania wstępnych pomiarów dokładnie identyfikujących charakterystykę przetwornika. Na bazie tych danych oblicza się wartości cyfrowe odpowiadające żądanej charakterystyce przetwarzania. Polega to na określeniu wszystkich wartości cyfrowych wynikających z rozdzielczości przetwornika. Wykonanie tych pomiarów metodami klasycznymi jest, przy rozdzielczości przetwornika powyżej 5—6 bitów, czynnością żmudną i długotrwałą. Celowym więc staje się zastosowanie automatycznych systemów pomiarowych posiadających możliwość samoczynnego przeprowadzania pomiarów oraz rejestracji wartości danych wejściowych i wyników [3].

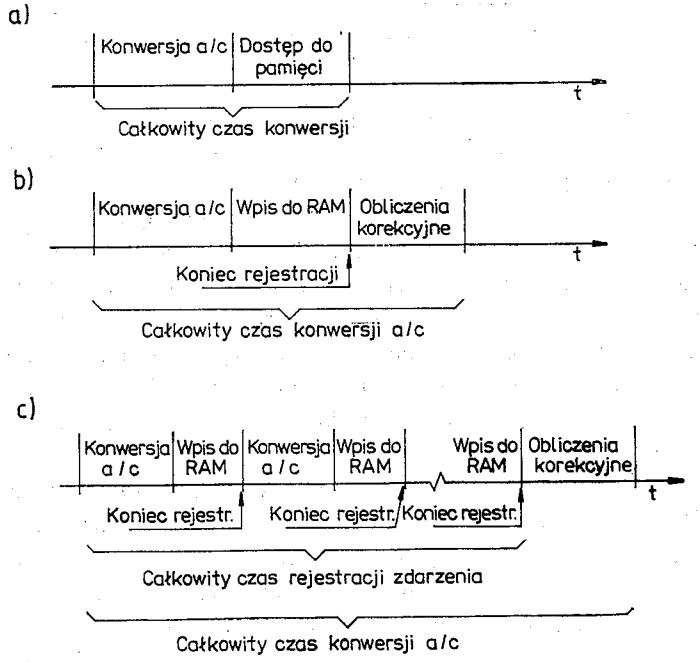
W większości zastosowań przetworników analogowo-cyfrowych wymagane jest, aby charakterystyka idealna przechodziła przez początek układu współrzędnych (tj. błąd niezrównoważenia był równy zeru) i była linią prostą w całym użytecznym przedziale zmienności. Nie dotyczy to specjalnych rozwiązań przetworników posiadających, z racji swojego przeznaczenia, charakterystykę nieliniową, np. logarytmiczną lub wykładniczą. Dla nich pojęcie idealnej charakterystyki musi być sformułowane oddzielnie. Ponieważ przetworniki tej grupy stosowane są bardzo rzadko, dlatego problemy związane z korektą tego typu charakterystyk nie będą omawiane w niniejszym opracowaniu.

Jest oczywiste, że wprowadzenie systemu cyfrowego zwiększa czas konwersji o czas potrzebny na obróbkę cyfrową wielkości wyjściowej przetwornika. Najmniejszy przyrost czasu można uzyskać stosując jako SC pamięć typu ROM. Przyrost czasu równy jest

wtedy czasowi dostępu do komórki pamięci i nie przekracza setek nsek. Ponadto jest on niezależny od rozdzielczości przetwornika (rys. 3a).

W przypadku metody programowej, zwiększenie czasu całkowitej konwersji A/C może być znaczne i będzie ściśle uzależnione od wielkości programu obsługi przetwornika oraz szybkości obliczeniowej samego systemu (rys. 3b).

Jednoznaczne ustalenie wielkości tego czasu jest możliwe tylko dla konkretnego zadania. W zadaniach, w których otrzymane z przetwornika dane nie muszą być natychmiast wykorzystywane w obliczeniach, można w pierwszej kolejności ograniczyć czynności SC tylko do zapamiętywania przez określony czas wartości dostarczanych przez przetwornik, a dopiero po tym czasie przeprowadzać obliczenia korekcyjne i ewentualnie inne obli-

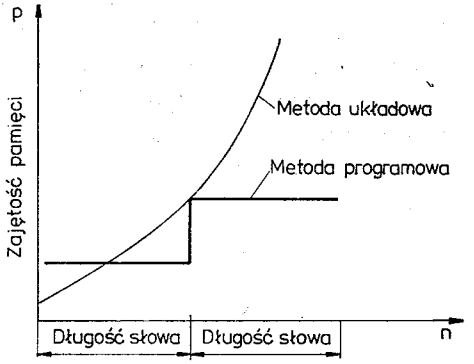


Rys. 3. Rozdział czasu w procesie konwersji a/c dla poszczególnych metod korekcji: a — metoda układowa z pamięcią typu ROM; b — metoda programowa ogólna; c — metoda programowa ze skomasowanym czasem rejestracji zdarzenia i obliczeń korekcyjnych

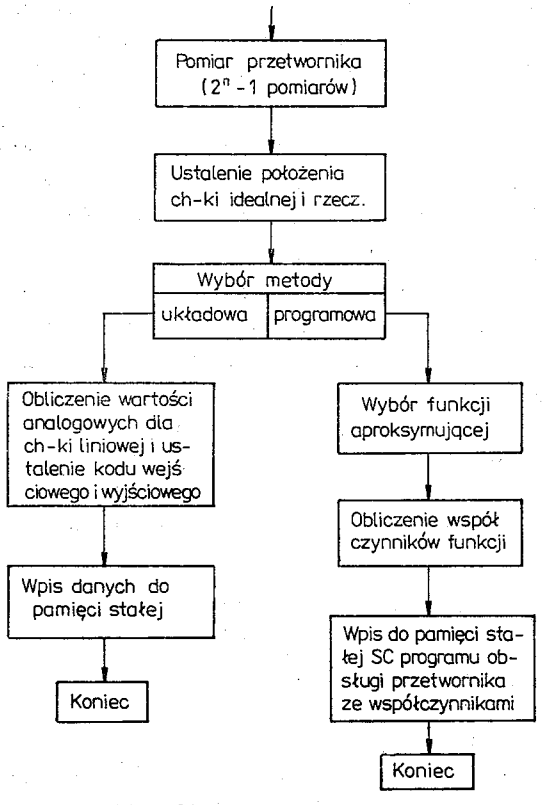
czenia (rys. 3c). Skomasowanie procesu rejestracji i przesunięcia w czasie obliczeń korekcyjnych po jego zakończeniu pozwala na zwiększenie częstotliwości wykonywanych pomiarów, a tym samym na dokładniejsze odwzorowanie przebiegu rzeczywistego. Ma to szczególne znaczenie przy analizie procesów szybkozmiennych o skończonym czasie trwania, np. stanów przejściowych.

Jeżeli czas konwersji nie jest czynnikiem jednoznacznie wyznaczającym metodę korekcji, to wyboru powinno dokonać się według kryterium zajętości pamięci stałej lub operacyjnej. Dla metody układowej zajętość pamięci jest wyznaczona jednoznacznie przez rozdzielczość przetwornika n_1 i wynosi $P_{max} = 2^{n_1}$ słów. Liczba bitów adresu musi być co najmniej równa rozdzielczości n_1, a długość słowa wyjściowego, rozdzielczości wyjściowej n_2.

Natomiast dla metody programowej, zajętość pamięci wyznaczona jest przez rozmiar programu obsługi przetwornika. Ponadto jest ona wielkością stałą i nie ulega zwiększeniu przy zmianie rozdzielczości przetwornika w obrębie długości słowa przetwarzanego przez SC (rys. 4). Skok w zajętości pamięci przy zwiększeniu rozdzielczości powyżej długości słowa przetwarzanego spowodowany jest koniecznością zmiany struktury programu, tak aby możliwe było wykonanie obliczeń na danych o zwiększonej długości słowa [2, 3].



Rys. 4. Zależność zajętości pamięci od rozdzielczości przetwornika



Rys. 5. Schemat blokowy procesu identyfikacji parametrów dla metod korekcji charakterystyki przetwornika A/C

Przy metodzie programowej ważne jest wybranie odpowiedniej funkcji aproksymującej charakterystykę rzeczywistą przetwornika, na podstawie której możliwe będzie przeprowadzenie korekcji. Jest oczywiste, że najlepsze przybliżenie otrzymuje się wtedy, gdy kształt naturalny funkcji odpowiada kształtowi funkcji aproksymowanej. Ponieważ pomiar charakterystyki przetwornika rzeczywistego obarczony jest błędem wynikającym ze skończonej dokładności przyrządów pomiarowych, dlatego współczynniki funkcji aproksymującej powinny być wyznaczone przez zastosowanie odpowiedniego kryterium minimalizującego błąd przybliżenia (np. metody najmniejszych kwadratów). Wynika stąd, że wymagane jest posiadanie dodatkowego systemu komputerowego z odpowiednim oprogramowaniem umożliwiającym wykonanie tego typu obliczeń.

Schemat blokowy procesu umożliwiającego uzyskanie w ostateczności zbioru danych, dla których SC będzie dokonywał przeliczeń wartości wyjściowych z przetwornika A/C, przedstawiony jest na rys. 5.

3. PRZYKŁADY

Powyższe metody korekcji charakterystyk przetwarzania przetworników A/C zostały sprawdzone praktycznie dla przetworników typu napięcie-cyfra, a otrzymane wyniki potwierdziły ich przydatność w zastosowaniu. Metoda układowa korekcji została sprawdzona dla przetwornika A/C wykorzystującego metodę z pojedynczym całkowaniem.

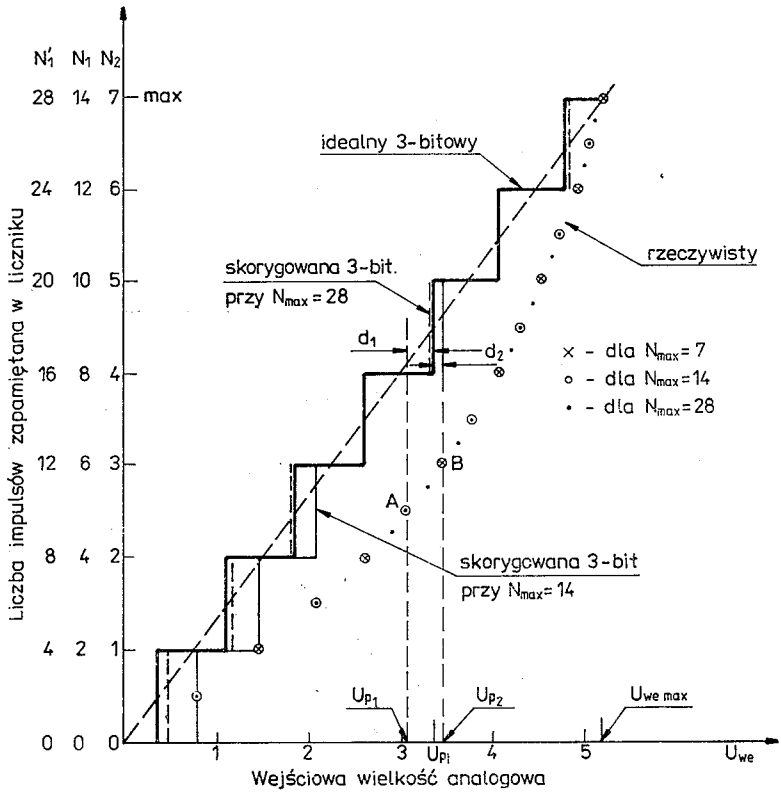
Przykładowe obliczenia wykonane zostały dla układu jak na rys. 2, przy założeniu, że rozdzielczość wyjściowa $n_2 = 3$. Przyjęto, że napięcie wzorcowe wytwarzane jest w układzie całkującym typu RC, a charakterystyka układu rzeczywistego i idealnego 3-bitowego, pokrywają się w punktach początkowym i końcowym.

Na charakterystyce rzeczywistej zaznaczone zostały punkty, którym odpowiadają wartości progowe napięć, tj. takich, przy których następuje zwiększenie zawartości licznika o 1. Liczba tych punktów wyznaczona jest przez częstotliwość generatora dobraną tak, aby w czasie potrzebnym na przetworzenie napięcia w pełnym zakresie liczba impulsów zarejestrowana w liczniku wynosiła kolejno 7, 14, 28.

Jeżeli założy się, że charakterystyka idealna 3-bitowa jest tą charakterystyką, którą powinno się otrzymać przez zastosowanie SC, to należy przyporządkować wyjściowym wartościom N_1 licznika wartości odpowiadające napięciom progowym układu rzeczywistego, znajdującym się najbliżej napięć progowych idealnego przetowrnika (rys. 6). Utworzona w ten sposób charakterystyka skorygowana nie w pełni pokrywa się z idealną.

Dla napięć wejściowych spełniających zależność $U_{p1} \leqslant U_{we} < U_{p2}$ wartość N_1 licznika wynosi 5, dla $U_{we} = U_{p2}$ stan N_1 wynosi 6. Między napięciami U_{p1} i U_{p2} jest napięcie progowe U_{pi} charakterystyki idealnej. Ponieważ wartość U_{p2} jest bliżej wartości U_{pi} niż napięcie $U_{p1}(d_2 < d_1)$, wartości binarne punktów A i B czyli wartości $N_1 = 5$ i $N_1 = 6$ system cyfrowy musi przetransponować na $N_2 = 4$.

Zwiększenie częstotliwości generatora powoduje zwiększenie liczby punktów na charakterystyce rzeczywistej i tym samym większe zbliżenie do charakterystyki idealnej 3-bitowej. Proces zwiększania częstotliwości generatora i tym samym rozdzielczości przetwor-

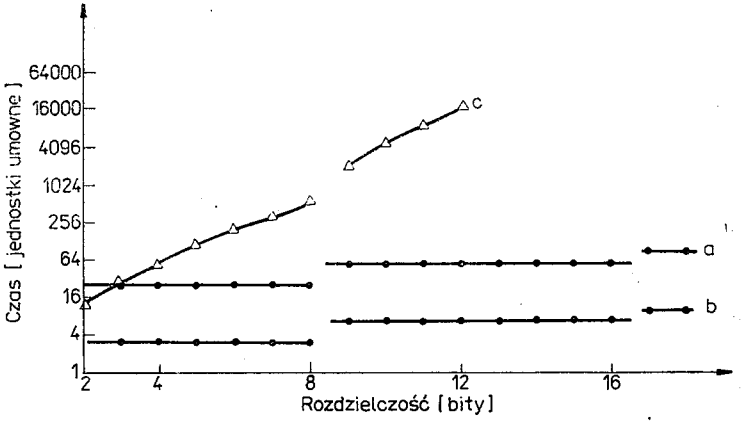


Rys. 6. Ilustracja wpływu rozdzielczości przetwornika rzeczywistego na charakterystykę wyjściową układu

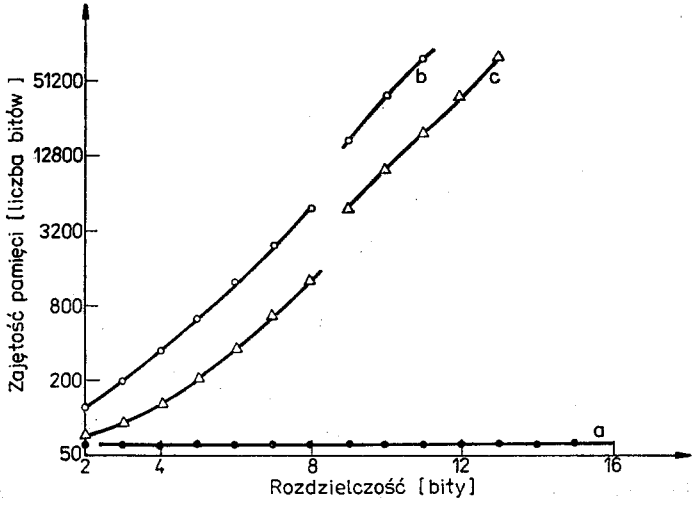
nika powinien być zakończony, jeżeli zostanie spełnione przyjęte kryterium oceny zgodności (np. błędu średniokwadratowego) napięć progowych charakterystyki idealnej i skorygowanej.

Dla omówionego przetwornika została zastosowana metoda programowa korekcji charakterystyki przejściowej i sprawdzona na systemie mikrokomputerowym. Równanie matematyczne charakterystyki rzeczywistej oraz jego współczynniki zostały ustalone na etapie przygotowania danych. Przebadano trzy metody korekcji programowej. Oszacowano zajętość pamięci i szybkość przetwarzania przy zastosowaniu każdej z metod. Sprawdzono także zakres zmian tych parametrów w zależności od rozdzielczości przetwornika (rys. 7 i 8). Metoda analityczna polega na obliczeniu rzeczywistej wartości napięcia wejściowego U_{wej} na podstawie równania charakterystyki rzeczywistej oraz aktualnej wartości wyjściowej przetwornika jako danej cyfrowej i zawarta jest w przedziale $[U_N, U_{N+1})$, gdzie U_N, U_{N+1} są wartościami napięć odpowiadającymi cyfrowym wartościom odpowiednio N, $N+1$. Zadaniem programu jest ustalenie przedziału napięć z charakterystyki skorygowanej, w której zawarte jest obliczone napięcie U, a następnie kodu, jaki temu przedziałowi odpowiada.

Dwie z proponowanych metod polegają na odczytywaniu z pamięci, wcześniej stab-



Rys. 7. Czas pozyskania wartości skorygowanej dla poszczególnych metod programowych: a — metoda analityczna; b — metoda adresowania elementu tablicy; c — metoda sekwencyjna



Rys. 8. Zajętość pamięci dla poszczególnych metod programowych: a — metoda analityczna; b — metoda adresowania elementu tablicy; c — metoda sekwencyjna

licowanych, wartości skorygowanych charakterystyki przetwornika. Przy stablicowaniu wszystkich możliwych wartości, uzyskanie wartości skorygowanej polega na bezpośrednim odczycie wartości spod adresu wskazywanego przez wartość cyfrową uzyskaną z wyjścia przetwornika rzeczywistego. Stablicowanie natomiast tylko tych wartości cyfrowych z wyjścia przetwornika, które odpowiadają zmianie wartości cyfrowej charakterystyki skorygowanej, powoduje, że poszukiwana wartość uzyskiwana jest w sposób sekwencyjny, tzn. że aktualna wartość z wyjścia przetwornika jest porównywalna z kolejnymi zapamiętanymi wartościami progowymi aż do odnalezienia wartości progowej najmniej się od niej różniącej. Wartością skorygowaną jest wtedy numer komórki, w której znajduje się ta wartość.

Z opisu proponowanych metod programowych oraz z wyników badań wynika, że najszybszą metodą jest metoda polegająca na stablicowaniu wszystkich wartości. Jej wadą jest największa zajętość pamięci. Program obsługi nie zajmie tu wiele miejsca. Najwięcej obszaru pamięci należy przeznaczyć na tablicę wartości. Czas pozyskania skorygowanej wartości jest stały dla wartości o długości nie przekraczającej wielokrotności bajtu. Natomiast zajętość pamięci zwiększa się dwukrotnie dla każdego następnego bitu zwiększającego rozdzielczość.

Metoda analityczna jest metodą najwolniejszą, przy czym czas uzyskania wartości skorygowanej jest stały dla rozdzielczości nie przekraczającej długości wielokrotności bajtu. Przekroczenie tej granicy zwiększa czas n-krotnie. Zajętość pamięci jest stała i najmniejsza ze wszystkich metod (rys. 8).

Metoda wykorzystująca tablice wartości progowych wymaga 4-krotnie mniejszego obszaru pamięci w stosunku do metody bazującej na stablicowaniu charakterystyki skorygowanej, jednak czas pozyskiwania skorygowanej wartości rośnie gwałtownie wraz ze wzrostem rozdzielczości przetwornika.

WNIOSKI KOŃCOWE

Powyższe przykłady potwierdzają możliwość zlinearyzowania charakterystyki rzeczywistej przetwornika A/C nawet znacznie odbiegającej od idealnej. Nie powodują zmiany takich parametrów przetwornika jak czas konwersji i zakres wielkości przetwarzanej. Wadą tej metody jest utrata rozdzielczości wyjściowej układu względem rozdzielczości samego przetwornika. Jednak jest to lepszy sposób poprawy liniowości niż preferowany w literaturze [5], a polegający na wykorzystaniu charakterystyki tylko w przedziale, w którym liniowość jest zadowalająca, co powoduje zmniejszenie zakresu wielkości przetwarzanej.

BIBLIOGRAFIA

1. Z. Kulka, A. Libura, M. Nadachowski: *Przetworniki analogowo-cyfrowe i cyfrowo-analogowe.* Warszawa: WKŁ, 1987
2. H. Małysiak, i inni: *Modułowe systemy mikrokomputerowe.* Warszawa: WNT, 1984
3. P. Misiurewicz: *Układy mikroprocesorowe.* Warszawa: WNT, 1983
4. J. R. Naylor: *Testing digital/analog and analog/digital converters.* IEEE Trans. on Circuits and Systems, 1978 July, vol. CAS-25, no 7, pp. 526—538
5. M. Łakomy, J. Zabrodzki: *Scalone przetworniki analogowo-cyfrowe i cyfrowo-analogowe.* Warszawa: PWN, 1985

R. ŚWIĄTEK, W. TARCZYŃSKI

LINEARISATION OF ANALOG-TO-DIGITAL CONVERSION

Summary

Some ways of improving the linearity of analog-digital converters are presented. A correction of the parameters of the converter is realized by digital processing of its output data so as to obtain linear characteristics at a given resolution. For this purpose, digital circuits such as code translators and microprocessor systems provided with a program of the converter operation.

R. ŚWIĄTEK, W. TARCZYŃSKI

ALIGNEMENT DU TRAITEMENT ANALOGIQUE-DIGITAL

Résumé

Dans l'article on a décrit les modes d'amélioration de la linéarité du convertisseur analogique-digital, à l'aide de l'application du circuits digitaux auxiliaires. La correction est effectuée par le traitement numérique de la valeur de sortie du convertiseur de telle sorte, qu'on obtienne la caractéristique linéaire pour un pouvoir séparateur donné. Comme circuits digitaux on utilise les traducteurs de code ou les systèmes microprocesseurs munis du programme de service du convertisseur.

R. ŚWIĄTEK, W. TARCZYŃSKI

GEWINNUNG DER LINEARITÄT BEI ANALOG-DIGITALER UMFORMUNG

Zusammenfassung

Im Beitrag werden Methoden beschrieben, die zur Verbesserung der Linearität des Analog-Digitalwandlers mittels zusätzlicher Digitalschaltkreise führen. Die Korrektur erfolgt über die Digitalverarbeitung des Eingangswertes des Wandlers in der Weise, damit die lineare Charakteristik des gegebenen Verteilungsvermögens erreicht wird. Als Digitalschaltungen betätigen sich hier Kodentransformatoren oder Mikroprozessorsysteme, die mit Bedienungsprogrammen des Wandlers ausgerüstet sind.

Р. СВИОНТЕК, В. ТАРЧИНЬСКИ

ЛИНЕАРИЗАЦИЯ ПРЕОБРАЗОВАНИЯ АНОЛОГ-ЦИФРА

Резюме

Описаны способы исправления линейности аналогово-цифровых преобразователей с применением добавочных цифровых цепей. Эти цепи реализуют цифровую обработку выходной величины преобразователя для получения линейной характеристики при данной разрешающей способности. Эти цепи состоят из трансляторов кода или из микропроцессорных систем, оснащенных программой обслуживания преобразователя.

ROZPRAWY ELEKTROTECHNICZNE, 1988, 34, z. 4, ss. 1087 — 1100

621.3—5:512.831

Algorytm generacji macierzy impedancyjnej węzłowej w obwodach zawierających źródła sterowane

MICHAŁ TADEUSIEWICZ, NASSER ASSI

Instytut Podstaw Elektrotechniki, Politechnika Łódzka

Otrzymano 1987.12.02

Autoryzowano do druku 1988.03.02

W pracy przedstawiono efektywną metodę generacji macierzy impedancyjnej węzłowej obejmującą obszerną klasę obwodów zawierających źródła sterowane. Wykorzystano macierzowy wzór Woodbury'ego oraz pewne koncepcje diakoptyki. Sformułowano, dogodny do implementacji komputerowej, algorytm oraz wskazano na jego zalety w stosunku do rozpowszechnionych metod El Abiada i Zollenkopfa. Przytoczono praktyczne przykłady zastosowania proponowanego algorytmu.

1. WPROWADZENIE

Symulacja komputerowa jest podstawowym narzędziem analizy i projektowania dużych i złożonych obwodów elektrycznych. W zakresie analizy AC problem polega na rozwiązaniu, często wielokrotnym, pewnego układu liniowych równań algebraicznych o współczynnikach zespolonych. W celu sformułowania tych równań często stosowana jest metoda napięć węzłowych prowadząca do opisu matematycznego o postaci

$$\boldsymbol{Yv} = \boldsymbol{j} \tag{1}$$

gdzie $\boldsymbol{v} = [V_1 \dots V_n]^t$ oraz $\boldsymbol{j} = [j_1 \dots j_n]^t$ oznaczają odpowiednio wektor napięć i prądów węzłowych, zaś $\boldsymbol{Y} = [Y_{ij}]_{n \times n}$ jest macierzą admitancyjną węzłową.

Równanie (1) można rozwiązać różnymi metodami numerycznymi jak np. metodą eliminacji Gaussa lub metodą Gaussa-Seidela. W celu wyznaczenia rozwiązania tego równania można również odwracać macierz otrzymując macierz impedancyjną węzłową $\boldsymbol{Z} = \boldsymbol{Y}^{-1}$ co pozwala sformułować zależność

$$\boldsymbol{v} = \boldsymbol{Zj} \tag{2}$$

Proces odwracania macierzy $\boldsymbol{Y}$ jest czasochłonny, ale znajomość macierzy $\boldsymbol{Z}$ pozwala rozpatrywać różne warianty problemu i daje pełną informację o równaniu (1).

W pewnych zastosowaniach macierz $\boldsymbol{Z}$ jest bardzo użyteczna lub wręcz niezbędna. Można tu wymienić obliczenia zwarciowe i rozpływ mocy w sieciach elektroenergetycznych [1—2], [4], [7], [9—10] oraz obliczanie wrażliwości wielu zmiennych przy dużej liczbie

Pracę wykonano w ramach CPBP 02.14

parametrów w układach elektronicznych [3]. W tego rodzaju zagadnieniach macierz Z odgrywa kluczową rolę.

Stały wzrost stopnia złożoności obwodów elektrycznych powoduje zapotrzebowanie na szczególnie efektywne algorytmy. Bezpośrednie metody odwracania macierzy, jak na przykład metoda eliminacji Gaussa wymagająca wykonania około n^3 długich operacji, są tu mało skuteczne. Dlatego opracowano pewne wyspecjalizowane metody odwracania macierzy Y, wśród których na szczególną uwagę zasługuje metoda Zollenkopfa [10] i metoda El-Abiada [2]. Obie te metody są dalece niewystarczające, zwłaszcza w przypadku złożonych układów zawierających źródła sterowane. Dlatego sprawą nadal otwartą jest opracowanie dostatecznie ogólnego szybkiego i efektywnego algorytmu obliczania macierzy impedancyjnej węzłowej dla obszernej klasy obwodów. W niniejszej pracy wykazano, że przy wykorzystaniu pewnych idei diakoptyki oraz macierzowego wzoru Woodbury'ego możliwe jest rozwiązanie sformułowanego wyżej problemu. Należy podkreślić, że zaproponowany algorytm wykazuje wyraźne zalety w stosunku do dotychczas stosowanych metod — jest szybszy oraz w mniejszym stopniu obciąża pamięć komputera.

2. ZALEŻNOŚCI PODSTAWOWE

Rozpatrzymy spójne obwody liniowe zawierające elementy RLC, źródła prądowe sterowane napięciowo (ŻPSN) oraz niezależne źródła prądowe i napięciowe. Założymy, że niezależne oraz/lub sterowane źródła prądowe nie tworzą żadnego przekroju, a źródła napięciowe nie tworzą żadnej pętli. Każde napięcie sterujące źródłem sterowanym będzie traktowane jako napięcie na gałęzi o zerowej admitancji dodatkowo włączonej do obwodu.

Zdefiniowany wyżej obwód zostanie opisany metodą węzłową, w sposób podobny jak w pracy [8].

1) Jeden z węzłów przyjmujemy za węzeł odniesienia. Spośród pozostałych węzłów wybieramy zbiór α wszystkich węzłów, do których nie jest dołączone żadne źródło napięcia. Następnie bierzemy pod uwagę zbiór węzłów połączonych źródłami napięcia i z każdego z tych zbiorów (wykluczając zbiór zawierający węzeł odniesienia, jeżeli taki istnieje), wybieramy dowolnie jeden węzeł. Zbiór węzłów wyznaczonych w ten sposób będzie oznaczony przez γ. Zbiór $\mu = \alpha \cup \gamma$ zawiera węzły niezależne. Napięcia między tymi węzłami a węzłem odniesienia tworzą wektor $\boldsymbol{v} = [V_1 \dots V_n]^t$. Pozostałe węzły (z wyłączeniem węzła odniesienia) w liczbie q są węzłami zależnymi. Napięcia między węzłami zależnymi a węzłem odniesienia tworzą wektor $(\boldsymbol{u}+\boldsymbol{e})_{q\times 1}$, gdzie składnikami wektora $\boldsymbol{u}$ są napięcia pomiędzy węzłami zbioru γ a węzłem odniesienia albo zera. Składowymi wektora $\boldsymbol{e}$ są kombinacje liniowe napięć źródłowych.

Wprowadzamy symbole Y, S, 0, E, J dla oznaczenia odpowiednio zbiorów gałęzi admitancyjnych, źródeł sterowanych, gałęzi o zerowych admitancjach, niezależnych źródeł napięcia oraz niezależnych źródeł prądu.

2) Tworzymy zredukowaną macierz incydencji A w ten sposób, że pierwsze wiersze odpowiadają niezależnym węzłom, a następne węzłom zależnym. Kolumny macierzy A odpowiadają kolejno elementom zbiorów Y, S, 0, E, J.

3) Przekształcamy macierz A dodając do wierszy odpowiadających węzłom zbioru γ odpowiednie wiersze związane z węzłami zależnymi w celu otrzymania zerowej podmacierzy

w przecięciu wierszy odpowiadających węzłom niezależnym i kolumn odpowiadającym źródłom napięcia. Tak przekształcona macierz incydencji będzie oznaczona przz A_r. W ten sposób podmacierz utworzona z n pierwszych wierszy macierzy A_r jest macierzą incydencji rozpatrywanego obwodu po zwarciu w nim źródeł napięcia.

4) Macierz A_r, o elementach 0, ± 1, można przedstawić następująco

$$A_r = \left[\begin{array}{c|c|c|c|c} A_Y & A_S & A_0 & 0 & A_J \\ \hline P_Y & P_S & P_0 & P_E & P_J \end{array}\right]. \tag{3}$$

Stąd otrzymujemy

$$A_r i_g = 0$$
$$A_r^t[v] = v_g$$

gdzie $i_g = [i_Y^t i_S^t i_0^t i_E^t i_J^t]^t$, $v_g = [v_Y^t v_S^t v_0^t v_E^t v_J^t]^t$ oznaczają odpowiednio wektory prądów i napięć gałęziowych. Stąd wynika:

$$v_Y = A_Y^t v + P_Y^t e$$
$$v_0 = A_0^t v + P_0^t e.$$

W rozpatrywanym obwodzie zachodzą następujące równania elementowe

$$i_Y = Y_Y v_Y \tag{4}$$

gdzie $Y_Y = \text{diag}\,(Y_1, \ldots, Y_b)$, przy czym b' jest liczbą gałęzi admitancyjnych obwodu,

$$i_S = C v_0 \tag{5}$$

gdzie $C = \text{diag}(c_1, \ldots, c_r)$, przy czym r jest liczbą ŻPSN oraz $c_i (i = 1, \ldots, r)$ oznacza współczynnik i-tego źródła sterowanego.

Powyższe zależności prowadzą do równania (1), gdzie

$$Y = A_Y Y_Y A_Y^t + A_S C A_0^t \tag{6}$$

$$j = -(A_Y Y_Y P_Y^t + A_S C P_0^t) e - A_J i_J. \tag{7}$$

3. METODA GENERACJI MACIERZY

Niżej zostanie omówiona metoda generacji macierzy impedancyjnej węzłowej $Z = Y^{-1}$, gdzie macierz Y jest określona wzorem (6). U podstaw proponowanego algorytmu leżą następujące zależności:

1) Macierzowy wzór Woodbury'ego [5].

Niech H, S, T, X będą odpowiednio macierzami o wymiarach $v \times w$, $w \times w$, $w \times v$, $v \times v$ oraz

$$\hat{X} = X + HST.$$

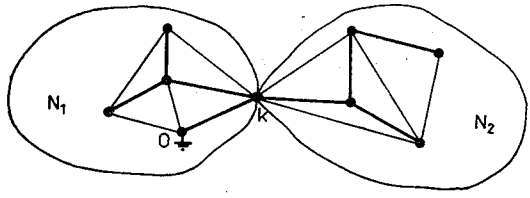
Zachodzi wzór Woodbury'ego

$$(X + HST)^{-1} = X^{-1} - X^{-1}H(S^{-1} + TX^{-1}H)^{-1}TX^{-1} \tag{8}$$

pozwalający wyznaczyć inwersję macierzy $\hat{X}$ w zależności od znanej inwersji macierzy X. Wzór jest bardzo efektywny jeżeli $w \ll v$.

2) *Twierdzenie 1*

Niech obwód N_1 o węzłach ponumerowanych od 0 do p (0 oznacza węzeł odniesienia) oraz obwód N_2 o węzłach $k, p+1, \ldots, p+l$; $(0 \leqslant k \leqslant p)$ mają tylko jeden wspólny węzeł k. Węzeł k przyjmujemy za węzeł odniesienia obwodu N_2 (rys. 1). Zakładamy, że sprzężenia



Rys. 1. Graf obwodu $N_{12} = N_1 \cup N_2$

(poprzez źródła sterowane) mogą wystąpić tylko w ramach poszczególnych obwodów N_1, N_2. Wówczas macierz impedancyjna węzłowa Z_{12} obwodu $N_{12} = N_1 \cup N_2$ ma następującą postać

$$Z_{12} = \left[\begin{array}{c|c} Z_{1_{(p\times p)}} & [Z_1^{(k)} Z_1^{(k)} \ldots Z_1^{(k)}]_{p\times l} \\ \hline \begin{bmatrix} {}_kZ_1 \\ {}_kZ_1 \\ \vdots \\ {}_kZ_1 \end{bmatrix}_{l\times p} & Z_{2_{(l\times l)}} + Z_{1_{kk}} \begin{bmatrix} 1 & 1 \ldots 1 \\ 1 & 1 \ldots 1 \\ . & . \ldots . \\ 1 & 1 \ldots 1 \end{bmatrix}_{l\times l} \end{array}\right] \tag{9}$$

gdzie

$$Z_1 = [Z_{1_{ij}}]_{p\times p} = [Z_1^{(1)} Z_1^{(2)} \ldots Z_1^{(p)}] = \begin{bmatrix} {}_1Z_1 \\ {}_2Z_1 \\ \vdots \\ {}_pZ_1 \end{bmatrix}, \quad i, j = 1, 2, \ldots, p$$

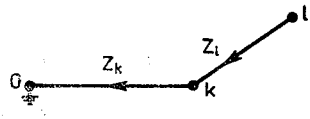
jest macierzą impedacyjną węzłową obwodu N_1; $Z_1^{(i)}$ $(i = 1, \ldots, p)$ jest i-tą kolumną macierzy Z_1; ${}_iZ_1$ jest i-tym wierszem macierzy Z_1; $Z_2 = [Z_{2_{ij}}]_{l\times l}$ $(i, j = 1, \ldots, l)$ jest macierzą impedancyjną węzłową obwodu N_2.

Dowód twierdzenia 1 podano w Dodatku.

W n i o s e k 1

Jeżeli obwód N_1 jest utworzony z gałęzi Z_k, a obwód N_2 z gałęzi Z_l połączonych w sposób pokazany na rys. 2, to

$$Z = \begin{bmatrix} Z_k & Z_k \\ Z_k & Z_l + Z_k \end{bmatrix}. \tag{10}$$



Rys. 2. Obwód elementarny

Wniosek 2

Jeżeli węzeł wspólny obwodów N_1 i N_2 jest węzłem odniesienia ($k = 0$), to

$$Z_{12} = \left[\begin{array}{c|c} Z_1 & \mathbf{0} \\ \hline \mathbf{0} & Z_2 \end{array}\right]. \tag{11}$$

Wniosek 3

Jeżeli obwody N_1 i N_2 nie zawierają źródeł sterowanych, to lewy dolny róg macierzy Z_{12} jest transpozycją prawego górnego rogu tej macierzy, czyli

$$\begin{bmatrix} {}_kZ_1 \\ \vdots \\ {}_kZ_1 \end{bmatrix} = [Z_1^k \dots Z_1^k]^t.$$

Rozpatrzymy obecnie obwód N zdefiniowany w rozdziale 2 i wyzerujemy w tym obwodzie wymuszenia źródeł niezależnych. Obwód podzielimy na drzewo T i dopełnienie L w ten sposób, żeby w dopełnieniu znalazły się wszystkie źródła prądowe sterowane napięciowo oraz gałęzie o zerowych admitancjach. Wzór (6) można wówczas przedstawić w postaci

$$Y = A_Y\left(\left[\begin{array}{c|c} Y_T & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right] + \left[\begin{array}{c|c|c} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & Y_{L_1} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right] + \dots + \left[\begin{array}{c|c} \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & Y_{L_m} \end{array}\right]\right) A_Y^t + \\ + A_s\left(\left[\begin{array}{c|c} c_1 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right] + \dots + \left[\begin{array}{c|c} \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & c_r \end{array}\right]\right) A_0^t, \tag{12}$$

gdzie $Y_T = \operatorname{diag}(Y_{T1}, \dots, Y_{T_n})$ oraz Y_{T_j} ($j = 1, \dots, n$) jest admitancją j-tej gałęzi drzewa; Y_{L_j} ($j = 1, \dots, m$) jest admitancją j-tej gałęzi dopełnienia, c_j ($j = 1, \dots, r$) jest współczynnikiem j-tego źródła sterowanego. Stąd wynika:

$$Y = A_T Y_T A_T^t + \sum_{j=1}^{m} A_{(n+j)} Y_{L_j} A_{(n+j)}^t + \sum_{j=1}^{r} A_s^{(j)} c_j (A_0^j)^t, \tag{13}$$

gdzie: A_T jest $n \times n$ podmacierzą zredukowanej macierzy incydencji A utworzoną z pierwszych n kolumn macierzy A odpowiadających gałęziom drzewa; $A_{(n+j)}$ oznacza $(n+j)$-tą kolumnę macierzy A związaną z j-tą admitancyjną gałęzią dopełnienia; $A_s^{(j)}$ jest $(n+m+j)$-tą kolumną macierzy A odpowiadającą j-temu źródłu sterowanemu; $A_0^{(j)}$ oznacza $(n+m+r+j)$-tą kolumnę macierzy A związaną z gałęzią o zerowej admitancji, której napięcie steruje j-tym źródłem prądowym. Obliczenia rozpoczynamy od określenia macierzy

$$Z^0 = (A_T Y_T A_T^t)^{-1}.$$

Następnie rozpatrujemy macierz

$$X_1 = A_T Y_Y A_T^t + A_{n+1} Y_{L_1} A_{n+1}^t$$

i dokonujemy następujących podstawień we wzorze (8):

$$X = A_T Y_Y A_T^t, \quad H = A_{n+1}, \quad T = A_{n+1}^t, \quad S = Y_{L_1}.$$

Otrzymujemy zależność

$$Z^1 = X_1^{-1} = Z^0 - Z^0 A_{n+1} (Y_{L_1}^{-1} + A_{n+1}^t Z^0 A_{n+1})^{-1} A_{n+1}^t Z^0,$$

gdzie macierz $W_1 = (Y_{L_1}^{-1} + A_{n+1}^t Z^0 A_{n+1})$, której inwersję należy obliczyć, ma wymiar 1×1. Powyższy proces obliczeniowy jest kontynuowany poprzez uwzględnienie, w analogiczny sposób, kolejnych składników sumy (13).

Znacznie lepsze rezultaty osiągnie się jeżeli opisaną metodę skojarzymy z pewnym koncepcjami diakoptyki. W tym celu dokonamy podziału obwodu N na dwa podobwody w następujący sposób. Obwód dzielimy na drzewo T i dopełnienie L według zasad omówionych wyżej. Następnie dzielimy drzewo T na dwa poddrzewa T_1 i T_2 mające jeden wspólny węzeł. Przyjmujemy, że węzeł odniesienia, jeżeli nie jest węzłem wspólnym, należy do poddrzewa T_1. Gałęzie dopełnienia dzielimy na trzy rozłączne zbiory L_1, L_2, L_{12} ($L = L_1 \cup L_2 \cup L_{12}$) tak, aby gałęzie podzbioru L_1 tworzyły fundamentalne pętle jedynie z gałęziami podrzewa T_1, gałęzie zbioru L_2 tylko z gałęziami poddrzewa T_2, a gałęzie zbioru L_{12} zarówno z gałęziami poddrzewa T_1 jak i T_2. Dla uproszczenia założymy, że dla każdego źródła sterowanego zarówno gałąź sterująca jak i sterowana należą do tego samego podzbioru dopełnienia. Proces obliczeniowy jest realizowany według następującego schematu:

1) Wyznaczamy macierz Z_1 dla obwodu N_1 złożonego z T_1 i L_1 stosując algorytm omówiony wyżej.

2) Analogicznie znajdujemy macierz Z_2 obwodu N_2 utworzonego z T_2 i L_2.

3) Obliczamy macierz Z_{12} obwodu $N_{12} = N_1 \cup N_2$ za pomocą wzoru (9).

4) Uwzględniamy poszczególne elementy zbioru L_{12} korzystając ze wzoru Woodbury'ego otrzymując ostatecznie poszukiwaną macierz Z.

U w a g a:

Jeżeli gałąź sterująca i sterowana jakiegoś źródła sterowanego należą do różnych zbiorów (np. L_1 oraz L_2), to źródło to usuwa się z obwodu na etapie wykonywania powyższego procesu obliczeniowego, a następnie uwzględnia się je stosując wzór Woodbury'ego.

Dalsze istotne usprawnienie i zwiększenie efektywności algorytmu osiągnie się dzieląc drzewo T na v poddrzew $T_1, T_2, \ldots, T_v$, tak aby jedno poddrzewo składało się tylko z dwóch (niekiedy jednej) gałęzi. Stosuje się tu zasady omówione wyżej, wobec czego dopełnienie zostaje podzielone na rozłączne podzbiory $L_1, \ldots, L_v, L_{12}, \ldots, L_{12\ldots v}$. Macierz impedancyjną węzłową wyznacza się stosując następujący algorytm:

Krok 1

Przygotowanie grafu obwodu i wprowadzenie danych wejściowych.

Krok 2

Wybór drzewa T, dopełnienia L oraz podział tych zbiorów na odpowiednie podstruktury.

Krok 3

Utworzenie macierzy impedancyjnej węzłowej dla podobwodów $N_{T_i L_i}$ w następujący sposób:

1) Obliczyć Z_{T_i} według wzoru (10).

2) Obliczyć Z_i' dla podobwodu $N_{T_i L_i}$ poprzez dołączenie admitancyjnych gałęzi dopełnienia stosując wzór Woodbury'ego.

3) Zmodyfikować za pomocą wzoru Woodbury'ego Z_i' uwzględniając odpowiednie źródła sterowane.

4) Powtórzyć czynności 1), 2), 3) aż $i = \nu$.

Krok 4

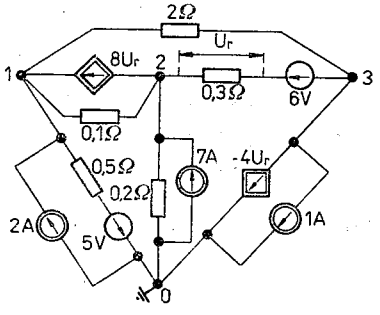
Utworzyć nowe podobwody poprzez łączenie dwóch kolejnych już istniejących, mających jeden węzeł wspólny, spełniających założenia twierdzenia 1. Wyznaczyć macierze impedancyjne węzłowe tych podobwodów stosując wzór (9) oraz wzór Woodbury'ego. Krok ten jest powtarzany do momentu otrzymania jednego podobwodu.

Uwaga:

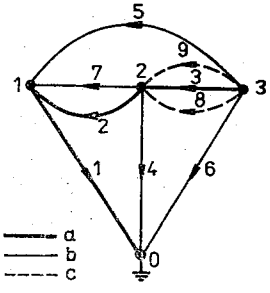
Gałęzi o zerowych admitancjach nie uwzględnia się w procesie obliczeniowym.

Przykład liczbowy

W obwodzie pokazanym na rys. 3 [3] należy wyznaczyć macierz impedancyjną węzłową Z. W tym celu dołączamy do obwodu dwie gałęzie o zerowych admitancjach (o numerach 8 i 9), usuwamy niezależne źródła prądowe, zwieramy niezależne źródła napięciowe i dzielimy graf obwodu na drzewo i dopełnienie (rys. 4). Dokonujemy podziału na podstruktury: $T_1 = \{1, 2\}$, $T_2 = \{3\}$, $L_1 = \{4\}$, $L_2 = \{\emptyset\}$, $L_{12} = \{5, 6, 7, 8, 9\}$.



Rys. 3. Obwód N



Rys. 4. Graf obwodu N
a — gałąź drzewa, b — gałąź dopełnienia, c — dodatkowa gałąź o zerowej admitancji

Macierze incydencji odpowiednich podzbiorów są następujące:

$$A_{T_1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \quad A_{T_2} = [1],$$

$$A_{L_1} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad A_{L_2} = [0], \quad A_{L_{12}} = \begin{bmatrix} -1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 \\ 1 & 1 & 0 & 1 & 1 \end{bmatrix}.$$

Rozpatrujemy odwód $N_{T_1L_1}$ i wyznaczamy kolejno

$$Z_{T_1} = \begin{bmatrix} 0{,}5 & 0{,}5 \\ 0{,}5 & 0{,}6 \end{bmatrix} \quad \text{oraz} \quad Z_1 = Z_{T_1L_1} = \begin{bmatrix} 0{,}1875 & 0{,}125 \\ 0{,}125 & 0{,}150 \end{bmatrix}.$$

Dla podobwodu $N_{T_2L_2}$ znajdujemy $Z_2 = 0{,}333$. Następnie tworzymy podobwód $N_{12} = N_{T_1L_1} \cup N_{T_2L_2}$ i wyznaczamy macierz Z_{12} ze wzoru (9):

$$Z_{12} = \begin{bmatrix} 0{,}1875 & 0{,}125 & 0{,}125 \\ 0{,}125 & 0{,}150 & 0{,}150 \\ 0{,}125 & 0{,}150 & 0{,}483 \end{bmatrix}.$$

Dalej dołączamy kolejno elementy zbioru L_{12}, t.j. gałęzie 5, 6 i 7 i stosując każdorazowo wzór Woodbury'ego otrzymujemy:

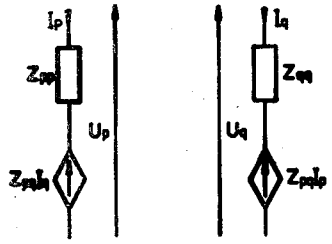
$$Z' = \begin{bmatrix} 0{,}1859 & 0{,}125 & 0{,}134 \\ 0{,}125 & 0{,}150 & 0{,}1464 \\ 0{,}134 & 0{,}1464 & 0{,}431 \end{bmatrix}, \quad Z'' = \begin{bmatrix} 0{,}151 & 0{,}138 & 0{,}961 \\ 0{,}087 & 0{,}165 & -1{,}057 \\ 0{,}022 & 0{,}191 & -3{,}114 \end{bmatrix},$$

$$Z = \begin{bmatrix} 0{,}147 & 0{,}140 & -1{,}072 \\ 0{,}110 & 0{,}155 & -0{,}313 \\ 0{,}072 & 0{,}170 & -1{,}503 \end{bmatrix}.$$

Gałęzi 8 i 9 o zerowych admitancjach nie uwzględnia się.

4. GENERACJA MACIERZY IMPEDANCYJNEJ WĘZŁOWEJ W OBWODACH ZAWIERAJĄCYCH CEWKI MAGNETYCZNIE SPRZĘŻONE

W omówionych w rozdziale 1 dziedzinach zastosowań macierzy impedacyjnej węzłowej zachodzi na ogół potrzeba rozpatrywania obwodów zawierających cewki magnetycznie sprzężone. Dotyczy to w szczególności sieci elektroenergetycznych, w których występują sprzężenia magnetyczne pomiędzy liniami. Przedstawiony w poprzednim rozdziale algorytm można stosować w omawianym przypadku po wprowadzeniu odpowiednich modeli cewek magnetycznie sprzężonych. W celu utworzenia tych modeli rozpatrzymy dwie gałęzie p oraz q zawierające cewki magnetycznie sprzężone. Ten układ dwóch gałęzi można przedstawić w sposób pokazany na rys. 5 wprowadzając źródła napięciowe sterowane



Rys. 5. Reprezentacja gałęzi p i q

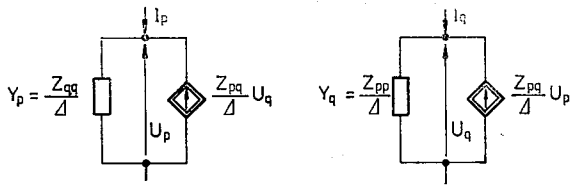
prądowo o współczynnikach $Z_{pq} = j\omega M_{pq}$. Obowiązuje tu następujące równanie

$$I_p = \frac{Z_{qq}}{\Delta} U_p - \frac{Z_{pq}}{\Delta} U_q$$
$$I_q = -\frac{Z_{pq}}{\Delta} U_p + \frac{Z_{pp}}{\Delta} U_q, \tag{14}$$

gdzie

$$\Delta = Z_{pp} Z_{qq} - Z_{pq}^2.$$

Równaniom (14) można przyporządkować obwód przedstawiony na rys. 6 zawierający dwa źródła prądowe sterowane napięciowo. W ten sposób możliwe staje się zastosowanie algorytmu z rodziału 3. Należy jednak zauważyć, że uwzględnienie w ten sposób sprzę-



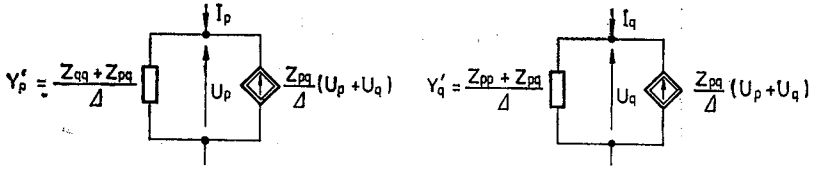
Rys. 6. Gałęziowe obwody odpowiadające równaniu (14)

żenia w procesie formułowania macierzy Z odbywa się w wyniku dwukrotnego zastosowania wzoru Woodbury'ego, co wydłuża obliczenia i jest niekorzystne. W celu usprawnienia tego fragmentu obliczeń zapiszemy równania (14) w sposób równoważny

$$\begin{bmatrix} I_p \\ I_q \end{bmatrix} = \begin{bmatrix} \frac{Z_{qq}+Z_{pq}}{\Delta} & 0 \\ 0 & \frac{Z_{pp}+Z_{pq}}{\Delta} \end{bmatrix} \begin{bmatrix} U_p \\ U_q \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \left(-\frac{Z_{pq}}{\Delta} \right) \begin{bmatrix} 1 \\ 1 \end{bmatrix}^t \begin{bmatrix} U_p \\ U_q \end{bmatrix}. \tag{15}$$

Równaniu (15) odpowiada obwód z rys. 7, a drugi składnik po prawej stronie tego równania ma postać umożliwiającą jednokrotne zastosowanie wzoru Woodbury'ego.

Zgodnie z powyższym, należy gałęzie p i q zawierające cewki magnetycznie sprzężone zastąpić modelem z rys. 7, co pozwoli uwzględnić w procesie formułowania macierzy Z obydwa źródła sterowane jednocześnie.



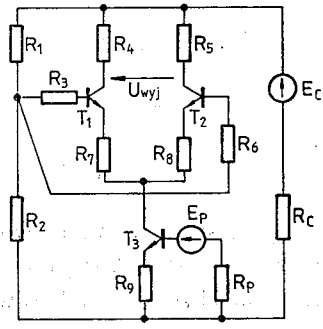
Rys. 7. Gałęziowe obwody odpowiadające równaniu (15)

Jeżeli gałąź jest sprzężona z m' innymi gałęziami ($m' > 1$), to należy wprowadzić $(m'-1)$ węzłów pozornych dzielących tę gałąź na m' podgałęzi o impedancjach $\frac{Z_i}{m'}$ tak, aby każda podgałąź była sprzężona wyłącznie z jedną inną gałęzią.

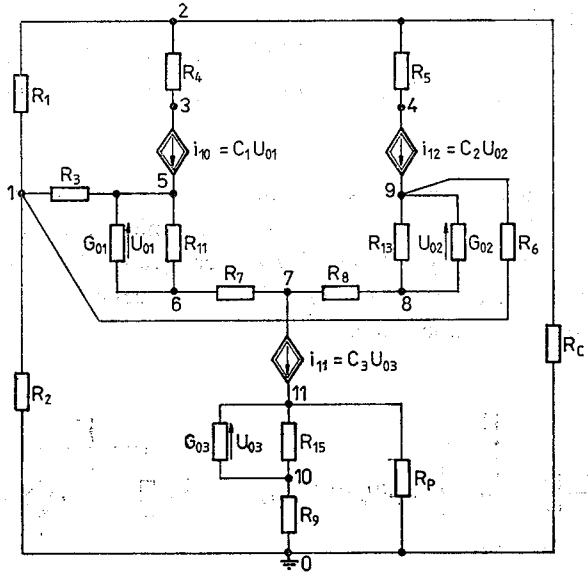
5. PRAKTYCZNE PRZYKŁADY ZASTOSOWANIA ALGORYTMU GENERACJI MACIERZY IMPEDANCYNEJ WĘZŁOWEJ

Przykład 1

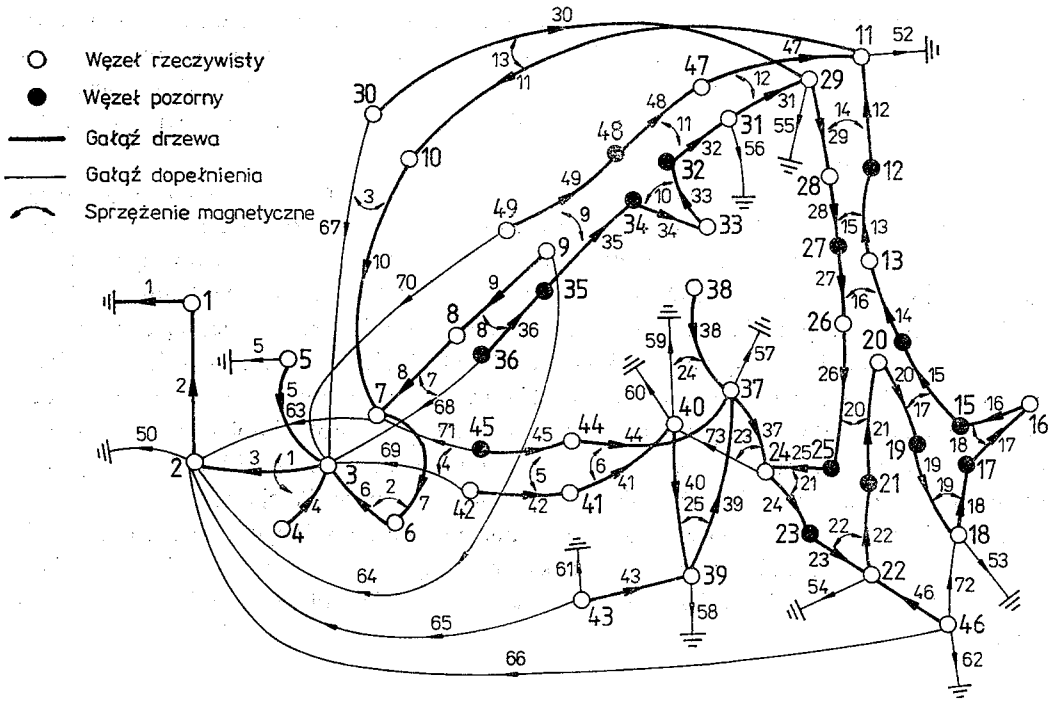
Na rys. 8 pokazano obwód wzmacniacza różnicowego, w którym należy przeprowadzić analizę wrażliwościową D.C. W tym celu zastąpiono tranzystory zredukowanymi modelami Ebersa-Molla, w następnie dokonano linearyzacji tego obwodu, otrzymując schemat z rys. 9 mający macierz impedancyjną węzłową Z. Znajomość tej macierzy



Rys. 8. Wzmacniacz różnicowy



Rys. 9. Obwód zlinearyzowany



Rys. 10. Graf przykładowej sieci elektroenergetycznej

pozwala natychmiast wyznaczyć wrażliwości każdego napięcia na zmianę dowolnego parametru. Stosując proponowany w pracy algorytm, wyznaczono macierz Z wykonując 344 długich operacji.

Przykład 2

Na rys. 10 przedstawiono graf schematu zastępczego sieci elektroenergetycznej podobnej do sieci ZE Warszawa-Miasto [6]. Po wprowadzeniu węzłów pozornych liczy ona 25 odcinków sprzężonych. Graf sieci zawiera 35 węzłów rzeczywistych, 15 węzłów pozornych, 75 gałęzi oraz 25 sprzężeń magnetycznych. Do obliczenia macierzy impedancyjnej węzłowej Z omawianego obwodu użyto prezentowany w pracy algorytm, co wymagało wykonania 30 492 długich operacji. Liczba ta jest kilkakrotnie mniejsza niż w metodzie El-Abiada.

UWAGI KOŃCOWE

W pracy przedstawiono algorytm generacji macierzy impedancyjnej węzłowej Z obejmujący obszerną klasę obwodów liniowych zawierających źródła prądowe sterowane napięciowo i cewki magnetycznie sprzężone. Zakres stosowalności algorytmu można rozszerzyć na praktycznie dowolne obwody liniowe dokonując wstępnie transformacji źródeł sterowanych w źródła prądowe sterowane napięciowo.

Macierz impedancyjna węzłowa $\boldsymbol{Z}$ odgrywa ważną rolę w zagadnieniach analizy i projektowania układów elektrycznych. Jest szeroko wykorzystywana w analizie sieci elektroenergetycznych przede wszystkim do obliczeń zwarciowych. Ponadto macierz $\boldsymbol{Z}$ jest bardzo przydatna do wyznaczania wrażliwości wielu zmiennych względem dużej liczby parametrów.

Przedstawiony w pracy algorytm pozwala wyznaczyć macierz $\boldsymbol{Z}$ bez potrzeby jawnego formułowania macierzy $\boldsymbol{Y}$. Algorytm jest skuteczny i charakteryzuje się małą liczbą długich operacji. Metoda wykazuje zalety w stosunku do powszechnie stosowanych metod Zollenkopfa i El-Abiada, a także w porównaniu do metody eliminacji Gaussa odwracania macierzy. Na przykład dla typowego obwodu drabinkowego zawierającego 50 węzłów metoda Zollenkopfa wymaga wykonania 11 350 długich operacji, metoda El-Abiada 67 326, metoda eliminacji Gaussa około 125 000, zaś proponowany w niniejszej pracy algorytm tylko 3 136. Efektywność algorytmu podnosi prostota implementacji komputerowej, oszczędność pamięci oraz łatwość wprowadzenia danych. Metoda, w sposób naturalny, pozwala uwzględnić zmiany spowodowane modyfikacją sieci.

Potencjalną wadą proponowanego algorytmu, podobnie jak innych metod, jest możliwość kumulowania błędów. Jednak liczne eksperymenty numeryczne wskazują, że dokładność otrzymanych wyników jest podobna jak w metodzie Zollenkopfa i nie budzi zastrzeżeń.

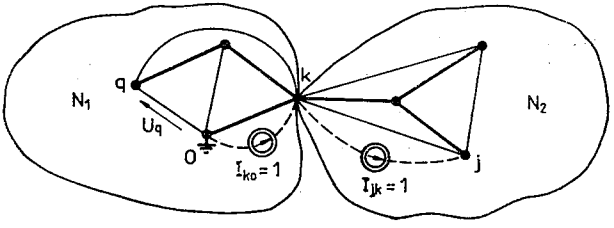
DODATEK

Dowód twierdzenia 1

Zauważmy, że element $(Z_{12})_{ij}$ macierzy $\boldsymbol{Z}_{12}$ jest określony zależnością

$$(Z_{12})_{ij} = U_i|_{I_j=1,\ I_k=0\ \ k\neq j\ \ k=1,\ldots,n^{\prime\prime}} \tag{D.1}$$

Wynika stąd wniosek, że w lewym górnym rogu macierzy $\boldsymbol{Z}_{12}$ występuje macierz $\boldsymbol{Z}_1$.



Rys. D.1. Obwód pomocniczy

Weźmiemy obecnie pod uwagę węzły i oraz j spełniające nierówności $1 \leqslant i \leqslant p$, $p+1 \leqslant j \leqslant p+l$ oraz określimy element $(Z_{12})_{ij}$. Zgodnie z (D.1) źródło prądowe $I_j = 1$ należy dołączyć do węzłów j oraz 0. Jest to równoważne dołączeniu dwóch źródeł $I_{jk} = 1$ oraz $I_{k0} = 1$ w sposób pokazany na rys. D.1. Ponieważ źródło I_{jk} nie wpływa na napięcie U_i, więc

$$(Z_{12})_{ij} = Z_{1_{ik}}$$

wobec czego w prawym górnym rogu macierzy Z_{12} wystąpi podmacierz utworzona z l jednakowych kolumn $Z_1^{(k)}$.

Jeżeli $p+1 \leqslant i \leqslant p+l$ oraz $1 \leqslant j \leqslant p$, to wszystkie napięcia pomiędzy węzłem k a węzłem i ($i = p+1, \ldots, p+l$) są równe zeru, a zatem $(Z_{12})_{ij} = Z_{1_{kj}}$. Tak więc w lewym dolnym rogu macierzy Z_{12} wystąpi podmacierz $[{}_kZ_1^t \ \ldots \ {}_kZ_1^t]^t$.

Do rozpatrzenia pozostał przypadek, gdy $p+1 \leqslant i \leqslant p+l$ oraz $p+1 \leqslant j \leqslant p+l$. Wówczas źródło prądowe I_j zastąpimy dwoma źródłami (rys. D.1). Ponieważ $U_i = U_{ik} + U_{k0}$, więc

$$(Z_{12})_{ij} = Z_{2_{ij}} + Z_{1_{kk}}$$

co wskazuje, że w prawym dolnym rogu macierzy Z_{12} występuje podmacierz

$$Z_2 + Z_{1_{kk}} \begin{bmatrix} 1 & 1 \ldots 1 \\ 1 & 1 \ldots 1 \\ 1 & 1 \ldots 1 \end{bmatrix}_{l \times l}$$

BIBLIOGRAFIA

1. D. Braess, E. Grebe: *A numerical analysis of load flow calculation methods.* IEEE Trans. Power Appar. & Systems, 1981, vol. PAS-100
2. H. E. Brown: *Solution of large networks by matrix methods.* New York—London—Sydney—Toronto: J. Wiley & Sons, 1975
3. L. O. Chua, P. M. Lin: *Komputerowa analiza układów elektronicznych.* Warszawa: WNT, 1981
4. H. H. Happ: *Piecewise methods and applications to power systems.* New York—Chichester—Brisbane—Toronto: J. Wiley & Sons, 1979
5. A. S. Householder: *The Theory of matrices in numerical analysis.* New York—Toronto—London: Blaisdell Publishing Company, 1965
6. P. Kacejko: *Obliczenie warunków zwarciowych w dużych systemach elektroenergetycznych w aspekcie wzrostu liczby sprzężonych ze sobą torów linii napowietrznych*, praca doktorska. Lublin 1982
7. R. Proulx, D. Crevier: *New interactive short-circuits calculation algorithm.* IEEE Trans. Power Appar. & Systems, 1982, vol. PAS-101
8. M. Tadeusiewicz: *DC analysis of piecewise-linear networks using the Gauss-Seidel method.* Int. J. Comp. Math. El. Electr. Eng., 1986, vol. 5, No. 4
9. M. Tadeusiewicz, N. Assi, G. Ciesielski: *An algorithm for generation of nodal impedance matrix for networks containing mutually conpled inductors.* I.J.EI. Pow. by. System, 1988, vol. 10, nr 1
10. K. Zollenkopf: *Sparse nodal impedance matrix generated by the bi-factorization method and applied to short circuit studies.* Proc. 5 th PSCC Conf., Cambridge, August 1975

M. TADEUSIEWICZ, N. ASSI

AN ALGORITHM FOR THE GENERATION OF A NODAL IMPEDANCE MATRIX FOR NETWORKS CONTAINING CONTROLLED SOURCES

Summary

An effective method for the generation of a nodal impedance matrix for a wide range of networks is developped. The method is formulated in terms of diacoptics and based Woodbury's formula. An algorithm usefull for computer implementation, having some advantages in comparison with the wide spread Zollenkopf and El-Abiad methods. Some practical examples are displayed.

M. TADEUSIEWICZ, N. ASSI

ALGORITHME DE GENERATION DE LA MATRICE IMPÉDANCE NOUEUSE POUR LES CIRCUITS CONTENANT DES SOURCES CONTRÔLÉES

Résumé

Dans l'article on a présenté la méthode effective ge génération de la matrice impédance noueuse pour une vaste classe de circuits, contenant des sources contrôlées. On a utilisé la formule de Woodbury et la conception de diacoptique. On a formé un algorithme, profitable pour l'implantation en ordinateur, qui a des avantages en comparaison avec des méthodes connues de Zollenkoopf et de El-Abiad. On a donné des exemples pratiques de l'application de l'algorithme proposé.

M. TADEUSIEWICZ, N. ASSI

ALGORITHMUS FÜR DIE GENERIERUNG EINER KNOTENINPEDANZMATRIX BEI STEUERUNGSQUELLEN ENTHALTENDEN STROMKREISEN

Zusammenfassung

In der vorliegenden Bearbeitung wurde eine effektive Methode für die Generierung einer Knoteninpedanzmatrix bei Steuerungsquellen enthaltenden Stromkreisen dargestellt. Man bediente sich dabei des Matrixmusters von Woodbury und gewisser Diakoptikkonzeptionen. Es wurde ein zur Computerimplementation geeigneter Algorithmus formuliert, und auf seine Vorteile den verbreiteten Methoden von El-Abiada und Zollenkopf gegenüber hingewiesen. Man hat praktische Beispiele gebracht, bei denen die Anwendung des gebotenen Algorithmus nützlich sein könnte.

М. ТАДЕУСЕВИЧ, Н. АССИ

АЛГОРИТМ ГЕНЕРАЦИИ ИМПЕДАНЦИОННОЙ УЗЛОВОЙ МАТРИЦЫ В ЦЕПЯХ СОДЕРЖАЩИХ УПРАВЛЯЮЩИЕ ИСТОЧНИКИ

Резюме

Изложен эффективный метод генерации импеданционной узловой матрицы, охватывающий широкий класс цепей содержащих управляемые источники. Использована матричная формула Вудбурега и некоторые концепции диакоптики. Сформулирован удобный для компьютерной имплементации алгоритм и подчёркнуты его достоинства по отношению к распространённым методом Эл Абеда и Золенкопфа. Приведены практические примеры предлагаемого алгоритма.

ROZPRAWY ELEKTROTECHNICZNE, 1988, 34, z. 4, ss. 1101 — 1112

621.372.852.11

Tworzenie układów z przełączanymi pojemnościami metodą morfologiczną

MAREK NAŁĘCZ, JAN MULAWKA

Instytut Podstaw Elektroniki, Politechnika Warszawska

Otrzymano 1987.01.05

Autoryzowano do druku 1988.01.04

W pracy zbadano transmitancje napięciowe wszystkich dwufazowych układów SC niewrażliwych na pojemności pasożytnicze, zbudowanych z jednego wzmacniacza operacyjnego i dwóch kondensatorów. Problem rozwiązano metodą morfologiczną, opisując układ za pomocą wielowrotników w dziedzinie z. Uzyskano dwie nowe ciekawe struktury. Ponadto sformułowano zaciskowe warunki pracy wielowrotników SC zapewniające eliminację wpływu pojemności pasożytniczych.

1. WSTĘP

W procesie projektowania filtrów z przełączanymi pojemnościami (SC) korzysta się z pewnych elementarnych bloków, jak np. integratory, opóźniacze, sumatory czy elementy mnożące. Bloki te tworzą podstawowe ogniwa, z których następnie budowane są bardziej skomplikowane układy. W literaturze podano już wiele różnych rozwiązań elementarnych bloków SC [3, 9], ale jak dotąd nie wiadomo, jakie jeszcze inne rozwiązania mogą istnieć. Dlatego interesujące wydaje się tworzenie pewnych klas układów SC. Rozważenie całej klasy struktur pozwala na kompletność ujęcia i umożliwia wybranie najlepszego (z punktu widzenia określonych kryteriów) rozwiązania spośród bloków o identycznych funkcjach układowych.

Zagadnienie to można zaliczyć do zadań kombinatorycznych, gdzie dane są reguły generacji pewnego przeliczalnego (zwykle skończonego) zbioru S zwanego przestrzenią rozwiązań, który zawiera zbiór wszystkich rozwiązań zadania. Rozwiązaniem jest każdy element S spełniający tzw. warunki zadania. Określone są także pewne warunki ograniczające, których niespełnienie powoduje usunięcie odpowiednich podzbiorów zbioru S ze zbioru potencjalnych rozwiązań. Problem polega na znalezieniu wszystkich rozwiązań. W praktyce może się to okazać bardzo trudne, gdyż moc zbioru S może być bardzo duża. Dlatego stosuje się przy tym zwykle dyrektywy heurystyczne. Podstawowe znaczenie ma jednak ograniczenie mocy przestrzeni rozwiązań na jak najwcześniejszym etapie poszukiwania, np. poprzez zmniejszenie jej wymiarowości.

W obszernej literaturze na temat układów SC znana jest praca Moschytza [5] doty-

cząca kombinatorycznych metod tworzenia nowych struktur. Wykorzystuje się w niej tzw. metodę morfologiczną do znalezienia bloków o określonej funkcji układowej. Tworzone układy SC opisywane są schematami w dziedzinie czasu i podejście to ograniczone jest tylko do klasy integratorów. W niniejszej pracy również za pomocą metody morfologicznej rozważono generację szerszej klasy układów. Zbadano mianowicie, jakie transmitancje napięciowe można uzyskać w klasie dwufazowych układów SC niewrażliwych na pojemności pasożytnicze, zbudowanych z jednego wzmacniacza operacyjnego (w.o.) i dwóch kondensatorów oraz pewnej liczby kluczy. Do opisu układów przyjęto schematy zastępcze w dziedzinie z, przez co jest to podejście alternatywne w porównaniu do metody opisanej w [5].

2. OKREŚLENIE KLASY GENEROWANYCH UKŁADÓW

Ze względu na ograniczenia technologiczne w technice układów SC dąży się do uzyskania rozwiązań niewrażliwych na pojemności pasożytnicze, co zostało wykazane w pracy [1]. Tylko takie bowiem układy umożliwiają wykorzystanie w pełni głównej zalety techniki SC, jaką jest możliwość bardzo dokładnej kontroli współczynników transmitancji napięciowej w technologii MOS.

W niniejszej pracy poszukiwać będziemy układów SC spełniających następujące warunki:

W1 — Układ zbudowany jest z jednego w.o., dwóch kondensatorów oraz pewnej liczby kluczy sterowanych zegarem dwufazowym o okresie T

W2 — Układ ma postać czwórnika o strukturze trójnikowej, którego wyjście jest wyjściem w.o., a interesującą funkcją układową — transmitancja napięciowa $H(z)$

W3 — Spełnione są warunki dostateczne niewrażliwości na pojemności pasożytnicze wg Haslera [2]

Przy założeniu idealności wszystkich elementów układ SC spełnia warunki dostateczne niewrażliwości na pojemności pasożytnicze jeśli:

a) w obu fazach zegara wszystkie węzły w układzie są postaci:
   typ 0 — masa
   typ I — masy pozorne
   typ II — źródła napięciowe (wejście układu lub wyjście w.o.)
   typ III — węzły incydentne w danej fazie tylko z pojedynczą okładką kondensatora, przy czym w drugiej fazie obie okładki tego kondensatora są incydentne z węzłami typu 0, I lub II

b) żadna okładka kondensatora nie jest przełączana pomiędzy węzłem typu I a węzłem typu II,

c) w układzie występuje minimalna liczba kluczy (tzn. nie ma kluczy redundancyjnych np. szeregowe połączenia kilku kluczy).

Należy zwrócić uwagę, że w pracy [2] jak również w innych pracach Haslera przeoczono możliwość istnienia węzłów typu III w układach niewrażliwych na pojemności pasożytnicze. Pominięcie tego typu węzłów oznaczałoby eliminację niektórych układów znanych już wcześniej jako niewrażliwe na pojemności pasożytnicze [3]. Ponieważ w niniej-

szej pracy nie zajmujemy się układami SC z kompensacją pojemności pasożytniczych [8], więc warunek W3 staje się również warunkiem koniecznym niewrażliwości układu na pojemności pasożytnicze [2].

3. OPIS UKŁADÓW SC ZA POMOCĄ PODSTAWOWYCH ELEMENTÓW SC

Przy założeniu, że wszystkie elementy układu SC są idealne, najdogodniejszym sposobem opisu są schematy zastępcze w dziedzinie z [4, 6, 7]. Najbardziej ogólny schemat zastępczy w dziedzinie z kondensatora przełączanego zegarem dwufazowym, podany przez Lakera [4], przedstawiony jest w tabl. 1 (poz. i). Jest to tzw. czterowrotnik SC. W większości praktycznych układów występują uproszczenia tego modelu, powstałe przez zwarcie do masy lub pozostawienie rozwartych pewnych wrót czterowrotnika. Wszystkie możliwe uproszczenia modelu Lakera tzw. podstawowe elementy SC zostały zebrane w tabl. 1 [7, 10]. Oczywiście elementy powstałe przez zamianę sterowania kluczy (e — 0) lub zmianę końcówek kondensatora (1—2) uważane są za równoważne. W omawianej tablicy przyjęto następujące oznaczenia: $q = z_z^{-1/2}$ oraz $p = 1 - z^{-1}$. Na schematach zastępczych w dziedzinie z podane zostały wartości admitancji elementów. Zauważmy, że element *ii* powstał przez zwarcie do masy jednych wrót czterowrotnika *i*, elementy *iii*, *iv* oraz *vi* przez zwarcie do masy pary wrót, zaś element *v* przez pozostawienie rozwartych jednych wrót lub pary wrót w tej samej fazie.

4. METODA TWORZENIA NOWYCH UKŁADÓW SC

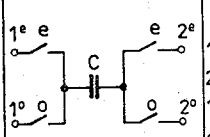
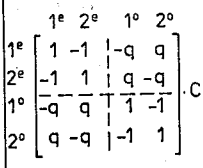
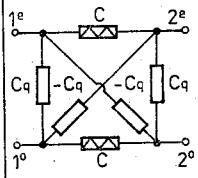
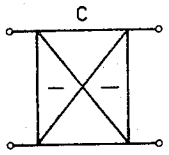
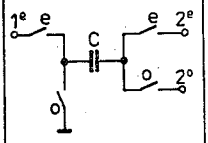
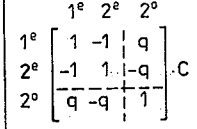
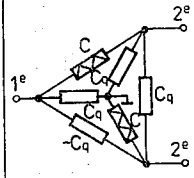
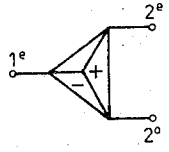
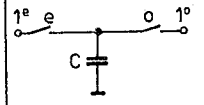
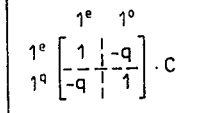
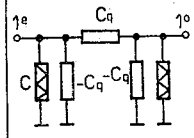
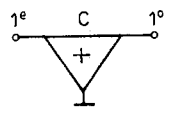
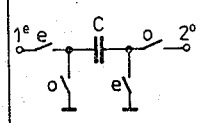
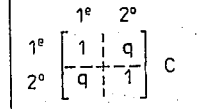
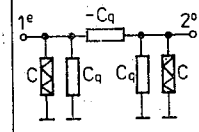
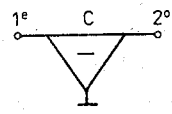
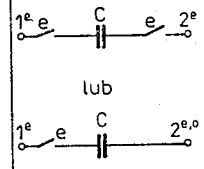
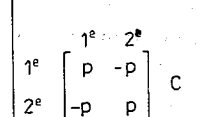
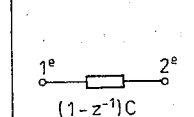
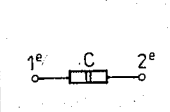
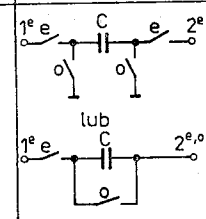
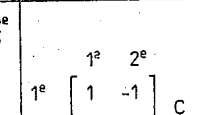
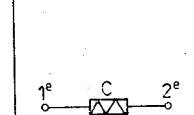
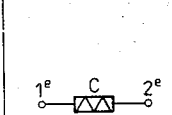
Jedną z metod rozwiązywania problemów kombinatorycznych jest metoda morfologiczna. Polega ona na rozbiciu rozwiązania na części składowe, znalezieniu dla każdej składowej wszystkich możliwych jej realizacji (tzw. zbioru formującego) oraz wyznaczeniu iloczynu kartezjańskiego zbiorów formujących odpowiadających wszystkim składowym. Wprowadzenie warunków ograniczających może spowodować usunięcie pewnych elementów tego iloczynu kartezjańskiego, skąd uzyskuje się ostateczny zbiór rozwiązań.

W rozważanej klasie układów SC (patrz punkt 2) można wyróżnić w dziedzinie z 3 elementy składowe: schematy zastępcze dla dwóch kondensatorów oraz jednego w.o. Zapewnienie poprawnej pracy w.o. wymaga, aby w każdej chwili (także w martwych strefach okresu zegara) miał on zamkniętą pętlę sprzężenia zwrotnego. Biorąc także pod uwagę warunek W3a przyjmiemy, że jeden z kondensatorów będzie na stałe włączony w pętlę ujemnego sprzężenia zwrotnego w.o., przy czym równolegle do niego może być przyłączony klucz. Wejście nieodwracające w.o. musi natomiast być umasione. Schemat zastępczy w dziedzinie z takiego podukładu można przyjąć za pierwszą składową rozwiązania, a schemat zastępczy drugiego kondensatora — za drugą składową.

4.1. OKREŚLENIE ZBIORÓW FORMUJĄCYCH

Zbiór formujący S_I odpowiadający w.o. i kondensatorowi C_0 w jego pętli sprzężenia zwrotnego został przedstawiony w tabl. 2. Zawiera ona wszystkie możliwe przypadki

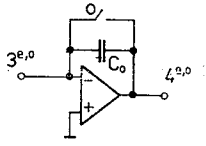
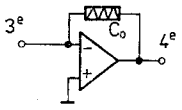
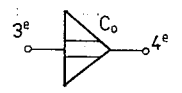
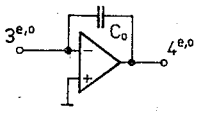
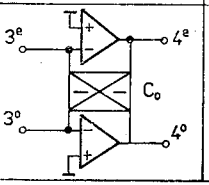
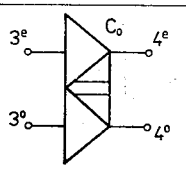
Podstawowe elementy SC

Typ	Schemat w dziedzinie t	Macierz admitancyjna	Schemat zastępczy w dziedzinie z	Symbol	Nazwa
i					Czterowrotnik SC
ii					Trzywrotnik SC
iii					Czwórnik + LPT
iv					Czwórnik - LPT
v					Dwójnik storancyjny
vi					Dwójnik oporowy

gdzie $q = z^{-1/2}$ $p = 1 - z^{-1}$

Tablica 2

Elementy zbioru formującego S_I

Typ	Schemat w dziedzinie t	Schemat zastępczy w dziedzinie z	Symbol
1			
2			

(podukłady powstałe przez zamianę sterowania kluczy uważamy za równoważne). Typ elementów S_I równy jest liczbie poziomów fazowych, na których występuje w.o. w schemacie zastępczym w dziedzinie z. W dalszej części pracy przyjmiemy C_0 za jednostkę pojemności i wprowadzimy oznaczenie $\alpha = C_1/C_0$, gdzie C_1 jest pojemnością drugiego kondensatora układu.

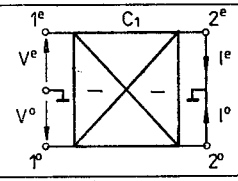
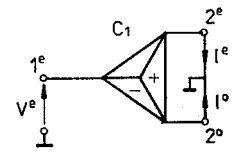
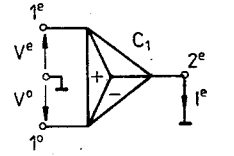
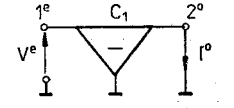
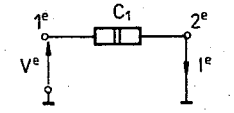
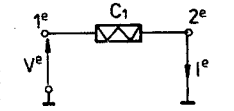
Jak wynika z tabl. 2 oraz z warunku W2, kondensator C_1 musi realizować macierz transmitancji od wejścia układu i wyjścia (wyjść) podukładów z S_I do wejścia (wejść) podukładów z S_I. Schemat zastępczy kondensatora C_1 w dziedzinie z może przybrać postać dowolnego wielowrotnika z tabl. 1. Przy tym o ile wyjścia i wejścia podukładów z S_I mogą być niewykorzystane (rozwarte), o tyle żadne wrota wielowrotnika odpowiadającego C_1 nie mogą być zwarte do masy ani pozostawione rozwartymi. Odpowiadałoby to bowiem innemu uproszczeniu modelu Lakera, o mniejszej liczbie wrót, a zatem powodowało niepotrzebną redundancję elementów iloczynu kartezjańskiego zbiorów formujących. Stąd oraz z warunku W3*b* wynika, że wrota $1^{e,o}(2^{e,o})$ wielowrotników z tabl. 1 mogą być połączone tylko z węzłem typu II (I). Otrzymany na podstawie powyższych rozważań zbiór formujący S_{II} odpowiadający kondensatorowi C_1 został przedstawiony w tabl. 3. Zawiera ona wszystkie możliwe przypadki (elementy powstałe przez zamianę sterowania kluczy $(e-o)$ uważamy za równoważne). W tabl. 3 przyjęto, że węzeł $1^{e,o}$ jest typu II, a węzeł $2^{e,o}$ — typu I.

4.2. OKREŚLENIE SKRZYNKI MORFOLOGICZNEJ

Do wyznaczenia iloczynu kartezjańskiego zbiorów formujących wykorzystuje się tzw. skrzynkę morfologiczną. W tabl. 4 przedstawiono skrzynkę morfologiczną $S_I \times S_{II}$. Zawiera ona 11 różnych struktur SC. Zgodnie z poprzednimi uwagami elementy A i B z S_{II} nie mogą być łączone z elementem 1 z S_I. Istnieją natomiast dwa możliwe połączenia elementu C z S_{II} z elementem 2 z S_I. Każda struktura została oznaczona symbolem,

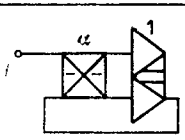
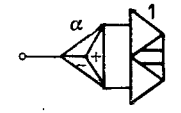
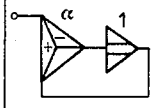
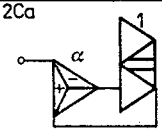
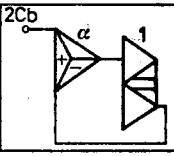
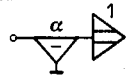
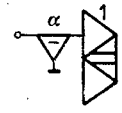
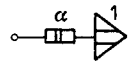
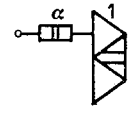
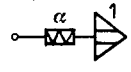
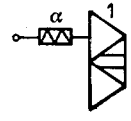
Tablica 3

Elementy zbioru formującego S_{II}

Typ	Schemat zastępczy w dziedzinie z
A	
B	
C	
D	
E	
F	

Tablica 4

Skrzynka morfologiczna $S_I \times S_{II}$

S_I / S_{II}	1	2
A	—	2A 
B	—	2B 
C	1C 	2Ca  2Cb 
D	1D 	2D 
E	1E 	2E 
F	1F 	2F 

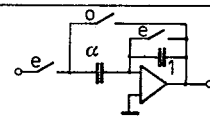
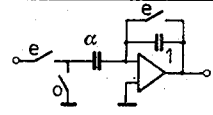
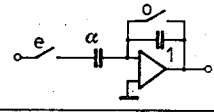
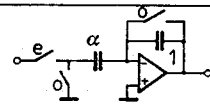
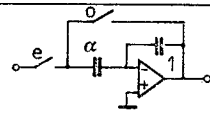
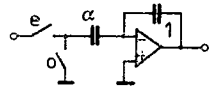
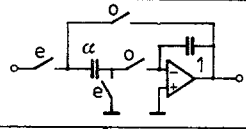
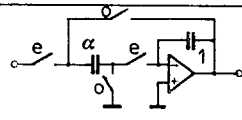
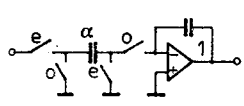
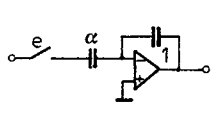
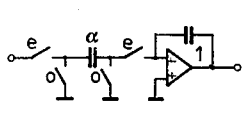
stanowiącym połączenie nazw typów jej elementów składowych. Wyniki analizy otrzymanych struktur zostaną zaprezentowane w następnym punkcie.

5. PRZEGLĄD OTRZYMANYCH ROZWIĄZAŃ

Dla każdej z otrzymanych w wyniku generacji struktur wyznaczono transmitancje napięciowe $H^{ee,oe}(z) = V_o^{e,o}(z)/V_i^e(z)$, gdzie indeksy o, i oznaczają odpowiednio wyjście

Tablica 5

Realizacje wynikające z Tablicy 4

Symbol	Transmitancje	Funkcja	Realizacja
1C	$H^{ee} = 0$ $H^{oe} = \frac{\alpha}{1+\alpha} z^{-1/2}$	+DLY	
1D	$H^{ee} = 0$ $H^{oe} = \alpha z^{-1/2}$	+DLY	
1E	$H^{ee} = -\alpha(1 - z^{-1})$ $H^{oe} = 0$	−BDD	
1F	$H^{ee} = -\alpha$ $H^{oe} = 0$	−MLX	
2A	$H^{ee} = -\alpha$ $H^{oe} = 0$	−MLX	
2B	$H^{ee} = -\alpha$ $H^{oe} = 0$	−MLX	
2Ca	$H^{ee} = \frac{\alpha z^{-1}}{(1+\alpha) - z^{-1}}$ $H^{oe} = \frac{\alpha z^{-1/2}}{(1+\alpha) - z^{-1}}$	+LPF +LPF	
2Cb	$H^{ee} = \frac{-\alpha}{1 - (1+\alpha) z^{-1}}$ $H^{oe} = \frac{-\alpha z^{-1/2}}{1 - (1+\alpha) z^{-1}}$	−AST −AST	
2D	$H^{ee} = \frac{\alpha z^{-1}}{1 - z^{-1}}$ $H^{oe} = \frac{\alpha z^{-1/2}}{1 - z^{-1}}$	+FDI +LDI	
2E	$H^{ee} = -\alpha$ $H^{oe} = -\alpha z^{-1/2}$	−MLX −DLY	
2F	$H^{ee} = \frac{-\alpha}{1 - z^{-1}}$ $H^{oe} = \frac{-\alpha z^{-1/2}}{1 - z^{-1}}$	−BDI −LDI	

i wejście układu. Obliczenia zostały wykonane wg metody opisanej w [6]. Następnie każdą z transmitancji sklasyfikowano pod względem funkcji wypełnianej w systemie dyskretnego przetwarzania sygnałów przez układ opisany taką transmitancją. Zastosowano przy tym następujące skróty:

MLX — układ mnożący
DLY — opóźniacz o $T/2$
BDD — układ różniczkujący BD
BDI — integrator BD
LDI — integrator LDI
FDI — integrator FD
LPF — filtr dolnoprzepustowy
AST — układ niestabilny
(+) — układ nieodwracający
(−) — układ odwracający

Otrzymane rozwiązania zostały zebrane w tabl. 5. Dla każdej struktury została określona ponadto jej realizacja w dziedzinie czasu, o minimalnej liczbie kluczy zapewniająca spełnienie warunku W3. Należy zauważyć, że dla drugiej realizacji elementarnego dwójnika oporowego SC (pozycja v_i tabl. 1) nie jest spełniony warunek W3*b* po zastosowaniu tego dwójnika w takich zaciskowych warunkach pracy, jak w pozycji *F* tabl. 3.

Bardziej szczegółowe rozważenie otrzymanych rozwiązań wskazuje, że nie wszystkie struktury nadają się do praktycznego zastosowania.

Z oczywistych względów należy np. wyeliminować niestabilny dla dowolnej wartości α układ 2*Cb*. Ponadto w układach 2*A*, 2*B* i 2*E* występuje niestabilność *DC* związana ze stałą degradacją ładunku w węźle wejściowym w.o. (ładunek ten zmniejsza się wskutek istnienia wejściowego prądu polaryzacji w.o. i nie może być w żaden sposób zregenerowany). Pozostałe struktury natomiast nadają się do praktycznego wykorzystania. Tworzą one pewien katalog bloków funkcjonalnych przedstawiony w tabl. 6.

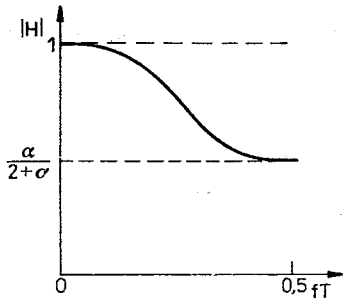
Tablica 6

Funkcje możliwe do zrealizowania w rozważanej klasie układów

Funkcja	Zalecana struktura
−MLX	1F
+DLY	1C, 1D
−BDD	1E
−BDI	2F
+LDI	2D
−LDI	2F
+FDI	2D
+LPF	2Ca

Należy stwierdzić, że większość otrzymanych struktur była już znana w literaturze (np. 2*D* i 2*F* w pracy [9], 1*D*, 1*E*, 1*F* i 2*B* w [3]). Nie są natomiast znane w dostępnej literaturze opracowania, w których byłyby prezentowane układy 1*C* lub 2*Ca* odznaczające się pewnymi interesującymi właściwościami. Otóż, układ 1*C* może zastąpić strukturę 1*D*, jeżeli wymagane wzmocnienie opóźniacza jest mniejsze od jedności. Dla takiego samego

wzmocnienia k układ $1C$ ma $1/(1-k)$ raz mniejszą wrażliwość $S_\alpha^{|H|}$ niż układ $1D$, przy czym wzrasta wtedy stosunek pojemności tyle samo razy. Z kolei układ $2Ca$ stanowi jedną z najprostszych realizacji filtru dolnoprzepustowego. Przybliżoną charakterystykę amplitudową tego układu przedstawiono na rysunku 1. Jak wynika z wykresu, odstęp między pasmem przenoszenia a tłumieniowym zależy od wartości parametru α. Przykładowo dla $\alpha = 0.1$ tłumienie na częstotliwości Nyquista przekracza 26 dB, co dla wielu praktycznych zastosowań może być wystarczające.



Charakterystyka amplitudowa układu $2Ca$

WNIOSKI

Przedstawione rozważania wykazują, że przyjęcie opisu układu SC za pomocą podstawowych elementów SC w dziedzinie z [7, 10] dzięki swej uniwersalności i prostej postaci dobrze nadaje się do tworzenia struktur metodami kombinatorycznymi. Widać ponadto, że przeanalizowanie w ten sposób szerszej klasy układów (np. z jednym w.o. i trzema kondensatorami) wymagałoby już użycia komputera. W wyniku przeprowadzonej generacji układów niewrażliwych na pojemności pasożytnicze, zbudowanych z jednego w.o. i dwóch kondensatorów, otrzymano obok szeregu znanych struktur także dwa nowe układy. Są to: opóźniacz i filtr dolnoprzepustowy o ciekawych właściwościach (odpowiednio pozycje $1C$ i $2Ca$ w tabl. 5). Ponadto jako produkt uboczny sformułowano zaciskowe warunki pracy wielowrotników SC, przy których spełnione są warunki W3 niewrażliwości układu na pojemności pasożytnicze (tabl. 3). Warunki te były dotychczas formułowane jedynie w dziedzinie czasu lub w języku macierzy admitancyjnych [2], w dodatku w sposób niekompletny.

BIBLIOGRAFIA

1. G. Fischer, G. S. Moschytz: *On the frequency limitations of SC filters.* IEEE J. Solid-State Circ., 1984, vol. SC-19, no. 4, pp. 510—518
2. M. Hasler, M. Saghafi: *Stray capacitance eliminating transformations for switched-capacitor circuits.* Int. J. Circ. Theor. Appl. 1983, vol. 11, np. 3, pp. 321—338
3. V. J. Hosticka, G. S. Moschytz: *Practical design of switched-capacitor networks for integrated circuit implementation.* Electron. Circ. Syst., 1979, vol. 3, no. 2, pp. 76—87
4. K. R. Laker: *Equivalent circuits for the analysis and synthesis of switched-capacitor networks.* Bell Syst. Tech. J., 1979, vol. 58, no. 3, pp. 729—769

5. G. S. Moschytz: *A morphological approach to switched-capacitor low-frequency integrator design.* AGEN — Mitteilungen, 1981, no. 32, pp. 7—15
6. J. J. Mulawka: *By — inspection analysis of switched-capacitor networks.* Int. J. Electron., 1980, vol. 49, no. 5, pp. 359—373
7. J. J. Mulawka, G. S. Moschytz: *Method of simplified analysis of switched-capacitor networks.* Electron. Lett., 1985, vol. 21, no. 4, pp. 138—139
8. J. J. Mulawka, M. Nałęcz: *Design of strays-insensitive switched-capacitor two-ports.* Proc. ECCTD'85, Prague, 1985, pp. 765—768
9. H. Weinrichter, J. A. Nossek: *Switched-capacitor-filters: a comparison i of the basic design principles.* Proc. ECCTD'80, Warsaw, 1980, vol. 1, pp. 17—28
10. J. Mulawka: *Zagadnienia syntezy układów analogowych z przełączanymi pojemnościami.* Prace Naukowe, Elektronika, Wydawnictwa Politechniki Warszawskiej, 1985, z. 68

M. NAŁĘCZ, J. MULAWKA

GENERATION OF SWITCHED-CAPACITOR NETWORKS WITH A MORPHOLOGICAL APPROACH

Summary

Voltage transfer functions of all two-phase SC networks unsensible to stray capacitance, comprising a single operational amplifier and two capacitors, are examined. The problem of circuit generation is solved by the morphological method using SC multiports described in the z-domain. In consequence, two new interesting structures are derived. Moreover, conditions for connecting SC multiports are established assuring the elimination of the influence of stray capacitance.

M. NAŁĘCZ, J. MULAWKA

MÉTHODE MORPHOLOGIQUE DE GENERATION DES SYSTÈMES À CAPACITÉS COMMUTEES

Résumé

On a étudié les fonctions de transfert pour tous les systèmes à capacités commutées à deux phases, comprenant un amplificateur opérationnel et deux condensateurs, ayant une faible sensibilité par rapport aux capacités parasites. Le problème a été résolu par méthode morphololigique, qui permet de décrire le système, utilisant des multipôles dans le domaine de z. On a obtenu deux nouvelles intéressantes structures et, en plus, on a formulé les conditions pour les bornes d'entrée et de sortie des multipôles à capacités commutées, assurant le fonctionnement lors de l'élimination de l'effet de capacités parasites.

M. NAŁĘCZ, J. MULAWKA

BILDUNG VON SYSTEMEN MIT GESCHALTETEN KAPAZITÄTEN MITTELS MORPHOLOGISCHER METHODE

Zusammenfassung

Bei dieser Bearbeitung wurden Netzwerkfunktionen aller, den Streukapazitäten gegenüber empflindlichen Zweiphasensysteme des Typs SC untersucht, die aus einem Operationsverstärker und zwei Kondensatoren bestehen. Das Problem wurde mittels der morphologischen Methode gelöst, indem das Netzwerk mit Hilfe von N-Polen (Vielpolen) im z-Bereich beschrieben wurde. Man gewann zwei interessante Strukturen. Darüber hinaus wurden Knotenbedingungen für die Funkion der SC-Vielpole formuliert, wodurch der Einfluß der Streukapazitäten ausgeschieden wird.

М. НАЛЭНЧ, Ю. МУЛЯВКА

СИНТЕЗ ЦЕПЕЙ С ПЕРЕКЛЮЧАЕМЫМИ КОНДЕНСАТОРАМИ МОРФОЛОГИЧЕСКИМ МЕТОДОМ

Р е з ю м е

Исследованы передаточные функции напряжения всех 2-фазных цепей SC построенных из одного операционного усилителя и двух конденсаторов. Эти цепи не чувствительны к паразитным ёмкостям. Проблема решена морфологическим методом на основе многополюсников описанных в z-области. Приведены два интересные решения цепей SC с одним усилителем. Кроме того предлогаются условия работы многополюсников SC гарантирующие исключение влияния паразитных ёмкостей.

ROZPRAWY ELEKTROTECHNICZNE, 1988, 34, z. 4, ss. 1113–1128

621.313.822

A Parallel-Plate Waveguide Containing Perpendicularly Magnetized Layers of Gyromagnetic Media

MICHAŁ MROZOWSKI, JERZY MAZUR

Instytut Telekomunikacji, Politechnika Gdańska

Received 1987.04.21

Authorized 1988.03.09

The results of the theoretical and numerical analysis of a stratified, gryomagnetic medium filled, parallel-plate waveguide are presented. The biasing d.c. magnetic field is assumed to be normal to interfaces between layers. The boundary value problem is expressed in terms of a linear operator equation. The eigenvalue equation is obtained using the transfer matrix approach. Approximate formulae for the interesting case of a thin gyromagnetic film are given. The results of a numerical investigation of several wavequiding structures are discussed. The influence of the anisotropy of a ferrite medium on the dispersion characteristics and the coupling among the modes existing in the guide is thoroughly examined.

INTRODUCTION

Stratified waveguiding structures are extensively utilized in microwave devices due to their technological advantages. For instance these structures can be found in MIC constructions. There are several methods [1—6] of the full-wave analysis of these lines among which the mode matching procedure [3—6] was found to be very powerful and particularly suitable for the rigorous examination of the waveguides having a complicated crossectional geometry. This procedure involves the expansion of fields in different regions of a waveguide, using complete sets of functions fulfilling both the wave equation and the boundary conditions on the screening walls. If the structure under investigation is isotropic, these functions may be found solving an appropriate Sturm-Liouville eigenvalue problem [7]. The Sturm-Liouville problem has been extensively investigated [11], neverthelles, papers various technical problems concerning the determination of its eigenvalues are still being published [8—9]. However, for stratified gyromagnetic media filled structures with perpendicular magnetization, no ready-to-use mathematical theory of a similar character exists and, in consequence, analytical methods allowing to deal with such waveguides are less developed. For this reason a new general approach to the analysis of an arbitrary parallel-plate line with inhomogeneous gyromagnetic filling was proposed by the authors [10].

In the present paper this new approach is used to treat more throughly the case of

a layered line. New results as well as formulae enabling the straightforward application of the new technique are included.

ANALYSIS

We assume that the structure under investigation, shown in Fig. 1, is lossless and infinite both in the y- and z-directions. The line is loaded with N layers of homogeneous gyromagnetic materials magnetized perpendicularly to the bounding screens (perfect electric of magnetic walls) located in the planes x_0, $x_N = const$. Each layer is charac-

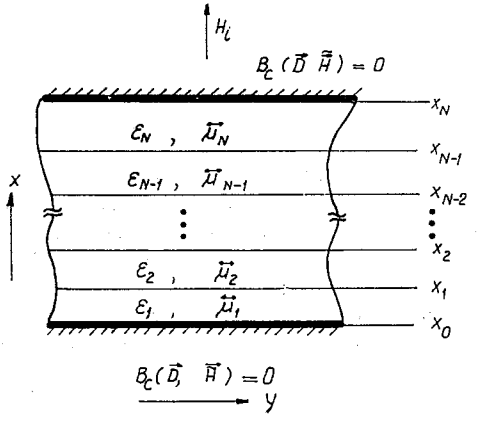


Fig. 1. Cross-section of a parellel-plate waveguide filled with layers of gyromagnetic media

terized by its scalar relative permittivity ε_i and tensor relative permeability $\overleftrightarrow{\mu_i}$. According to the definition of the structure $\overleftrightarrow{\mu_i}$ has the following form:

$$\overleftrightarrow{\mu_i} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \mu_i & -j\mu_{ai} \\ 0 & j\mu_{ai} & \mu_i \end{bmatrix}. \tag{1}$$

In the theoretical study presented in [10], it was shown that the electromagnetic wave propagation in such a structure can be alternatively expressed in terms of the following eigenvalue problem:

$$\boldsymbol{L}\underline{\Phi} - \delta\underline{\Phi} = 0 \tag{2a}$$

$$\boldsymbol{B}_c(\underline{\Phi}) = 0 \tag{2b}$$

where:

$$\underline{\Phi}(x, y, z) = \underline{\Phi}(x)e^{k_y y + k_z z}, \tag{3}$$

$$\underline{\Phi}(x) = \begin{bmatrix} \Phi_1(x) \\ \Phi_2(x) \end{bmatrix} = \begin{bmatrix} D_x(x) \\ \tilde{H}_x(x) \end{bmatrix} \quad \text{and} \quad \delta = k_y^2 + k_z^2 \tag{4}$$

with

$$\tilde{H}_x = \sqrt{\varepsilon_0 \mu_0}\ H_x.$$

In the above expressions $\boldsymbol{L}$ is a self-adjoin operator $\boldsymbol{B}_c(\underline{\Phi})$ represents boundary conditions for the electric flux density D_x and the magnetic field $\tilde{H}_x$ at the screening walls, ε_0, μ_0 stand for the premittivity and permeability of the vaccum, whereas k_y and k_z are the wavenumbers in the y- and z-directions, respectively.

The detailed expressions for the elements of operator $\boldsymbol{L}$ are given in [10]. The properties of $\boldsymbol{L}$ were thoroughly investigated in [10] and [18] by means of the spectral theory of linear operators [11—13]. In particular, it was found that the spectrum of $\boldsymbol{L}$ is real, discrete, infinite and semi bounded regardless of the combination of the boundary conditions at $x = x_0, x_N$.

In order to determine solutions satisfying both the operator equation (2a) and boundary conditions (2b) one has to find first general solutions of the wave equation in each layer. Next, the characteristic equation, determining the eigenvalues δ_i of the operator, is derived by consequent matching of the field constituents at the cross-sectional interfaces. Once the characteristic equation has been solved, the eigenfunction $\underline{\Phi}_i$ is constructed for each eigenvalue as a superposition of solutions obtained for subsequent layers.

The general solution to the wave equation in the i-th layer is given by:

$$D_x(x) = S_1 \lambda_1^{-1} \frac{\mathrm{d}}{\mathrm{dx}} f_1(x) + f_3(x)$$
$$\tilde{H}_x(x) = f_1(x) + R_3 \lambda_3^{-1} \frac{\mathrm{d}}{\mathrm{dx}} f_3(x), \tag{5}$$

where:

$$f_1(x) = A_1 \sinh[\lambda_1(x - x_{i-1})] + A_2 \cosh[\lambda_1(x - x_{i-1})]$$
$$f_3(x) = A_3 \sinh[\lambda_3(x - x_{i-1})] + A_4 \cosh[\lambda_3(x - x_{i-1})] \tag{6}$$

with

$$\lambda_1 = \sqrt{\frac{1}{2}\left[g_2 - \sqrt{g_2^2 - 4g_0}\right]}$$
$$\lambda_3 = \sqrt{\frac{1}{2}\left[g_2 + \sqrt{g_2^2 - 4g_0}\right]} \tag{7}$$

$$S_1 = -\frac{1}{k_0 \mu_{ai} \lambda_1}\left[-\lambda_1^2 + \mu_i(\delta + k_0^2 \varepsilon_i)\right]$$
$$R_3 = \frac{\mu_i}{k_0 \mu_{ai} \varepsilon_i \lambda_3}(\lambda_3^2 + \delta + k_0^2 \varepsilon_i \mu_{effi}), \tag{8–9}$$

where k_0 is the wavenumber in the free space, $\mu_{effi} = (\mu_i^2 - \mu_{ai}^2)/\mu_i$ is the effective permeability, and

$$g_2 = -\delta(1 + \mu_i) - 2k_0^2 \varepsilon_i \mu_i \quad g_0 = \mu_i(\delta + k_0^2 \varepsilon_i \mu_{effi})(\delta + k_0^2 \varepsilon_i).$$

Using the above expressions, we may formulate the characteristic equation. For this purpose a novel procedure similar to the transverse resonance method [14—15] is applied. In this procedure, a transfer matrix is defined using the new set continuity conditions for field components normal to the cross-sectional interfaces and parallel to the magnetization direction [10].

Applying the continuity conditions for all interfaces in the structure under investigation, we obtain the following definition of the global transfer matrix $\boldsymbol{T}$.

$$\boldsymbol{T} = \prod_{i=N}^{1} \boldsymbol{T}^{(i)} \tag{10}$$

where $\boldsymbol{T}^{(i)}$ is the transfer matrix for the i-th layer of gyromagnetic medium. Detailed expressions for the elements of $\boldsymbol{T}^{(i)}$ are given in Appendix A. Also, in this appendix an interesting case of a thin layer of gyromagnetic medium is considered and the transfer matrix for such a layer is derived.

The characteristic equation depends on the combination of boundary conditions at the screening walls. For instance, if the structure is bounded by electric walls located at the planes $x = x_0$ and $x = x_N$, the characteristic equation takes the form:

$$t_{13}t_{24} - t_{14}t_{23} = 0 \tag{11}$$

where t_{kl} are the elements of global transfer matrix $\boldsymbol{T}$.

NUMERICAL RESULTS

As the structures we deal with in this paper are infinite, we may assume that the electromagnetic field is homogeneous with respect either to the y- or z-direction. Accordingly we may put $k_y = 0$. Let us recall that eigenvalues are related with the wavenumbers of a parallel-plate line by the expression $\delta_i = (k_y^{(i)})^2 + (k_z^{(i)})^2$. Thus, having determined the set of eigenvalues $\{\delta_i\}$, we may investigate the electromagnetic wave propagation along the z-axis of the rectangular coordinate system.

In order to discuss the behaviour of eigenvalues of various layered structures, the FORTRAN computer program implementing the method briefly described in the prewious section was prepared. Roots of the characteristic equation were found by means of a new numerical procedure designed for finding zero locci of a real function of two variables [19]. This novel pocedure minimizes the number of function calls and facilitates the separation of curves. The computations were carried out on R-32 mainframe computer and an 8 MHz IBM-AT compatible computer with a numerical coprocessor. The computation time on both computers was very similar and depended mainly on the number of strata. Typically, one point in a three layers configuration was evaluated in less then 0.1 second.

First, a two layer ferrite-air parallel-plate line was examined. The characteristic equation, obtained using (A.1—A.16) together with (10) and (11) reads as follows:

$$S_1 R_3 \left(\frac{\lambda_d}{\varepsilon_d}\overline{C}_1\overline{S}_d + \frac{\lambda_1}{\varepsilon_f}\overline{S}_1\overline{C}_d\right)\left(\frac{\lambda_3}{\mu}\overline{S}_d\overline{C}_3 + \lambda_d\overline{C}_d\overline{S}_3\right) + S_1\frac{\overline{S}_d\overline{C}_d}{\varepsilon_f}\mu_a k_0(\lambda_1\overline{S}_1\overline{C}_3 - \lambda_3\overline{S}_3\overline{C}_1) -$$
$$\left(\lambda_d\overline{S}_1\overline{C}_d + \frac{1}{\mu}\overline{C}_1\overline{S}_d\right)\left(\frac{\lambda_3}{\varepsilon_f}\overline{S}_3\overline{C}_d + \frac{\lambda_d}{\varepsilon_d}\overline{S}_d\overline{C}_3\right) = 0 \tag{12}$$

with

$$\overline{S}_{1(3)} = \sinh(\lambda_{1(3)}h_f) \quad \overline{S}_d = \sinh(\lambda_d h_d)$$
$$\overline{C}_{1(3)} = \cosh(\lambda_{1(3)}h_f) \quad \overline{C}_d = \cosh(\lambda_d h_d),$$

where h_f, ε_f and h_d, ε_d is the thickness and relative permittivity of the ferrite and dielectric layer, respectively, while $\lambda_d = -\varepsilon_d k_0^2 - \delta$. It can be easily shown that as the biasing d.c. magnetic field H_i tends to infinity, the two first terms in (12) vanish and one gets two independent equations for the E_x and H_x modes, similar to those obtained by Emert [3]. It should be emphasized that even a thin layer of gyrotropic medium causes the coupling betwen the E_x and H_x modes. Assuming the thickness d of a ferrite film much smaller than the wavelength and using approximate formulae derived in Appendix A we get:

$$G^E G^H + G^{EH} = 0, \tag{13}$$

where

$$G^E = \lambda_d \sinh(\lambda_d h_d) + \frac{\mathrm{d}}{\varepsilon_d} \cosh(\lambda_d h_d)[\lambda_d^2 - k_0^2(\varepsilon_f \mu - 1($$

$$G^H = \frac{\sinh(\lambda_d h_d)}{\lambda_d} + \mathrm{d}\mu \cosh(\lambda_d h_d)$$

$$G^{EH} = k_0^2 \mu_a^2 \mathrm{d}^2 \cosh^2(\lambda_d h_d).$$

In the above expression, the first and the second term correspond to the perturbed E_x and H_x modes of an empty parallel-plate line, respectively, whereas component G^{EH} describes the coupling between them due to the anisotropic properties of the ferrite film.

In Figs. 2a and 2b the roots of equations (12) and (13) are compared for $f = 11$ GHz and $f = 25$ GHz. Provided the ratio d/λ_0 (λ_0 — wavelength in the free space) is smaller than 0.02, the low order eigenvalues are well approximated by eq. (13). The error increases faster for the eigenfunctions for which the H_x component is predominant. Comparing Figs. 2a and 2b one may observe that the results obtaines for $f = 25$ GHz are accurate in a greater range than for $f = 11$ GHz. This effects is due to the anisotropy of the ferrite medium defined by the ratio μ_a/μ which, in this case, is smaller for $f = 25$ GHz. Hence, the approximate formulae (A.20) are not only sensitive to the ratio d/λ_0 but also to the physical parameters of the ferrite film.

Figure 3 shows the behaviour of δ as a function of the ferrite layer thickness. At $f =$ $= 11$ GHz (Fig. 3a) where μ_a has a high value, the eigenvalues vary in a far larger range than at $f = 25$ GHz (Fig. 3b). All but the first eigenvalue form pairs and they degenerate, i.e. converge to the same value when the thickness of the ferrite layer becomes zero. It can be also observed that the starting point $((n\pi/h_d) - k_0^2)$ of each wavy line determines the number of its crests.

In order to determine the influence of a gyromagnetic layer on the frequency behaviour of the eigenvalues, a three layer structure was examined. In the subsequent figures (Fig. 4a-d) the solutions of the characteristic equation are plotted for the increasing thickness of the gyromagnetic layer while $h_1 + h_2$ remains constant. In the first diagram there is no anisotropic medium in the structure ($h_2 = 0$) and, therefore, pure E_x and H_x modes exist independently. Note, that only the fundamental mode propagates, i.e. $\delta = k_z^2 < 0$, at $f = 11$ GHz. As the thickness of the bottom dielectric layer decreases, the cut-off frequencies of the modes become lower. At first, the gyromagnetic layer affects mainly the higher order modes. If the layer is thick enough (Figs. 4c and 4d) the curves approach

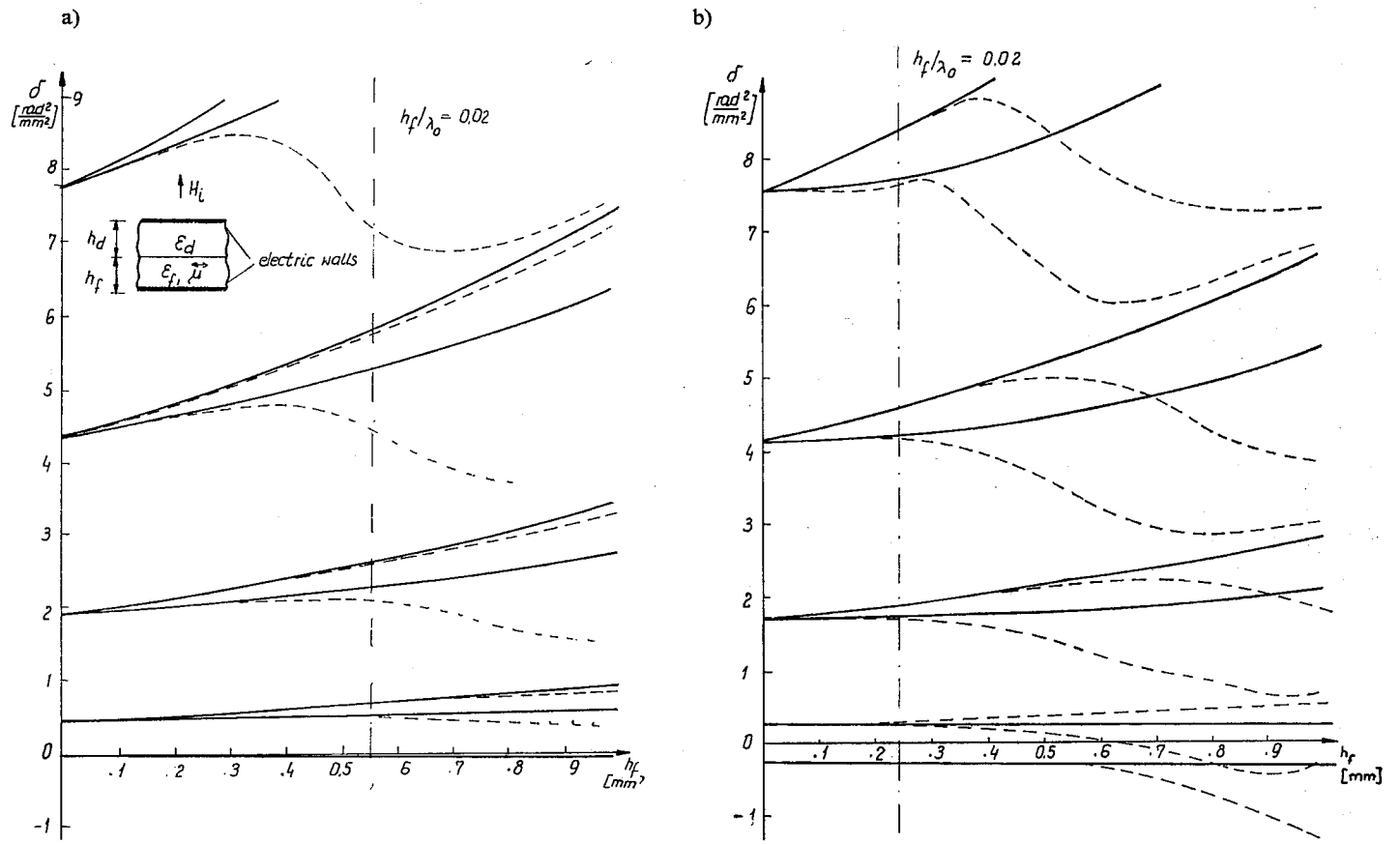


Fig. 2. Solutions of the eigenvalue equation as a function of the ferrite layer thickness (two layer ferrite-dielectric configuration: $M_s = 398.09$ kA/m, $H_i = 159.24$ kA/m, $\varepsilon_f = 13.5$, $\varepsilon_d = 1$, $h_f + h_d = 4.5$ mm) Solid line — eq. (13), dashed line — eq. (12); a) $f = 11$ GHz, $\mu = .136$, $\mu_a = 1.7$, $\mu_a/\mu = 12.46$, b) $f = 25$ GHz, $\mu = .87$, $\mu_a = 0.59$, $\mu_a/\mu = 0.674$

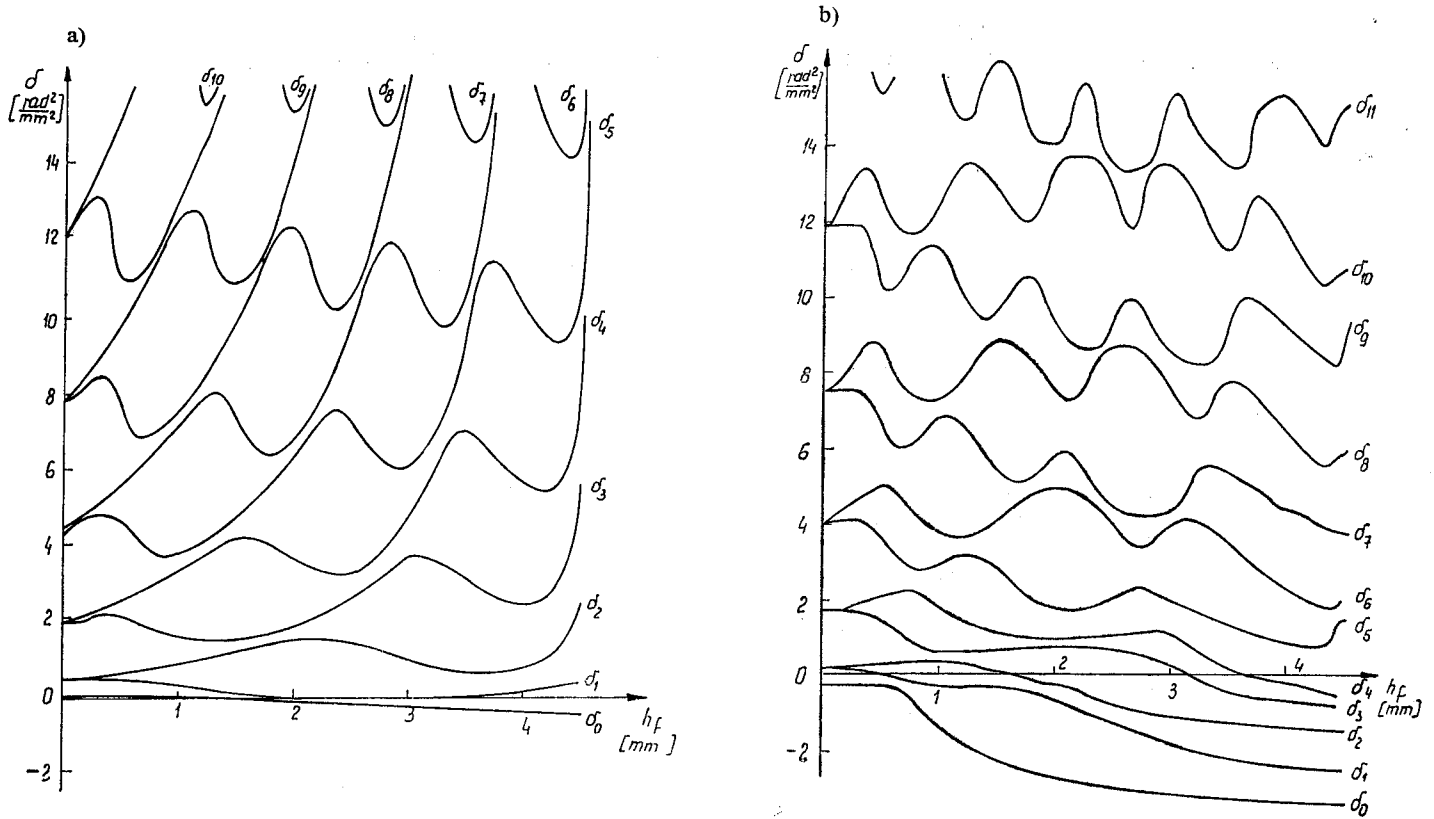


Fig. 3. Eigenvalues of the ferrite-dielectric parallel-plate line shown in Fig. 2 vs. the ferrite layer thickness ($h_f + h_d = 4.5$ mm); a) $f = 11$ GHz, $\mu = .136$, $\mu_a = 1.7$, $\mu_a/\mu = 12.46$, b) $f = 25$ GHz, $\mu' = .87$, $\mu_a = 0.59$, $\mu_a/\mu = 0.674$

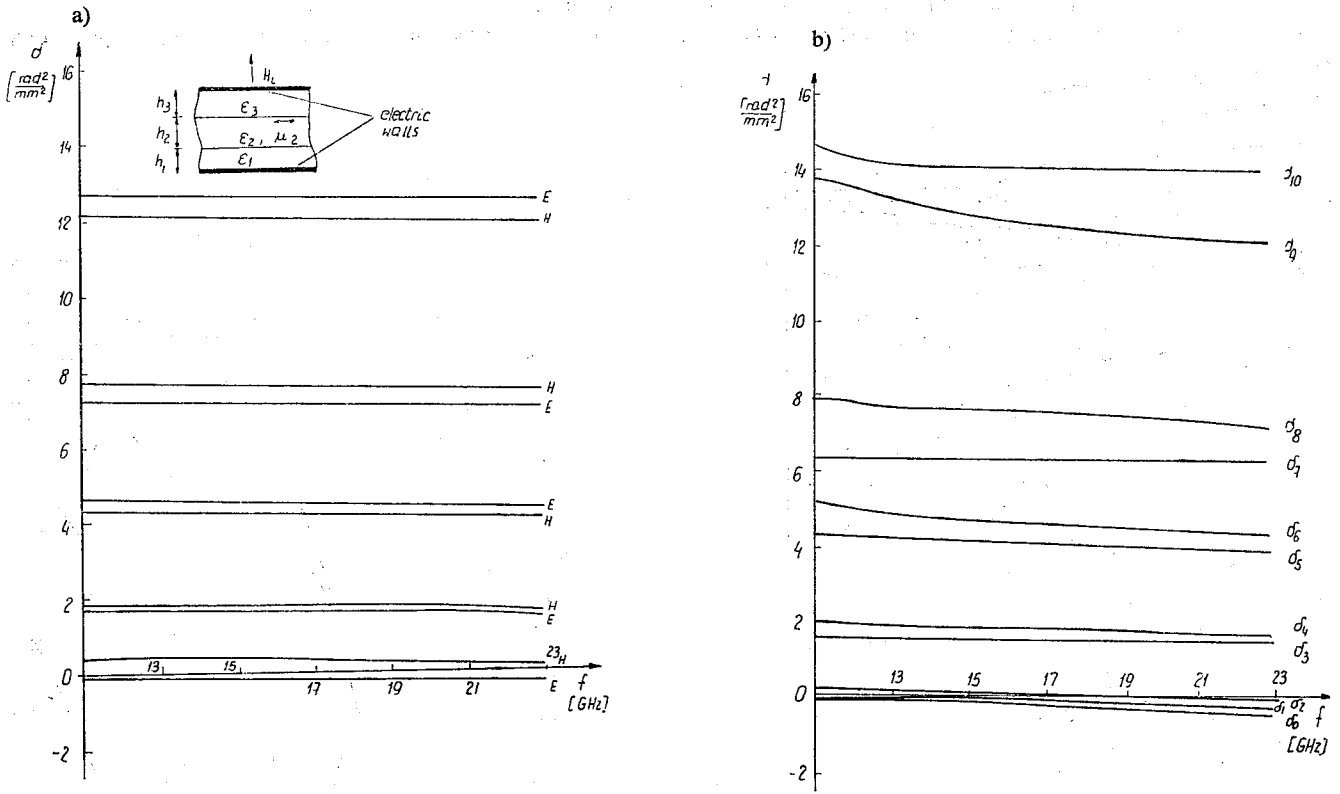


Fig. 4. Eigenvalues of a three layer parallel-plate line vs. frequency $M_s = 398.09$ kA/m, $H_i = 159.24$ kA/m, $\varepsilon_1 = 2.35$, $\varepsilon_2 = 13.5$. $\varepsilon_3 = 1$, $h_1 + h_2 = 2$ mm, $h_3 = 2.5$ mm; a) $h_1 = 2$ mm, b) $h_1 = 1.5$ mm

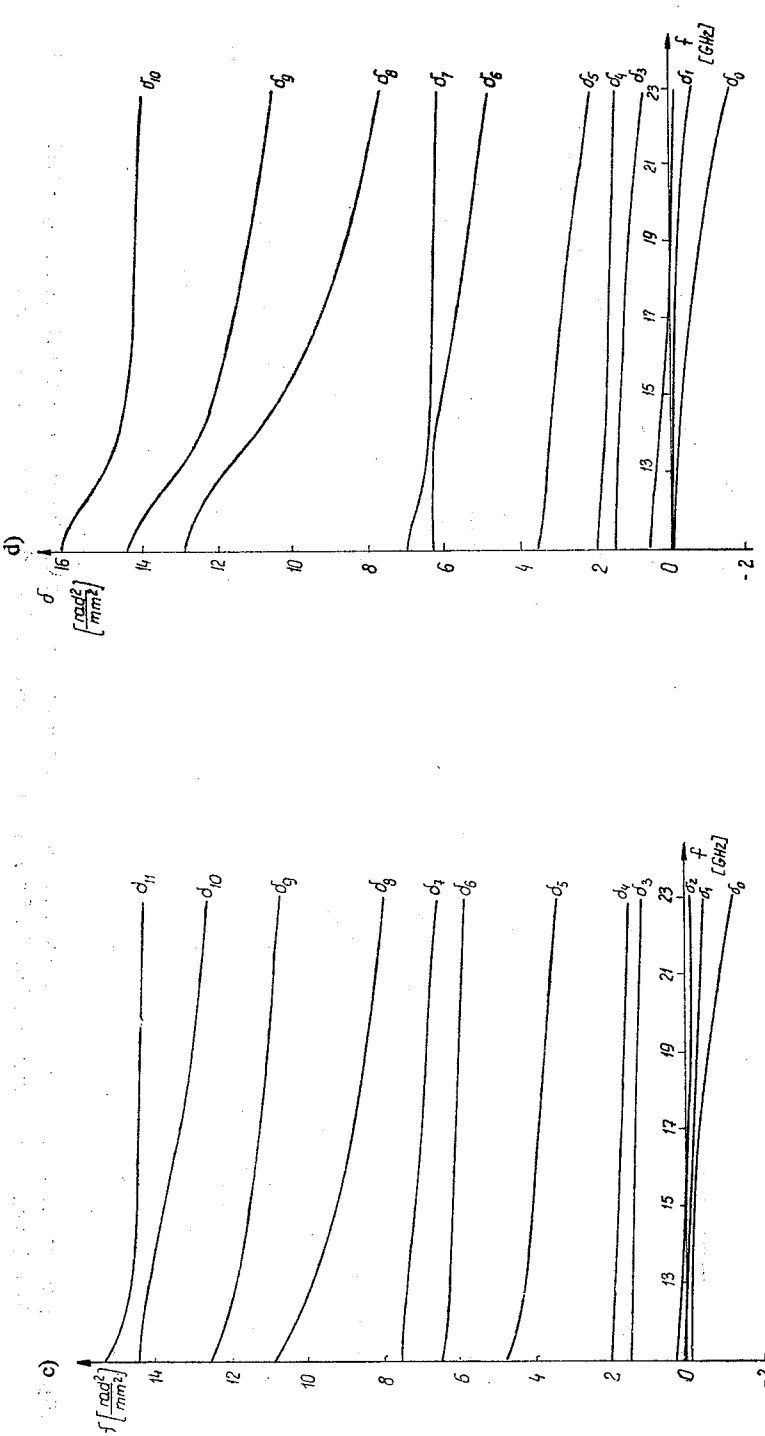


Fig. 4. Eigenvalues of a three layer parallel-plate line vs. frequency $M_s = 398.09$ kA/m, $H_i = 159.24$ kA/m, $\varepsilon_1 = 2.35$, $\varepsilon_2 = 13.5$, $\varepsilon_3 = 1$, $h_1 + h_2 = 2$ mm, $h_3 = 2.5$ mm; c) $h_1 = 1$ mm, d) $h_1 = 0.5$ mm

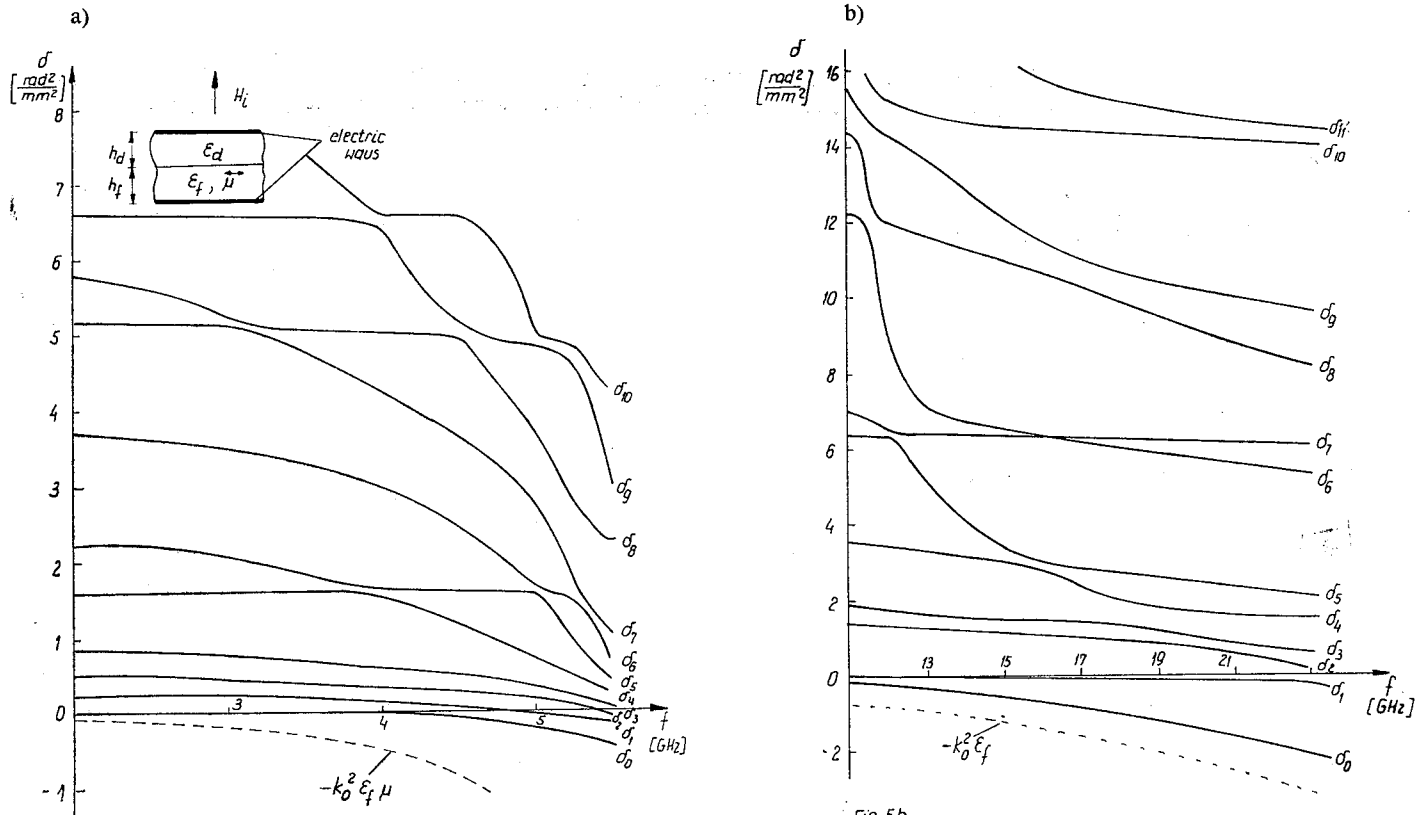


Fig. 5b

Fig. 5. Eigenvalues of a ferrite-dielectric parallel-plate line vs. frequency $M_s = 398.09$ kA/m, $H_i = 159.24$ kA/m, $\varepsilon_f = 2.35$, $\varepsilon_d = 13.5$, $h_f = 2$ mm, $h_d = 2.5$ mm a) below the ferromagnetic resonance, b) above the ferromagnetic resonance

Table 1

E_x mode purity coefficient C^i for several lowest order modes for the line shown in Fig. 5

μ, μ_a	f[GHz]	C^0	C^1	C^2	C^3	C^4	C^5	C^6	C^7	C^8
	3.0	.99	.07	.92	.01	.99	.00	.00	.88	.12
	3.2	.99	.10	.89	.01	.99	.00	.00	.20	.81
	3.4	.99	.15	.85	.01	.99	.00	.00	.03	.96
$\mu > 1$;	3.6	.98	.20	.80	.01	.99	.00	.00	.01	.98
$\mu > 0$	3.8	.96	.26	.75	.02	.96	.02	.00	.00	.98
	4.0	.94	.33	.70	.02	.01	.98	.00	.00	.98
	4.2	.90	.42	.65	.02	.01	.99	.00	.00	.98
	4.4	.83	.52	.62	.02	.01	.99	.00	.00	.96
$\mu < 0$										
	11.0	.16	.96	.76	.32	.52	.98	.12	.68	.77
	11.6	.15	.98	.68	.40	.48	.98	.19	.54	.83
	12.2	.15	.98	.60	.49	.46	.69	.68	.56	.60
	12.8	.15	.99	.51	.58	.44	.64	.95	.31	.53
	13.4	.15	.99	.44	.66	.42	.72	97	.21	.47
	14.0	.16	.99	.37	.73	.42	.73	.97	.19	.42
	14.6	.16	.99	.33	.78	.44	.70	.97	.19	.37
$\mu_a > 0$;	15.2	.17	.99	.29	.83	.50	.62	.98	.19	.32
$\mu < 1$	15.8	.18	.99	.27	.85	.69	.43	.97	.22	.28
	16.4	.18	.99	.25	.88	.76	.33	.32	.86	.27
	17.0	.19	.99	.23	.89	.77	.31	.24	.95	.29
	17.6	.19	.99	.23	.90	.76	.30	.23	.95	.31
	18.2	.20	.99	.22	.90	.76	.30	.23	.96	.34
	18.8	.21	.98	.23	.89	.77	.29	.24	.95	.37
	19.4	.21	.98	.24	.87	.79	.29	.24	.95	.38
	20.0	.22	.98	.26	.82	.81	.29	.25	.95	.39
	20.6	.22	.98	.33	.75	.82	.28	.25	.95	.40
	21.2	.23	.97	.46	.60	.84	.28	.26	.94	.41
	21.8	.23	.97	.62	.44	.85	.28	.26	.94	.41

one another which means that the gyromagnetic medium causes the cross-coupling among different modes of the line.

In order to make a quantitative discussion of the coupling phenomenon, we introduce the „E_x — mode purity coefficient" C^i which we define in the following way:

$$C^i = \frac{\int_{x_0}^{xN} \varepsilon^{-1}(x) D_x^i(x) D_x^i(x) \mathrm{d}x}{\int_{x_0}^{xN} [\varepsilon^{-1}(x) D_x^i(x) D_x^i(x) + \tilde{H}_x^i(x) \tilde{H}_x^i(x)] \mathrm{d}x}. \tag{14}$$

Since in the parallel-plate line filled with gyromagnetic media each eigenfunction is a vector with two components, namely E_x and $\tilde{H}_x$ [10], the value of C^i varies from zero to one.

The E_x — mode purity coefficients for several low order modes of the ferrite-air waveguide presented in Fig. 5 are shown in Table 1. Note, that below the ferromagnetic reso-

nance ($\mu > 1$, $\mu_a > 0$) almost pure modes exist. As the frequency increases, the H_x modes approach the E_x modes yielding a passive coupling. Additionally, we may observe that the only propagating mode is similar to the fundamental pure E_x mode of an empty parallel-plate line. This is not true for the lowest order mode above the ferromagnetic resonance in the frequency region in which $0 < \mu < 1$, $\mu_a < 0$. From the table it is seen that for this wave, the magnetic field x-component is predominant which implies that it is rather the HE_x mode. Finally, the figures in the table also illustrate a phenomenon discussed earlier, namely the coupling between the components of the vector eigenfunction. Above the ferromagnetic resonance only few modes may be regarded as „pure". The majority of them are the hybrid ones. For this reason it is difficult to determine precisely the frequency band in which a coupling between given modes takes place.

$$T_{34}^i = \frac{S_1}{M_2}\left[-\cosh(\lambda_1 h_i)+\cosh(\lambda_3 h_i)\right]$$

$$T_{44}^i = \frac{1}{M_2}\left[-G_1\cosh(\lambda_1 h_i)+G_3 S_1\cosh(\lambda_3 h_i)\right], \qquad \text{(A1–A16)}$$

where:

$$M_1 = R_3 S_1\,\lambda_1 - \lambda_3, \qquad M_2 = S_1 G3 - G1$$

with:

$$G_1 = \frac{\mu_{ai} k_0 S_1 + \lambda_1}{\mu_i}$$

$$G_3 = \frac{\mu_{ai} k_0 + R_3 \lambda_3}{\mu_i}$$

CONCLUSIONS

A detailed discussion of the electromagnetic wave propagation in a parallel-plate line loaded with perpendicularly magnetized layers of gyromagnetic media is presented. The boundary value problem is expressed in term of a self-adjoint operator equation. The transverse resonance approach is used to determine the eigenvalue equation. The expressions for the elements of the transfer matrix are given. Several stratified structures are numerically examined. The influence of a gyromagnetic layer on the mode coupling phenomenon is investigated. The results are applicable to the full-wave mode-matching analysis of shielded waveguides inhomogeneously filled with gyromagnetic media [18].

ACKNOWLEDGMENT

This work was supported by the Polish Academy of Sciences under contract CPBP-02.02.3.2.

APPENDIX A

THE TRANSFER MATRIX OF A THICK GYROMAGNETIC LAYER

Transfer matrix $T^{(i)}$ links the continuity vector F_i between the extreme height $x = x_i^-$ and $x = x_{i-1}^+$ of the i-th layer

$$F_{i|x=x_i^-} = T^{(i)} F_{i|x=x_{i-1}^+}$$

where:

$$F_i(x) = \begin{bmatrix} 0 & \frac{1}{\varepsilon_i}\frac{\mathrm{d}}{\mathrm{dx}} \\ 1 & 0 \\ 0 & 1 \\ \frac{1}{\mu_i}\frac{\mathrm{d}}{\mathrm{dx}} & k_0\frac{\mu_{ai}}{\mu_i} \end{bmatrix} \begin{bmatrix} D_x^i \\ \tilde{H}_x^i \end{bmatrix}$$

and the elements of transfer matrix $T^{(i)}$ are defined according to the following expressions

$$T_{11}^i = \frac{1}{M_1}[\lambda_1 S_1 R_3 \cosh(\lambda_1 h_i) - \lambda_3 \cosh(\lambda_3 h_i)]$$

$$T_{21}^i = \frac{\varepsilon_i R_3}{M_1}[\cosh(\lambda_1 h_i) - \cosh(\lambda_3 h_i)]$$

$$T_{31}^i = \frac{\varepsilon_i}{M_1}[S_1 R_3 \sinh(\lambda_1 h_i) - \sinh(\lambda_3 h_i)]$$

$$T_{41}^i = \frac{\varepsilon_i}{M_1}[G_1 R_3 \sinh(\lambda_1 h_i) - G_3 \sinh(\lambda_3 h_i)]$$

$$T_{12}^i = \frac{S_1 \lambda_1 \lambda_3}{\varepsilon_i M_1}[-\cosh(\lambda_1 h_i) + \cosh(\lambda_3 h_i)]$$

$$T_{22}^i = \frac{1}{M_1}[-\lambda_3 \cosh(\lambda_1 h_i) + R_3 S_1 \lambda_1 \cosh(\lambda_3 h_i)]$$

$$T_{32}^i = \frac{S_1}{M_1}[-\lambda_3 \sinh(\lambda_1 h_i) + \lambda_1 \sinh(\lambda_3 h_i)]$$

$$T_{42}^i = \frac{1}{M_1}[-\lambda_3 G_1 \sinh(\lambda_1 h_i) + G_3 S_1 \lambda_1 \sinh(\lambda_3 h_i)]$$

$$T_{13}^i = \frac{1}{\varepsilon_i M_2}[\lambda_1 S_1 G_3 \sinh(\lambda_1 h_i) - \lambda_3 G_1 \sinh(\lambda_3 h_i)]$$

$$T_{23}^i = \frac{1}{M_2}[G_3 \sinh(\lambda_1 h_i) - G_1 R_3 \sinh(\lambda_3 h_i)]$$

$$T_{33}^i = \frac{1}{M_2}[S_1 G_3 \cosh(\lambda_1 h_i) - G_1 \cosh(\lambda_3 h_i)]$$

$$T_{43}^i = \frac{G_1 G_3}{M_2}[\cosh(\lambda_1 h_i) - \cosh(\lambda_3 h_i)]$$

$$T_{14}^{i} = \frac{S_1}{\varepsilon_i M_2}\left[-\lambda_1 \sinh(\lambda_1 h_i) + \lambda_3 \cosh(\lambda_3 h_i)\right]$$

$$T_{24}^{i} = \frac{1}{M_2}\left[-\sinh(\lambda_1 h_i) + R_3 S_1 \sinh(\lambda_3 h_i)\right]$$

THE TRANSFER MATRIX OF A THIN GYROMAGNETIC LAYER

Presently, let us derive a transfer matrix for the case of a thin film of a gyromagnetic medium inserted between two layers, say i and $i+1$, of a parallel-plate line (Fig. A.1). We assume that the gyromagnetic layer may be considered as a thin one if its thickness is much smaller than the wavelength. In this case, the functions expressing field components continuous across the interface and their first derivatives with respect to the x-direction may be computed inside such a layer according to the approximate formulae:

$$F_f = \frac{F_{i+1}^{+} + F_i^{-}}{2}, \qquad \frac{\mathrm{d}}{\mathrm{d}x} F_f = \frac{F_{i+1}^{+} - F_i^{-}}{d}, \tag{A.17}$$

where d is the thickness of a gyromagnetic film, F_f represents field components (D_x, $\tilde{H}_x$, E_z, $\tilde{H}_z$, E_y, $\tilde{H}_y$) inside the film and F_{i+1}^{+} and F_i^{-} are the field components at the upper x_i^{+} interface in layer $i+1$ and at the lower x_i^{-} interface in layer i, respectively.

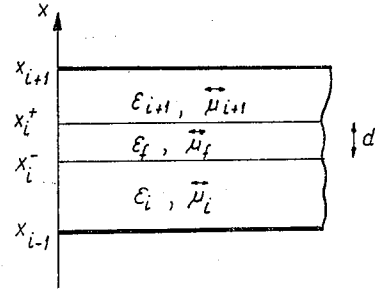


Fig. A.1. A gyromagnetic film inserted between two layers of gyromagnetic media

Next, from Maxwell's equations one obtains, under assumption (A.17), formulae expressing the transfer of the tangential field components through the layer of a gyromagnetic medium [17].

$$2\varepsilon_0(E_z^{+} - E_z^{-}) = \frac{\mathrm{d}k_z}{\varepsilon_f}(D_x^{+} + D_x^{-}) + jk_0 d[\mu_f(\tilde{H}_y^{+} + \tilde{H}_y^{-}) - j\mu_{af}(\tilde{H}_z^{+} + \tilde{H}_z^{-})]$$

$$2\varepsilon_0(E_y^{+} - E_y^{-}) = \frac{\mathrm{d}k_y}{\varepsilon_f}(D_x^{+} + D_x^{-}) - jk_0 d[\mu_f(\tilde{H}_z^{+} + \tilde{H}_z^{-}) + j\mu_{af}(\tilde{H}_y^{+} + \tilde{H}_y^{-})]$$

$$2(\tilde{H}_z^{+} - \tilde{H}_z^{-}) = \mathrm{d}k_z(\tilde{H}_x^{+} + \tilde{H}_x^{-}) - jk_0 d\varepsilon_0 \varepsilon_f(E_y^{+} + E_y^{-})$$

$$2(\tilde{H}_y^{+} - \tilde{H}_y^{-}) = \mathrm{d}k_y(\tilde{H}_x^{+} + \tilde{H}_x^{-}) + jk_0 d\varepsilon_0 \varepsilon_f(E_z^{+} + E_z^{-}), \tag{A.18}$$

where ε_f, μ_f, μ_{af} are the quantities describing physical parameters of the film. Equations

(A.18) may be transformed to the form in which only the normal field components are involved, yielding:

$$\boldsymbol{F}_{i+1|x=x_i^+} = (\boldsymbol{T}_f^{+-1}\boldsymbol{T}_f^-)\boldsymbol{F}_{i|x=x_i^-} \tag{A.19}$$

where:

$$\boldsymbol{T}_f^{\pm} = \begin{bmatrix} 1 & 0 & \pm\frac{d}{2\varepsilon_f}(\delta + k_0^2\varepsilon_f\mu_f) & \mp\frac{d\mu_{af}k_0}{2} \\ 0 & 1 & \pm\frac{d\mu_{af}k_0}{2} & \mp\frac{d\mu_f}{2} \\ \pm\frac{d\varepsilon_f}{2} & 0 & 1 & 0 \\ 0 & \pm\frac{d}{2}(\delta + k_0^2\varepsilon_f) & 0 & 1 \end{bmatrix}. \tag{A.20}$$

Expression $\boldsymbol{T}_f^{+-1}\boldsymbol{T}_f^-$ defines the transfer matrix for a thin layer of a gyromagnetic medium.

REFERENCES

1. E. Yamashita and K. Atsuki: *Analysis of microstrip-like transmission lines by nonuniform discretization of integral equation.* Transact. IEEE MTT-24, 1976, pp. 195—200
2. J. B. Davies and P. Mirshekar-Syakal: *Spectral domain solution of arbitrary coupled transmission lines with multilayer substrate.* Transact. IEEE MTT-25, 1977, pp. 143—146
3. H. Emert: *Guided modes and radiation characteristics of covered microstrip lines.* AEÜ 30, 1976, pp. 65—70
4. R. Mittra, Y. L. Hou, and V. Jamenjad: *Analysis of open dielectric waveguides using mode-matching technique and variational methods.* Transact. IEEE MTT-28, 1980, pp. 36—43
5. S. T. Peng and A. A. Oliner: *Guidance and leakage properties of a class of open dielectric waveguides: Part I — mathematical formulations.* Transact. IEEE MTT — 29, 1981, pp. 843—855
6. U. Crombach: *Analysis of single and coupled rectangular dielectric waveguides.* Transact. IEEE MTT-29, 1981, pp. 870—874
7. R. E. Collin: *Field theory of guided waves.* New York, McGraw-Hill, 1960
8. V. Bilik: *One root intervals of dispersion relations for a resonator with dielectric sample.* Transact. IEEE MTT-32, 1984, pp. 197—198
9. H. Brandiss, U. Crombach and R. Gesche: *Bounds for the eigenvalues of a multilayered parallel-plate line.* AEÜ 37, 1983, pp. 113—116
10. M. Mrozowski and J. Mazur: *General analysis of a parallel-plate waveguide inhomogeneously filled with gyromagnetic media.* Transact. IEEE MTT-34, 1986, pp. 388—395
11. R. Courant and D. Hilbert: *Methods of mathematical physics.* vol. I. New York: Interscience Publishers, 1953
12. B. Friedman: *Principles and techniques of applied mathematics.* New York: John Wiley and Sons Inc., 1956
13. W. Kołodziej: *Selected topics on mathematical analysis.* Warsaw: PWN, 1982, (in Polish)
14. F. E. Gardiol: *Anisotropic slabs in rectangular waveguides.* Transact. IEEE MTT-18, 1970, pp. 461—467
15. H. Seidel: *Ferrite slabs in transverse electric mode wave guide.* J. Appl. Phys. 28, 1957, pp. 218—226
16. E. F. Kurushin and E. I. Nefiodov: *Electrodynamics of anisotropic waveguiding structures.* Moscow: Nauka, 1983 (in Russian)
17. M. Mrozowski and J. Mazur: *Full-wave analysis of a rectangular waveguide containing layers of ferrite media.* Proc. URSI Int. Symp., Budapest, Ang. 1986, pp. 483—485

18. M. Mrozowski and J Mazur: *The lower bound for the eigenvalues of the characteristic equation for an arbitrary multilayered gyromagnetic structure with perpendicular magnetization.* Submitted for publication in Transact. IEEE MTT
19. M. Mrozowski: *An efficient algorithm for finding zeros of a real function of two variables.* Transact. (IEEE MTT., March. 1988. (in print)

M. MROZOWSKI, J. MAZUR

LINIA PŁASKORÓWNOLEGŁA ZAWIERAJĄCA WARSTWY OŚRODKÓW ŻYROMAGNETYCZNYCH MAGNESOWANYCH PROSTOPADLE

Streszczenie

W pracy przedstawiono wyniki analizy teoretycznej linii płaskorównoległej zawierającej warstwy ośródków żyromagnetycznych. Przyjęto, iż kierunek pola polaryzującego ośrodek był prostopadły do warstw. Problem brzegowy sprowadzono do postaci liniowego równania operatorowego. Równanie charakterystyczne otrzymano stosując metodę macierzy transmisji. Podano przybliżone zależności dla interesującego przypadku linii zawierającej warstwę materiału żyromagnetycznego. Omówiono wyniki badań numerycznych dla kilku przykładowych struktur linii.

M. MROZOWSKI, J. MAZUR

LIGNE À BANDES PARALLÈLES CHARGÉE DE COUCHES GYROMAGNÉTIQUES MAGNÉTISÉES PERPENDICULAIREMENT AUX INTERFACES

Résumé

On a présenté les résultats de l'analyse théorique d'une ligne à bandes parallèles chargés des couches gyromagnétiques magnétisée perpendiculairement aux interfaces. Le problème est reduit à une équation opératoire linéaire. L'équation caractéristique est obtenue par méthode des matrices de transfert. On a dérivé les formules approximées pour les cas d'une couche mince. On a décrit les résultats de recherches numériques pour quelques lignes exemplaires.

M. MROZOWSKI, J. MAZUR

BANDLEITUNG MIT SENKRECHT VORMAGNETISIERTER GYROMAGNETISCHER FÜLLUNG

Zusammenfassung

Die Bandleitung mit geschichteter gyromagnetischer Füllung (Vormagnetisierung senkrecht den Schichten gegenüber) wird theoretisch untersucht. Das Problem wird mittels einer linearen Operatorgleichung ausgedrückt. Die Eigenwertgleichung wird mit Hilfe der Kettenmatrix gebildet.

Annäherungsformeln werden für einen interessanten Fall einer dünnen Ferritschicht hergeleitet. Die Ergebnisse der numerischen Untersuchungen für einige Beispiele linearer Struktur werden diskutiert.

М. Мрозовски, Я. Мазур

ПЛОСКИЙ ВОЛНОВОД С ГИРОМАГНИТНЫМИ СЛОЯМИ

Резюме

Представлен метод анализа плоского волновода с гиромагнитными слоями и пленками. Поставленная задача решена методом операторных уравнений. На основе полученных численным методом дисперсионных характеристик линии обсуждены их полевые свойства.

ROZPRAWY ELEKTROTECHNICZNE, 1988, 34, z. 4, ss. 1129–1147

621.372.632

Wpływ metod sterowania bezpośrednim przemiennikiem częstotliwości na pobór mocy biernej

TADEUSZ CITKO, ZOFIA DASZUTA, ANTONI BOGDAN

Katedra Elektrotechniki Przemysłowej, Politechnika Białostocka

Otrzymano 1987.12.16

Autoryzowano do druku 1988.01.08

W artykule zaprezentowano zmiany mocy biernej pobieranej przez bezpośredni przemiennik częstotliwości przy dwóch alternatywnych metodach regulacji napięcia wyjściowego: metodzie modulacji amplitudowo zależnej i metodzie regulacji szerokości impulsów. Prezentowane rezultaty uzyskano na drodze symulacji cyfrowej przebiegów oraz obliczeń komputerowych. Uzyskane wyniki potwierdziły możliwość (oszacowały poziom) minimalizacji poboru mocy biernej przy metodzie modulacji amplitudowo zależnej drogą sterowania kątem zmiany kształtu fali napięciowej.

1. WSTĘP

Głównym parametrem użytkowym każdego przemiennika częstotliwości jest wartość skuteczna I_0 wymaganej składowej harmonicznej prądu na wyjściu. O tym na ile od tej wartości odbiega wartość skuteczna I_{wy} całkowitego prądu na wyjściu świadczy prądowy współczynnik zniekształcenia $\mu_0 = \frac{I_0}{I_{wy}}$. Jeżeli impedancja odbiornika określona będzie jako stosunek wymaganych harmonicznych napięcia i prądu wyjściowego, to moc czynną pobieraną z trójfazowego przemiennika można określić następująco:

$$P = 3I_{wy}^2 \frac{U_0}{I_0} \cos\varphi_0, \tag{1}$$

gdzie φ_0 — kąt przesunięcia pomiędzy wymaganymi harmonicznymi napięcia i prądu wyjściowego.

W przypadku mostkowej konfiguracji przekształtnika sześciopulsowego wartość skuteczną napięcia U_0 można przedstawić w postaci:

$$U_0 = \frac{3}{\pi} \frac{U_m}{\sqrt{2}} \cdot r, \tag{2}$$

gdzie U_m — maksymalna wartość międzyfazowego napięcia zasilającego przemiennik,

r — względna wartość skuteczna wymaganej składowej napięcia wyjściowego odniesiona do maksymalnej wartości skutecznej tej składowej, $r \in \langle 0,1 \rangle$.
Po uwzględnieniu zależności (2):

$$\frac{P}{\frac{9}{\sqrt{2}\pi} U_m I_0} = \frac{r}{\mu_0^2} \cos\varphi_0. \tag{3}$$

Z założenia, że przemiennik jest urządzeniem bezstratnym wynika, że moc czynna pobierana z sieci zasilającej nie zależy od rodzaju przemiennika ani sposobu jego sterowania i jest jednoznacznie określona przez parametry prawej strony równania (3).

W przypadku mocy biernej powyższa prawidłowość nie obowiązuje. Moc bierna wejściowa związana z przesunięciem podstawowej harmonicznej prądu wejściowego względem napięcia sieci zasilającej jest ściśle związana ze sposobem sterowania przemiennika. W powszechnie znanym przypadku prostownika sterowanego przesunięcie to jest równe stałemu kątowi wysterowania α. W bezpośrednim przemienniku częstotliwości z sinusoidalnym średnim napięciem wyjściowym (sinusoidal mean output voltage) kąt wysterowania przekształtnika jest liniową funkcją czasu. Czas, w którym impedancje obciążenia są przyłączone do określonych faz sieci zasilającej, jest w tym przypadku stały i równy t_p [1, 2]. Zmiany sposobu dołączenia odbiornika do odpowiednich faz napięcia zasilającego poprzez wysterowanie kolejnych zaworów odbywają się po upływie kolejnych przedziałów p czasu t_p. Związek czasu t_p z pulsacją wyjściową ω_0 jest następujący [1]:

$$\omega_0 = \frac{2\pi}{qt_p} - \omega, \tag{4}$$

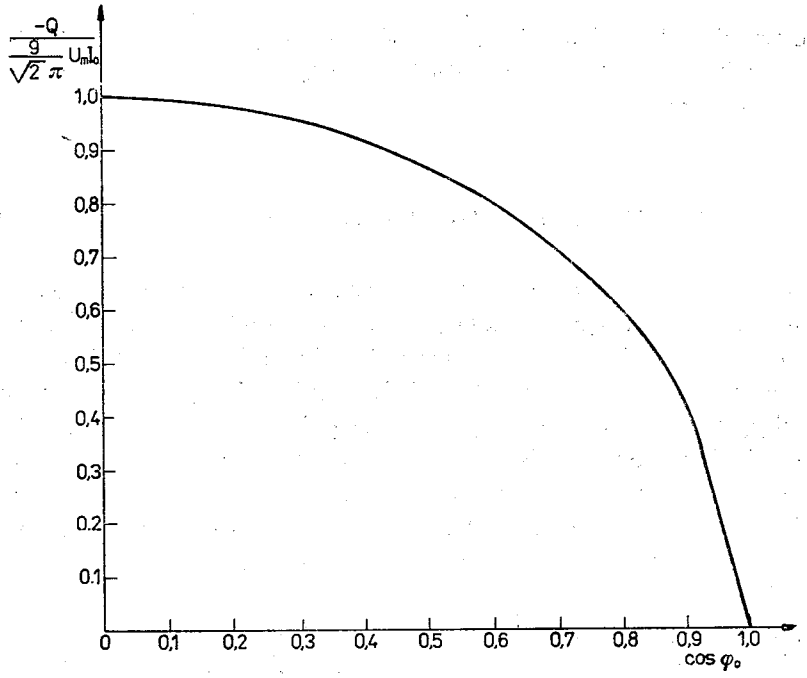
gdzie q — pulsacyjność przekształtnika (w rozważanym przypadku $q = 6$). ω — pulsacja napięcia zasilającego.
Zmiana czasu t_p powoduje zmianę pulsacji wyjściowej ω_0, przy niezmienionej amplitudzie wymaganej składowej napięcia wyjściowego, równej maksymalnej wartości $\sqrt{2}U_0 = \frac{3U_m}{\pi}$.

Można wykazać teoretycznie [1, 2], a świadczą o tym również wyniki symulacji cyfrowej przedstawione na rys. 1, że dla powyższego sterowania moc bierna pobierana przez przemiennik z sieci jest równa mocy biernej odbiornika i ma charakter przeciwny. To znaczy: dla odbiornika pobierającego moc bierną indukcyjną, moc bierna pobierana z sieci jest pojemnościowa i odwrotnie. Wartość mocy biernej można określić z zależności:

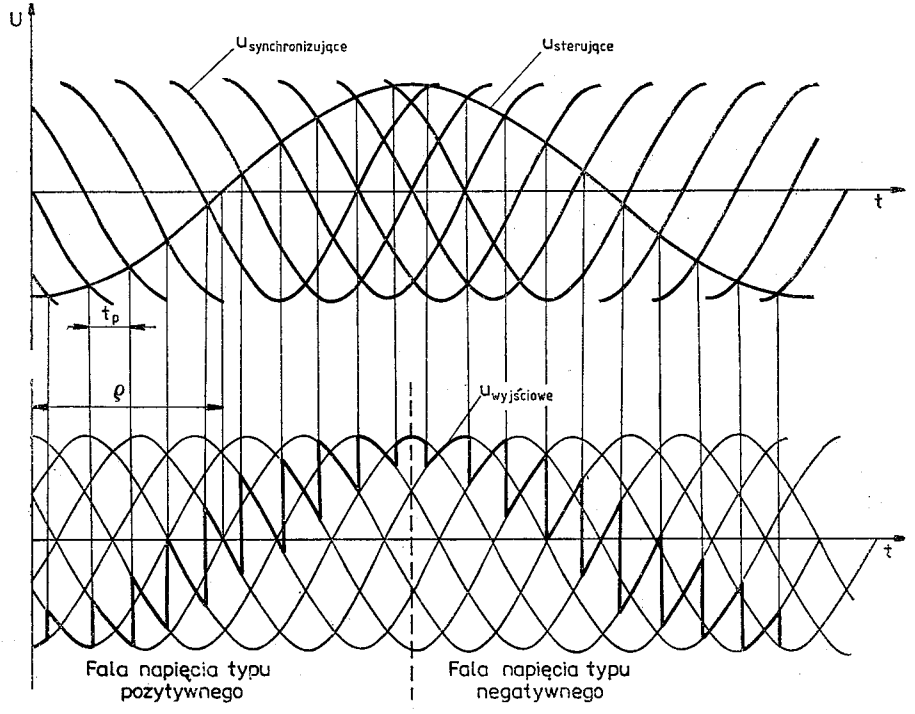
$$Q = \frac{9}{\sqrt{2}\pi} U_m I_0' \sin\varphi_0 \tag{5}$$

W praktyce rzadko jest stosowany przemiennik, w którym zmienia się tylko częstotliwość wymaganej harmonicznej napięcia. Najczęściej wymagana jest także niezależna regulacja amplitudy napięcia wyjściowego. Praktycznie istnieją dwie możliwości regulacji amplitudy wymaganej składowej napięcia wyjściowego bez zmiany jej częstotliwości przez:

1 — naruszenie liniowej zmiany kąta wysterowania (stałości czasu przewodzenia, np. przy metodzie modulacji amplitudowo zależnej)



Rys. 1. Zależność mocy biernej pobieranej przez przemiennik od współczynnika $\cos\varphi_0$ odbiornika przy sterowaniu t_p = const. zapewniającym $f_0 = 5$ Hz



Rys. 2. Zasada kształtowania napięcia wyjściowego przemiennika przy liniowej zmianie kąta wysterowania (t_p = const.)

2 — odłączenie odbiornika od sieci i zwieranie ich, w celu zapewnienia ciągłości prądu, na czas $(1-\gamma)t_p$ w każdym przedziale czasu $t_p(\gamma \in \langle 0,1\rangle$, przy metodzie modulacji szerokości impulsów).

7. REGULACJA NAPIĘCIA METODĄ MODULACJI AMPLITUDOWO ZALEŻNEJ

Praktycznie liniową zmianą kąta wysterowania można uzyskać poprzez porównanie sinusoidalnej funkcji sterującej o pulsacji ω_0 z cosinusoidalną funkcją synchronizującą o pulsacji napięcia zasilającego i amplitudzie równej amplitudzie funkcji sterującej w taki sposób, że w przedziałach, w których funkcja sterująca jest rosnąca, następuje porównanie z malejącym zboczem funkcji synchronizującej i odwrotnie — rys. 2. Takie sterowanie zapewnia falę napięcia wyjściowego typu pozytywnego (positive output vaveform) dla rosnącego napięcia sterującego oraz falę typu negatywnego (negative output vaveform) przy malejącym napięciu sterującym. Trzymając się powyższej zasady, zmiany amplitudy wymaganej harmonicznej na wyjściu można uzyskiwać poprzez modulację amplitudowo zależną czyli przez zmianę amplitudy napięcia sterującego.

Zmniejszenie amplitudy napięcia sterującego prowadzi do wzrostu poboru mocy biernej praktycznie niezależnie od charakteru obciążenia. Względna wartość mocy biernej Q odniesionej do biernej mocy obciążenia przy maksymalnej amplitudzie wymaganej składowej napięcia wyjściowego przedstawiona jest na rys. 3 zawierającym wykresy:

$$Q_w = \frac{Q}{\frac{9}{\sqrt{2}\pi} U_m I_0 \sin\varphi_0} = f(\cos\varphi_0) \quad \text{przy } r = const. \tag{6}$$

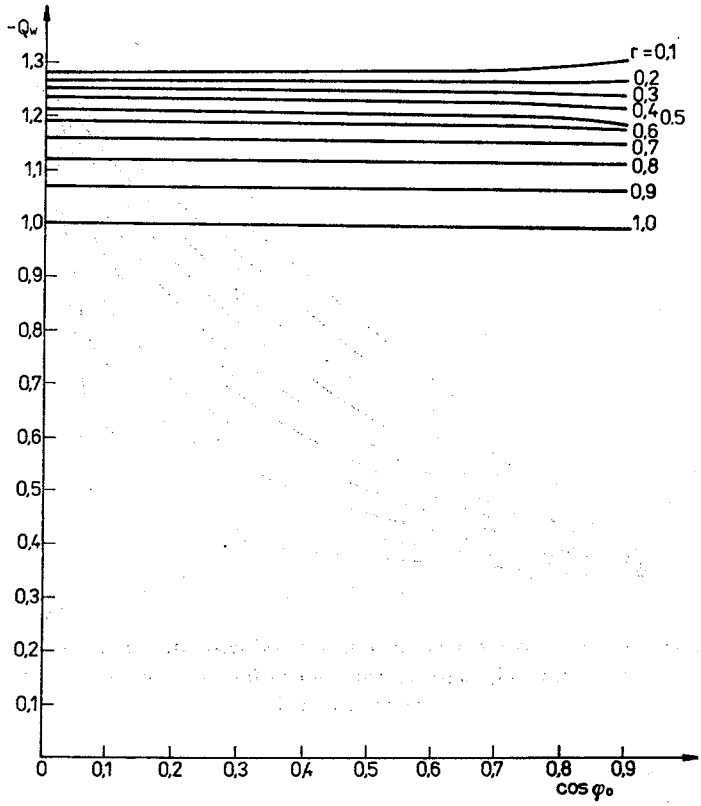
Wykresy te są wynikiem symulacji cyfrowej i podane są dla $\cos\varphi_0 \neq 1$, gdyż dla $\cos\varphi_0 = 1$ mianownik wyrażenia (6) jest równy zero.

Biorąc pod uwagę fakt, że zgodnie z (3), wraz ze zmniejszaniem się względnej amplitudy napięcia r zmniejsza się moc czynna pobierana z sieci, uświadomić sobie można, jak silnie zmienia się w tych warunkach wejściowy współczynnik przesunięcia:

$$\cos\varphi_{we} = \frac{P}{\sqrt{P^2+Q^2}} \tag{7}$$

Ilościowe zmiany tego współczynnika w funkcji współczynnika obciążenia $\cos\varphi_0$, przy różnych wartościach r zaprezentowane są na rys. 4. Prezentowane wykresy dotyczą obciążenia o charakterze rezystancyjno-indukcyjnym ($\varphi_0 \geqslant 0$). Moc bierna pobierana przez przemiennik ma charakter pojemnościowy.

Istotne jest to, że omawiany sposób regulacji amplitudy wymaganej składowej napięcia wyjściowego różni się od powszechnej metody sterowania klasycznego cyklokonwertora tylko chwilą zmiany typu fali napięciowej, określoną względem zerowej fazy początkowej napięcia sterującego. W cyklokonwertorze chwila zmiany typu fali napięciowej jest określona chwilą zmiany kierunku prądu. Jeżeli więc chwila zmiany typu fali wyjściowej w cyklokonwertorze określona jest przez kąt $\varrho = \varphi_0$, to w omawianym przy-

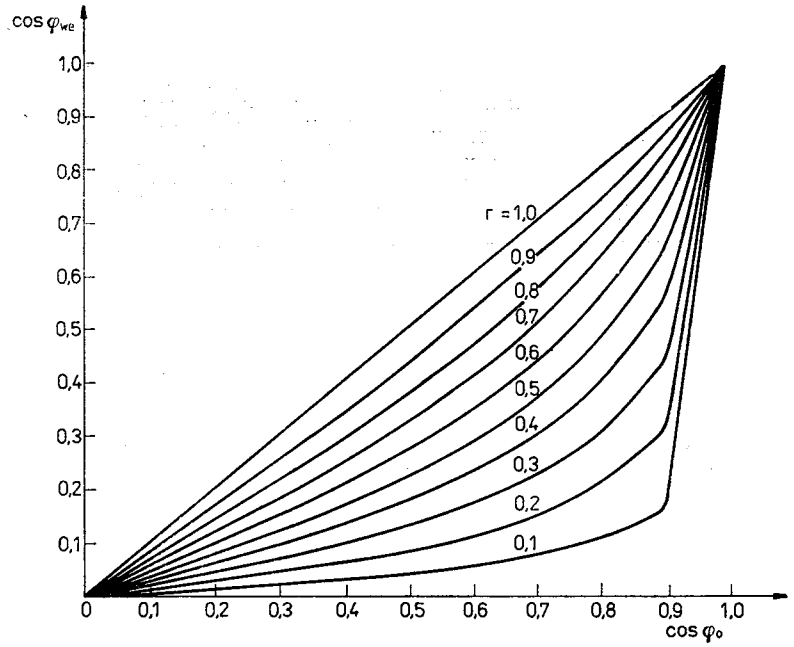


Rys. 3. Zależność względnej wartości wejściowej mocy biernej od charakteru odbiornika ($\cos\varphi_0$), przy różnych wartościach względnej amplitudy napięcia wyjściowego (r) — sterowanie metodą modulacji amplitudowo zależnej

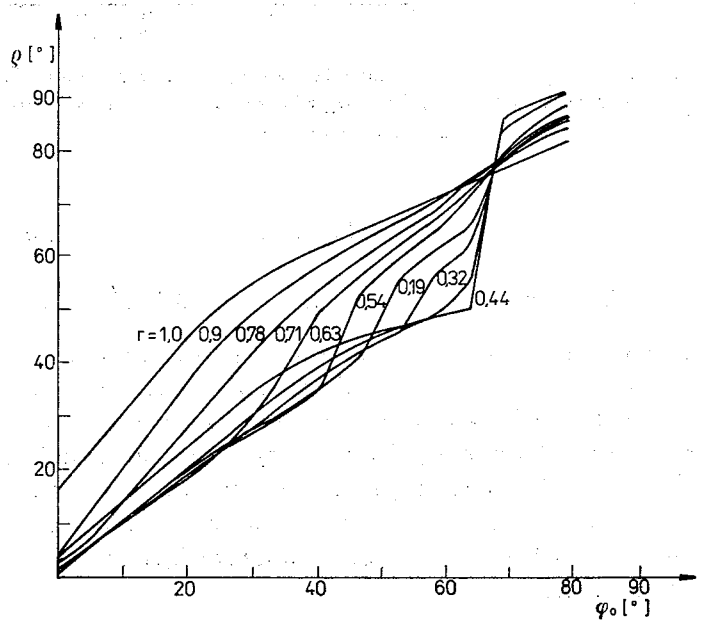
padku sterowania $\varrho = -\frac{\pi}{2}$. Ponieważ cyklokonwertor pobiera z sieci moc bierną indukcyjną niezależnie od charakteru odbiornika, przeto w przedziale $\left\langle -\frac{\pi}{2}, \varphi_0 \right\rangle$ istnieje taka wartość kąta ϱ, przy której nie ma poboru mocy biernej z sieci zasilającej ($Q = 0$). Odpowiednie wartości kąta ϱ w konkretnych warunkach można wyliczyć na modelu cyfrowym przemiennika. Na rys. 5 przedstawiona jest wymagana zmienność kąta ϱ zapewniająca pracę przemiennika bez poboru mocy biernej przy zmianach charakteru obciążenia ($\cos\varphi_0$) i względnej amplitudy wymaganej składowej napięcia wyjściowego (r), uzyskana drogą symulacji cyfrowej dla przykładowej częstotliwości wyjściowej $f_0 = 50$ Hz. Szacunkowo można określić tendencję zmian kąta ϱ umożliwiającą minimalizację mocy biernej na wejściu przemiennika:

$$\varrho = -\frac{\pi}{2} + \varphi_0 \tag{8}$$

Uwzględnienie tak sformułowanej zasady zmian kąta ϱ w obliczeniach na modelu cyfro-



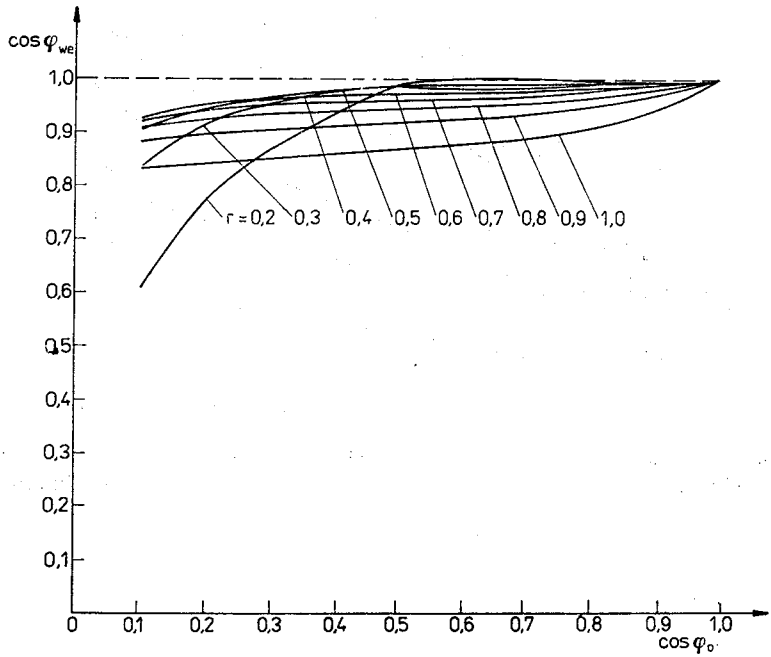
Rys. 4. Zależność wejściowego współczynnika przesunięcia od charakteru odbiornika ($\cos\varphi_o$), przy różnych wartościach względnej amplitudy napięcia wyjściowego (r) — sterowanie metodą modulacji amplitudowo zależnej



Rys. 5. Wymagana zmienność kąta ϱ przy zmianach charakteru obciążenia (φ_o) kiedy przekształtnik nie pobiera mocy biernej ze źródła, przy różnych wartościach względnej amplitudy napięcia wyjściowego (r), częstotliwość wyjściowa $f_o = 50$ Hz

wym, pozwala prześledzić kształtowanie się charakterystycznych wielkości wejściowych przemiennika.

Wejściowy współczynnik przesunięcia $\cos\varphi_{we}$, który w całkowicie „*skompensowanym*" układzie powinien być równy jedności, w przypadku sterowania zgodnie z zależnością (8) ulega zmianom przedstawionym na rys. 6.



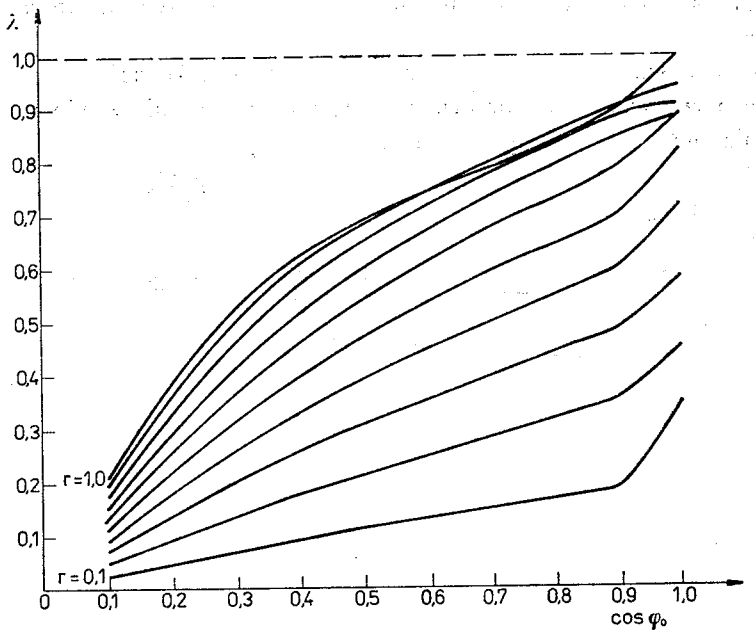
Rys. 6. Zależność wyjściowego współczynnika przesunięcia od charakteru odbiornika ($\cos\varphi_0$), przy różnych wartościach względnej amplitudy napięcia wyjściowego (r) — sterowanie metodą modulacji amplitudowo zależnej, przy $\varrho = -\frac{\pi}{2}+\varphi_0$

Wejściowy współczynnik mocy przemiennika zdefiniowany jako stosunek mocy czynnej do mocy pozornej:

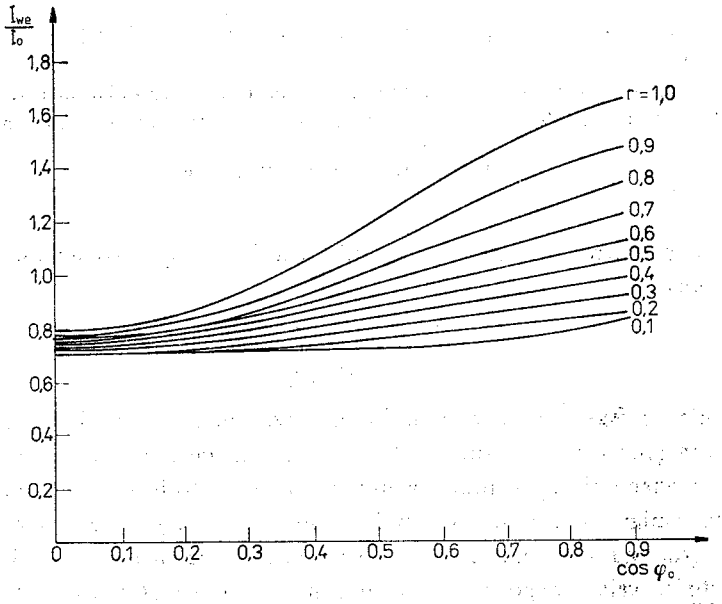
$$\lambda = \frac{P}{S} = \frac{P}{\sqrt{P^2+Q^2+D^2}} \tag{9}$$

różni się znacznie od wejściowego współczynnika przesunięcia ze względu na wpływ mocy deformacji D występującej w zależności (9). Moc deformacji D pobierana z sieci związana jest z istnieniem harmonicznych prądu wejściowego o częstotliwościach różnych od częstotliwości sieci. Zmiany wejściowego współczynnika mocy uzyskane w wyniku symulacji cyfrowej, w przypadku spełnienia zasady (8) przedstawione są na rys. 7. Uzyskanie największych wartości wejściowego współczynnika mocy można utożsamiać z minimalizacją wartości skutecznych prądów pobieranych z sieci zasilającej.

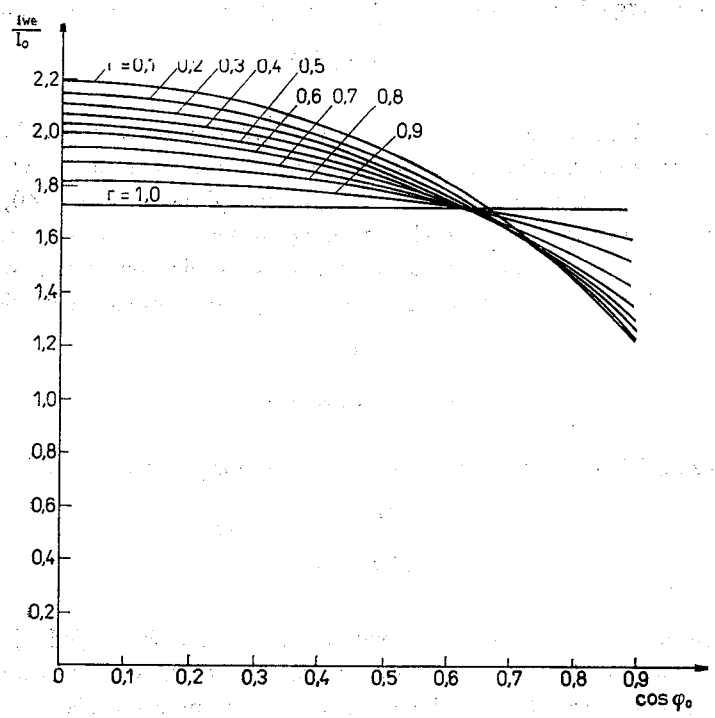
Efekty uzyskane w zakresie zmniejszania wartości skutecznej prądów wejściowych poprzez sterowanie zgodne z zasadą (8) przedstawione są na rys. 8, na którym umieszczone



Rys. 7. Zależność wejściowego współczynnika mocy od charakteru odbiornika ($\cos\varphi_0$), przy różnych wartościach względnej amplitudy napięcia wyjściowego (r) — sterowanie metodą modulacji amplitudowo zależnej, przy $\varrho = -\frac{\pi}{2} + \varphi_0$



Rys. 8. Zależność względnej wartości skutecznej prądu wejściowego od charakteru odbiornika ($\cos\varphi_0$), przy różnych wartościach względnej amplitudy napięcia wyjściowego (r) — sterowanie metodą modulacji amplitudowo zależnej, przy $\varrho = -\frac{\pi}{2} + \varphi_0$



Rys. 9. Zależność względnej wartości skutecznej prądu wejściowego od charakteru odbiornika ($\cos\varphi_0$), przy różnych wartościach względnej amplitudy napięcia wyjściowego (r) — sterowanie metodą modulacji amplitudowo zależnej, przy $\varrho = -\frac{\pi}{2}$

są wykresy zmian wartości skutecznej prądu pobieranego z sieci odniesione do wartości skutecznej I_0, w zależności od charakteru obciążenia i względnej amplitudy wymaganej składowej napięcia wyjściowego. Aby ocenić poziom omawianych efektów wyniki z rys. 8 należy porównać z wynikami zamieszczonymi na rys. 9, które odpowiadają sterowaniu przy stałym kącie $\varrho = -\frac{\pi}{2}$ niezależnie od charakteru obciążenia. W przypadku obciążenia rezystancyjnego $\cos\varphi_0 = 1$ wyniki obu przypadków sterowania (rys. 8, 9) powinny być identyczne.

3. REGULACJA NAPIĘCIA METODĄ MODULACJI SZEROKOŚCI IMPULSÓW

W metodzie, w której regulacja napięcia wyjściowego oparta jest na zmianie szerokości impulsów składowych, stały czas t_p związany z pulsacją wyjściową ω_0(4), określony przy maksymalnym napięciu wyjściowym dzieli się na dwa przedziały. Jeden o czasie trwania γt_p, w którym przekształtnik pobiera prąd z sieci i drugi $(1-\gamma)t_p$, w którym elementy przekształtnika zawierają odpowiednio odbiornik i odłączają go od sieci zasilającej, ($\gamma \in \langle 0,1\rangle$). Związek pomiędzy względną amplitudą wymaganej składowej napięcia

wyjściowego r a względną szerokością impulsów γ jest następujący [2, 3]:

$$r = 2\sin\frac{\gamma\pi}{6}. \tag{10}$$

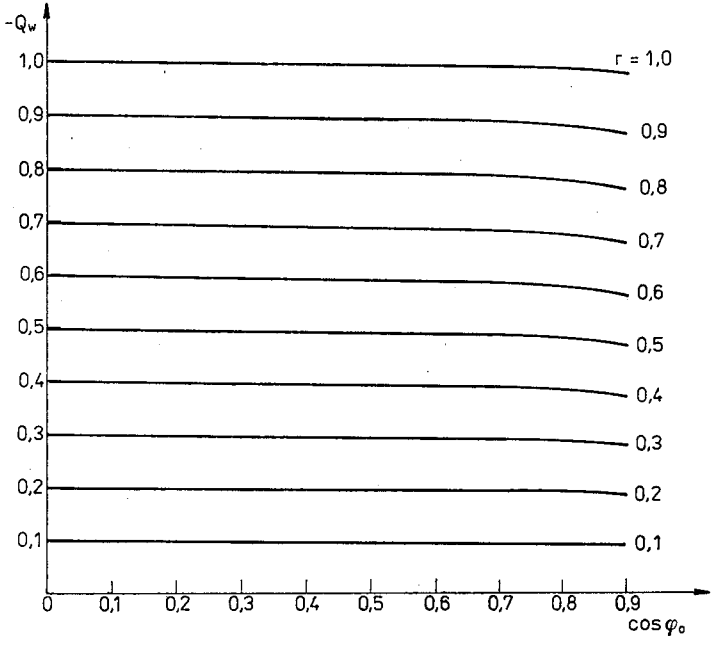
Można wykazać analitycznie [3], ponadto potwierdzają to wyniki obliczeń numerycznych, że w przypadku regulacji szerokości impulsów przemiennik — w odniesieniu do mocy biernej — w całym zakresie zmian napięcia wyjściowego ($r \in \langle 0, 1\rangle$) zachowuje właściwości takie jak przy maksymalnym napięciu wyjściowym. Oznacza to, że moc bierna pobierana z sieci i bierna moc obciążenia są sobie równe lecz mają przeciwne znaki. Związek między mocą bierną wejściową i wyjściową można wyrazić jako:

$$\frac{Q}{r\dfrac{9}{\sqrt{2}\pi}U_m I_0 \sin\varphi_0},$$

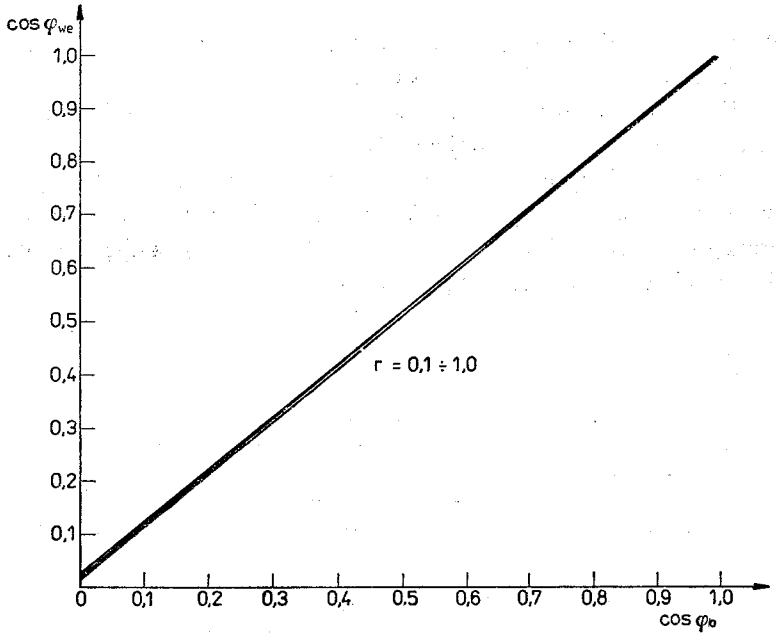
albo pomiędzy wejściową mocą bierną u maksymalną mocą bierną na wyjściu (przy $r = 1$):

$$\frac{Q}{\dfrac{9}{\sqrt{2}\pi}U_m I_0 \sin\varphi_0} = Q_w. \tag{11}$$

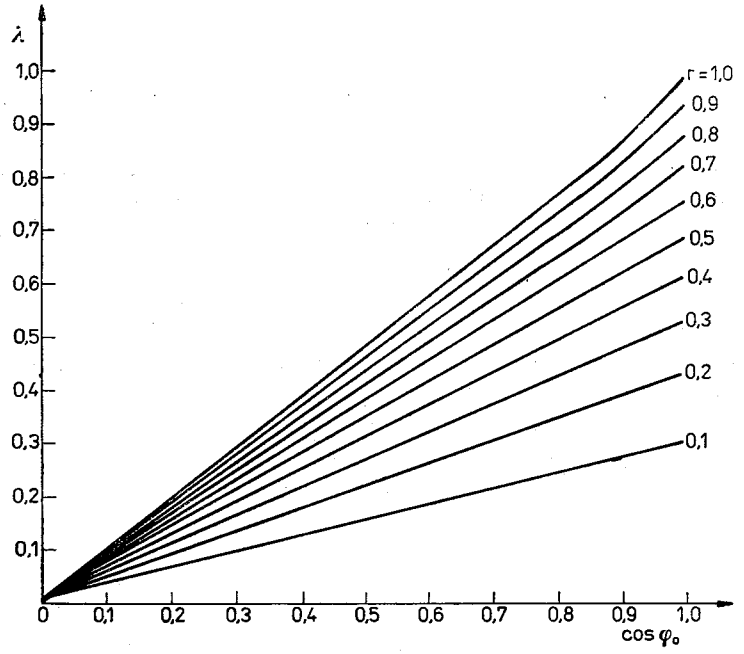
Na rys. 10 przedstawione są wyniki obliczeń numerycznych względnej wartości mocy biernej, określonej zgodnie z lewą stroną wyrażenia (11). W omawianym przypadku sterowania oczywiste jest, że wejściowy współczynnik przesunięcia $\cos\varphi_{we}$ jest równy



Rys. 10. Zależność względnej wartości wejściowej mocy biernej od charakteru odbiornika ($\cos\varphi_0$), przy różnych wartościach względnej amplitudy napięcia wyjściowego (r) — sterowanie metodą szerokości impulsów



Rys. 11. Zależność wejściowego współczynnika przesunięcia od charakteru odbiornika ($\cos\varphi_o$), przy różnych wartościach względnej amplitudy napięcia wyjściowego (r) — sterowanie metodą szerokości impulsów

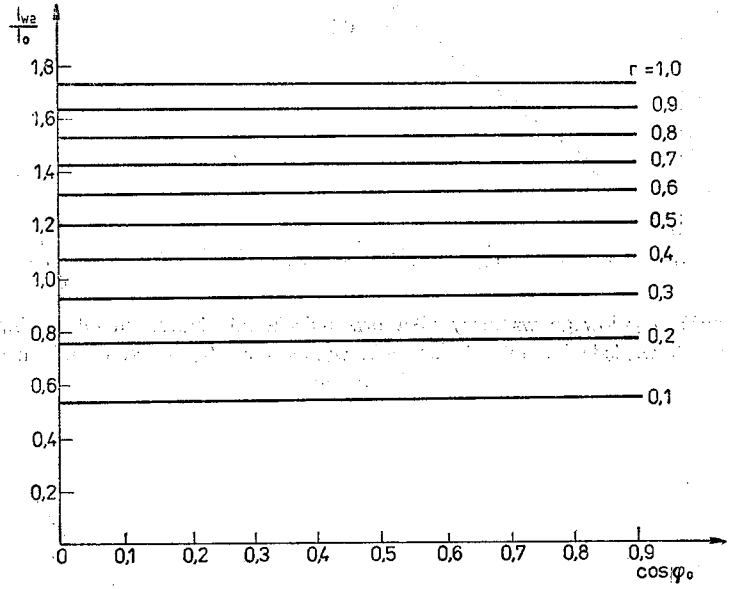


Rys. 12. Zależność wejściowego współczynnika mocy od charakteru odbiornika ($\cos\varphi_o$), przy różnych wartościach względnej amplitudy napięcia wyjściowego (r) — sterowanie metodą szerokości impulsów

współczynnikowi obciążenia $\cos\varphi_0$, lecz ma przeciwny charakter. Właściwość tę potwierdzają wyniki obliczeń przedstawione na rys. 11.

Zmiany wejściowego współczynnika mocy λ określonego zgodnie z zależnością (9), a uzyskane w efekcie symulacji cyfrowej przedstawione są na rys. 12.

Ostatnią z przedstawionych wielkości charakterystycznych świadczących o właściwościach przemiennika sterowanego metodą szerokości impulsów jest wartość skuteczna prądu pobieranego z sieci odniesiona do wartości I_0. Zmiany względnej wartości skutecznej prądu wejściowego przedstawione są na rys. 13.



Rys. 13. Zależność względnej wartości skutecznej prądu wejściowego od charakteru odbiornika ($\cos\varphi_0$), przy różnych wartościach względnej amplitudy napięcia wyjściowego (r) — sterowanie metodą szerokości impulsów

Porównanie ze sobą dwóch omówionych kolejno metod sterowania można przeprowadzić analizując wyniki z rys. 8 i rys. 13 lub z rys. 7 i rys. 12. Ze względów praktycznych należy zwrócić uwagę na to, że w sterowaniu metodą szerokości impulsów niezbędna jest dwukrotnie większa częstotliwość przełączeń przekształtnika niż w sterowaniu metodą modulacji amplitudowo zależnej. Częstotliwość przełączeń związana jest z istotnymi dodatkowymi stratami w układzie przemiennika rzeczywistego.

PODSUMOWANIE

Prezentowane wyniki obliczeń numerycznych obrazują kształtowanie się parametrów wejściowych bezpośredniego przemiennika częstotliwości przy dwóch konkurencyjnych metodach sterowania. Główny wniosek wynikający z porównania tych metod nie powinien być bez zastrzeżeń utożsamiany ze stosowaniem metody modulacji amplitudowo

zależnej z uwzględnieniem zasady $\varrho = -\frac{\pi}{2}+\varphi_0$ (8). Realizacja praktyczna tej zasady może okazać się kłopotliwa. Istotny jest natomiast fakt, że poprzez zmianę wartości kąta ϱ można minimalizować moc bierną lub wartość skuteczną prądu pobieranego ze źródła. Praktycznie minimalizacja taka może być zrealizowana np. w układzie regulacji, w którym moc bierna będzie wielkością mierzoną, a kąt ϱ — regulowaną.

DODATEK

W dodatku jest naszkicowany i skomentowany sposób obliczania kolejnych wielkości wejściowych przemiennika.

1. Moc czynna na wejściu jest obliczana za pomocą całki z mocy chwilowej w okresie powtarzalności przebiegu. Okres powtarzalności jest związany z częstotliwością wyjściową (napięcia zasilającego). Jest to moc uśredniona w okresie T_0 (T_0 — okres wymaganej składowej napięcia wyjściowego).

$$P = \frac{\omega_0}{2\pi}\int_0^{T_0} [u_R u_S u_T]\begin{bmatrix} i_R \\ i_S \\ i_T \end{bmatrix}. \tag{1}$$

Przy założeniu, że napięcia sieciowe u_R, u_S, u_T są symetryczne i sinusoidalne, po przejściu z wartości chwilowych prądów fazowych i_R, i_S, i_T do wektora prądu wejściowego definiowanego jako:

$$i_{we} = \frac{2}{3}[1aa^2]\begin{bmatrix} i_R \\ i_S \\ i_T \end{bmatrix};\left(a = e^{j\frac{2\pi}{3}}\right), \tag{2}$$

moc czynną na wejściu można obliczać za pomocą całkowania rzutu wektora prądu wejściowego na wektor napięć sieciowych określony zależnością analogiczną do (2) [4].

$$P = \frac{1}{T_0}\frac{\sqrt{3}}{2}\int_0^{T_0} U_m Re(i_{we}^*) \tag{3}$$

$Re(i_{we}^*)$ — część rzeczywista wektora prądu i_{we}^* przeniesionego do układu wirującego współfazowo z wektorem napięć sieciowych.

2. Moc bierną na wejściu, uśrednioną w okresie T_0, można wyznaczyć (w sposób analogiczny do (3)) przez całkowanie rzutu wektora prądu wejściowego na oś prostopadłą do wektora napięć sieciowych.

$$Q = \frac{1}{T_0}\frac{\sqrt{3}}{2}\int_0^{T_0} U_m I_m(i_{we}^*) \tag{4}$$

$I_m(i_{we}^*)$ — część urojona wektora i_{we}^*.

3. Wejściowy współczynnik przesunięcia zdefiniowany jest jako:

$$\cos\varphi_{we} = \frac{P}{\sqrt{P^2+Q^2}} \tag{5}$$

Jest to współczynnik uśredniony w okresie T_0.

4. Moc pozorna przy napięciach sieci sinusoidalnych i symetrycznych jest określona sumą wartości skutecznych prądów fazowych:

$$S = \frac{U_m}{\sqrt{2}}(I_R + I_S + I_T) \tag{6}$$

W przypadku, gdy wartości skuteczne prądów w poszczególnych fazach zasilających są sobie równe (brak mocy asymetrii), moc pozorną można obliczać na podstawie znajomości wektora prądu zasilającego w sposób następujący:

$$S = \frac{3U_m}{2\sqrt{2}} \sqrt{\frac{1}{T_0}\int_0^{T_0} |i_{we}|^2 \mathrm{d}t}. \tag{7}$$

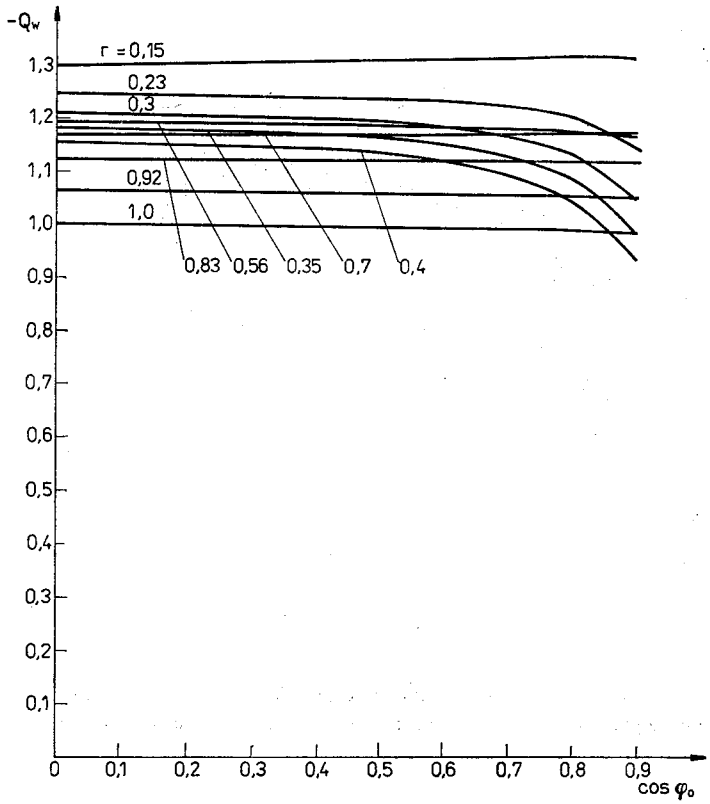
5. Wartości skuteczne prądów I_R, I_S, I_T są obliczane z definicji, przez całkowanie kwadratów wartości chwilowych i_R, i_S, i_T w okresie T_0.
6. Wejściowy współczynnik mocy jest równy:

$$\lambda = \frac{P}{S}. \tag{8}$$

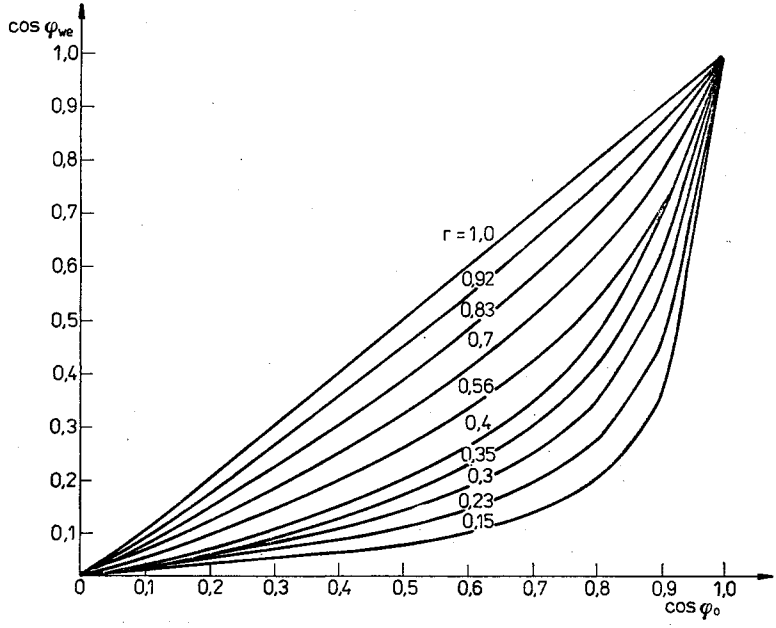
Wszystkie wyniki prezentowane w artykule zostały uzyskane drogą symulacji cyfrowej przebiegów wielkości wejściowych przemiennika, przy złożeniu natychmiastowej komutacji zaworów, pominięciu strat mocy w przekształtniku i wykorzystaniu zależności: (3), (4), (5), (6), (8).

Pewnego komentarza wymaga założenie powtarzalności przebiegów wejściowych po okresie wyjściowym T_0. Teoretycznie, przebiegi prądów wejściowych są powtarzalne w okresie T_0 tylko dla takich pulsacji wyjściowych ω_0, dla których pulsacja wejściowa ω jest całkowitą wielokrotnością. Przebiegi wyjściowe powtarzają się po okresie T_0 również tylko dla takich pulsacji wyjściowych ω_0, dla których liczba $q\omega$ jest całkowitą wielokrotnością (q — pulsacyjność przemiennika). Założenie o powtarzalności przebiegów po okresie T_0, dla dowolnych pulsacji wyjściowych, można przyjąć w przypadku, gdy sterowanie przemiennika odbywa się w układzie zamkniętym i składowe podharmoniczne w rozpatrywanych przebiegach są eliminowane.

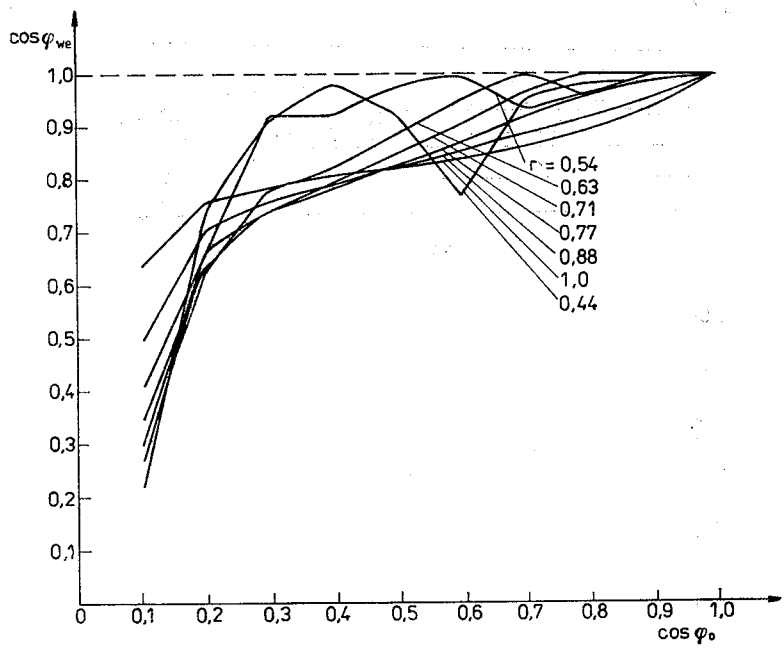
Istotne jest również wyjaśnienie, czy prezentowane wyniki dla konkretnej częstotliwości wyjściowej $f_0 = 5$ Hz można uogólnić na dowolne częstotliwości wyjściowe, a szczególnie — bliskie częstotliwości wejściowej f. W przypadku sterowania metodą regulacji szerokości impulsów prezentowane wyniki są identyczne dla wszystkich pulsacji wyjściowych. W przypadku sterowania metodą modulacji amplitudowo zależnej, przy innych częstotliwościach mogą wystąpić pewne zmiany ilościowe możliwe do wyeliminowania w zamkniętych układach sterowania. Wyniki uzyskane przy stosunkowo niskiej częstotliwości wyjściowej zachowują swój charakter także i dla innych pulsacji. Wskazuje na to porównanie wyników uzyskanych dla częstotliwości wyjściowej $f_0 = f = 50$ Hz, przedstawionych na rys. 1D—6D z wynikami dla $f_0 = 5$ Hz zamieszczonymi w głównej części artykułu odpowiednio na rys. 3, 4, 6, 7, 8, 9.



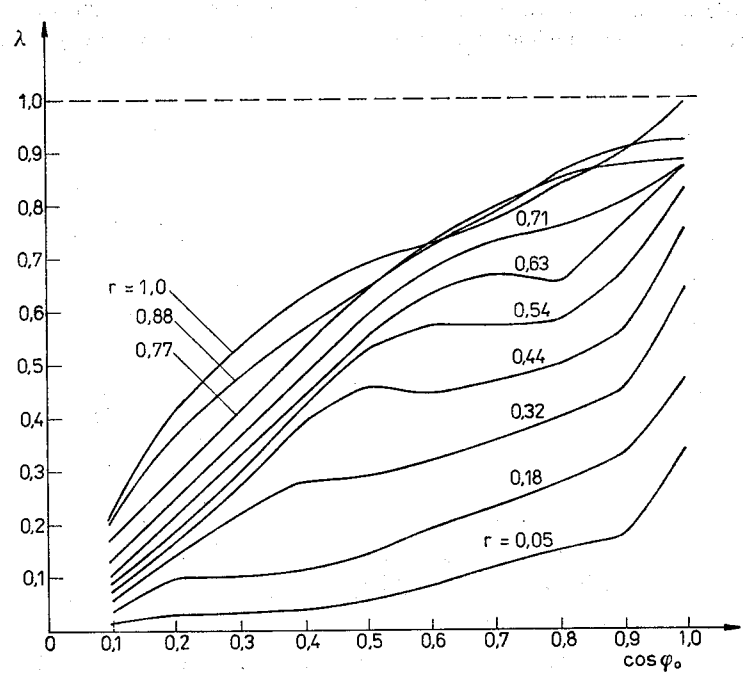
Rys. 1D. Zależność względnej wartości wejściowej mocy biernej od charakteru odbiornika ($\cos\varphi_o$) przy różnych wartościach względnej amplitudy napięcia wyjściowego, dla częstotliwości wyjściowej $f_o = 50$ Hz



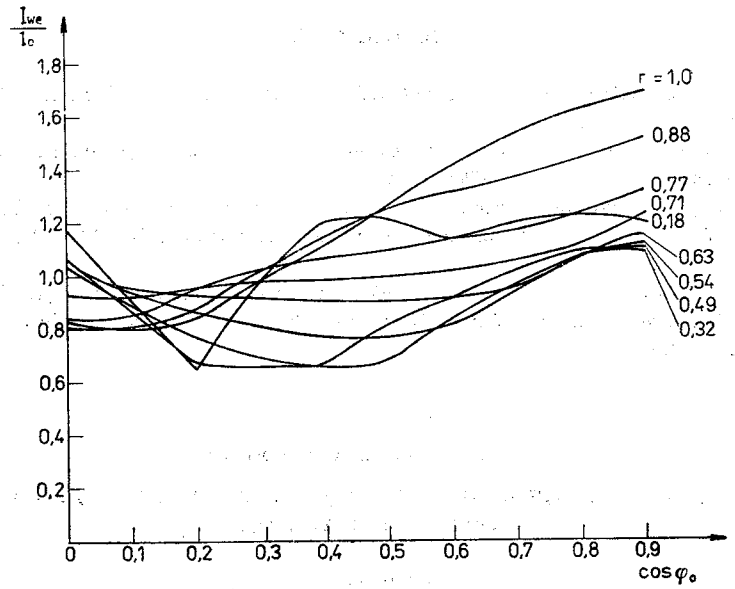
Rys. 2D. Zależność wejściowego współczynnika przesunięcia od charakteru odbiornika ($\cos\varphi_o$) przy różnych wartościach względnej amplitudy napięcia wyjściowego, dla częstotliwości wyjściowej $f_o = 50$ Hz



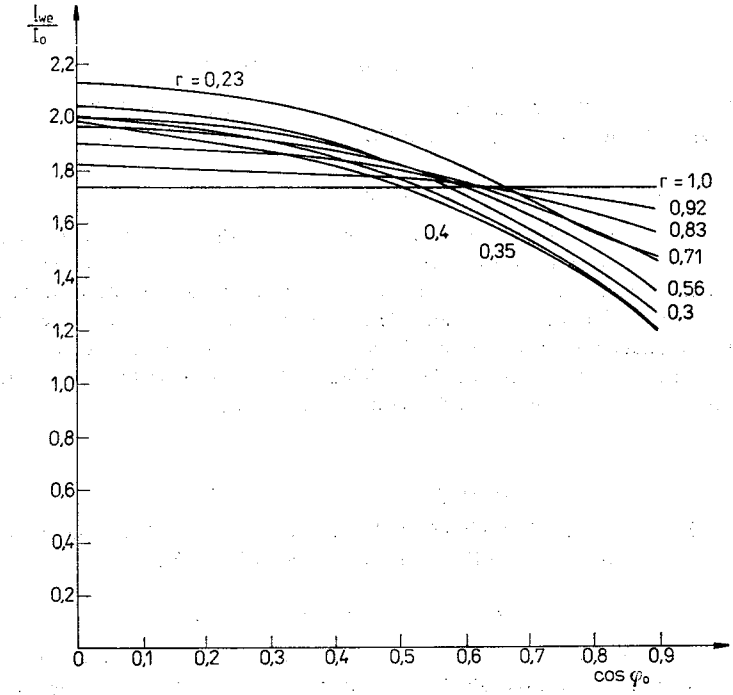
Rys. 3D. Zależność wejściowego współczynnika przesunięcia od charakteru odbiornika ($\cos\varphi_o$), przy różnych wartościach względnej amaplitudy napięcia wyjściowego, dla częstotliwości wyjściowej $f_o = 50$ Hz



Rys. 4D. Zależność wejściowego współczynnika mocy od charakteru odbiornika ($\cos\varphi_o$), przy różnych wartościach względnej amplitudy napięcia wyjściowego, dla częstotliwości wyjściowej $f_o = 50$ Hz



Rys. 5D. Zależność względnej wartości skutecznej prądu wejściowego od charakteru odbiornika ($\cos\varphi_0$), przy różnych wartościach względnej amplitudy napięcia wyjściowego, dla częstotliwości wyjściowej $f_0 = 50$ Hz



Rys. 6D. Zależność względnej wartości skutecznej prądu wejściowego od charakteru odbiornika ($\cos\varphi_0$), przy różnych wartościach względnej amplitudy napięcia wyjściowego, dla częstotliwości $f_0 = 50$ Hz

BIBLIOGRAFIA

1. T. Citko: *Bezpośrednie przemienniki częstotliwości.* Zeszyty Naukowe Politechniki Białostockiej — Elektryka Nr 2 Białystok 1986
2. L. Gyugi, B. R. Pelly: *Static power frequency changers theory performance and aplication.* New York, Sydney, Toronto, John Wiley 1976
3. *Analiza wejściowych i wyjściowych parametrów bezpośrednich przemienników częstotliwości przy różnych metodach ich sterowania.* Sprawozdanie z pracy badawczej w ramach CPBR. Politechnika Białostocka 1986
4. T. Citko: *Sposób obliczania parametrów układu bezpośredniego przemiennika częstotliwości o komutacji naturalnej obciążonego maszyną indukcyjną przy ciągłym prądzie stojana.* Archiwum Elektrotechniki 1983, Zeszyt 3/4

T. CITKO, Z. DASZUTA, A. BOGDAN

EFFECT OF THE CONTROL OF A DIRECT FREQUENCY CHANGER ON THE REACTIVE POWER CONSUMPTION

Summary

In the article the changes of the input reactive power of the ac to ac convertors controlled by: — pulse width modulation method or by: amplitude — dependent frequency modulation method — are presented. The results were obtained by using computer simulation and calculations.
These results conformed the possibility of minimalization of the reactive power consumption using the control by angle ϱ (ϱ — determines the moment of exchange the complementary voltage waves) in the amplitude — dependent frequency method.

T. CITKO, Z. DASZUTA, A. BOGDAN

INFLUENCE DES MÉTHODES DE COMMANDE DIRECTE À L'AIDE D'UN CONVERTISSEUR DE FREQUENCE SUR LA PRISE DE PUISSANCE REACTIVE

Résumé

Dans l'étude on présente les changements de la puissance réative prise directement par un convertisseur de fréquence en appliquant les deux méthodes alternatives de régulation de la tension de sortie: méthode de modulation dépendant de l'amplitide et méthode de régulation de la largeur d'impulsion. Les résultats y présentés ont été obtenus à l'aide de la simulation numérique des parcours et des calculs sur calculateur. Les résultats obtenus ont prouvé la possibilité de la minimisation de la prise de puissance réactive lors de l'application de la méthode de modulation dépendante de l'amplitude au moyen de la commande de l'angle diu changement de la forme d'onde de tension.

T. CITKO, Z. DASZUTA, A. BOGDAN

EINFLUSS DER STEUERUNGSMETHODE EINES UNMITTELBAREN FREQUENZUMSETZERS AUF DIE BLINDLEISTRUNGSENTNAHME

Zusammenfassung

In Beitrag werden Änderungen der Blindleistung geschildert, die durch einen unmittelbaren Frequenzumsetzer unter Anwendung von zwei alternativen Regelungsmethoden der Ausgangsspannung entnommen wird: der amplitudenabhängigen Modulationsmethode und der Ragelungsmethode der Impulsbreite. Die dargestellten Ergebnisse wurden im Wege der Simulierung der Digital verlaufe sowie rechner-

gestützten Berechnungen gewonnen. Die gewonnenen Ergebnisse bestätigten die Möglichkeit (schätzten das Niveau ein) einer Minimalisierung der Blindleistungsentahme bei der amplitudenabhängigen Modulationsmethode mittels Steuerung mit dem Änderungswinkel der Spannungswellenform.

Т. ЦИТКО, З. ДАШУТА, А. БОГДАН

ВЛИЯНИЕ МЕТОДА УПРАВЛЕНИЯ НЕПОСРЕДСТВЕННЫМ ПРЕОБРАЗОВАТЕЛЕМ ЧАСТОТЫ НА ПОТРЕБЛЕНИЕ РЕАКТИВНОЙ МОЩНОСТИ

Р е з ю м е

Представлены изменения реактивной мощности потребляемой непосредственным преобразователем частоты для двух методов регулирования выходного напряжения:
— метода регуляции ширины импульсов,
— метода синусоидального ШИМ.

Представлены результаты полученные путем цифровой симуляции процессов и машинных вычислений. Полученные результаты показали, что для управления по закону синусоидального ШИМ возможно сведение к минимуму потребления реактивной мощности путем подбора угла изменения волны выходного напряжения.

ROZPRAWY ELEKTROTECHNICZNE, 1988, 34, z. 4, ss. 1149–1158

621.372.8:535

Propagation characteristics of multiclad F-doped silica fibers

ADAM MAJEWSKI, ZBIGNIEW WAWRZYNIAK

Instytut Podstaw Elektroniki, Politechnika Warszawska

Received 1987.07.20

Authorized 1988.03.09

We report some considerations on triple-clad fluorine-doped silica fibers designed for a strictly single-mode behaviour over a wide spectral range. It is possible to design multiclad single-mode fibers only with moderate levels of fluorine and realize most of the benefits of dispersion-flattened fibers at the 1.55 μm range with the HE_{12} mode cutoff below 1.4 μm. Such fibers seem to be especially necessary for the WDM when installing the ISDN.

1. INTRODUCTION

Optical fibers with a low dispersion over a wide wavelength range are very attractive for high-capacity long-distance optical transmission links. So far, the greatest efforts have been concentrated mainly not only on step-index and W-profile fibers but also on multiple-index ones [1, 2, 3] for this purpose. However, the dispersion performance of the dispersion-flattened single-mode fibers is still well-known only almost numerically for these more sophisticated index-profiles [4, 5] that have also a considerable microbending sensitivity [6].

The use of fluorine as dopant in optical fiber fabrication has been demonstrated to significantly improve the properties of the fibers [7] lowering the OH excess loss reduing hydrogen degradation effects and allowing the preparation of dipfree pure-silica fiber structures. The fluorine dopants seem to prove their evidence for high quality fibers especially for the PCVD method [8]. In the present paper we report some characteristics of triple-clad F-doped silica fibers designed for a strictly single-mode behaviour over a wide wavelength range.

2. CALCULATIONS

To investigate an arbitrary index-profile and the influence of the profile distortions on the propagation characteristics we use a model of the multistep-index fiber [9] based on an accurately solved system of Maxwell's equations subjected to proper boundary conditions. The real dispersion characteristics of the F-doped silica [10] are included

in the calculations. Detail considerations on the numerical aspects are published elsewhere [11]. However, it should be pointed out that the approach has proved its accuracy and flexible effectiveness for the real fibers [12] also avoiding a numerical unaccuracy found in computation method due to use of the wave equation instead of the Maxwell's ones that is seen in other papers concerning the fiber in question [13, 14, 15, 16].

3. RESULTS

The concept of providing the attractive option of low dispersion over a certain range of the wavelengths suggested initially for the W-fibers has proved to be more properly suited to the triple- (TC) and the quadruple-clad (QC) fiber giving a dispersion curve which turns over two or three wavelengths of zero dispersion. However, QC-fibers, that have demonstrated the high degree in achieving the desired dispersion performance, are extremely sensitive to small changes in fiber design parameters [4, 6] and the excellent dispersion performance can be achieved only at the expense of a reduction in the spot size [17].

The TC-fibers seem not to suffer from considerable bending sensitivity and prevent a low fundamental cutoff wavelength. Although the maximum concentation of the F dopants in pure silica by the PCVD method enables that the refractive-index depresion relative to pure silica reaches the value of 2.3% [18], the presented analysis of the TC-fibers is made for the moderate values of F-dopant concentration not exceeding that of 2 M% [19], for which the exact dispersion characteristics are already known [10]. A specific set of fiber design parameters to confine the dispersion to $\pm$ 0.3 ps/km/nm over the spectral range of $1.31 \div 1.61$ μm (see Fig. 1) is proposed, which demonstrates that the

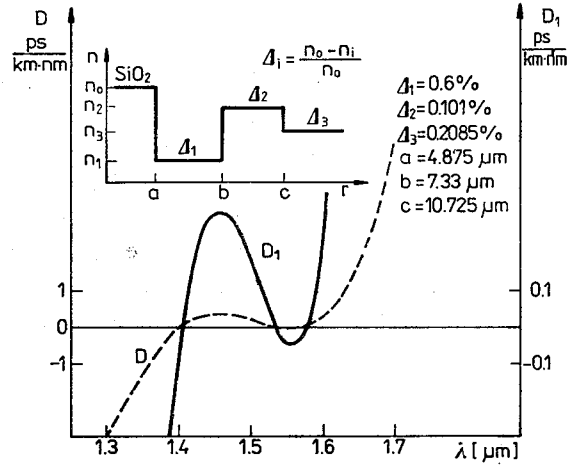


Fig. 1. The TC-fiber design with its dispersion characteristic

penetration of light in the second clad is very weak (Fig. 2). This results from the flattened dispersion characteristics of the fiber in Fig. 1 (see that the scale of D_1 is one tenth of that of D) and the phase characteristics optimized for the 1.55 μm-range having a long cutoff

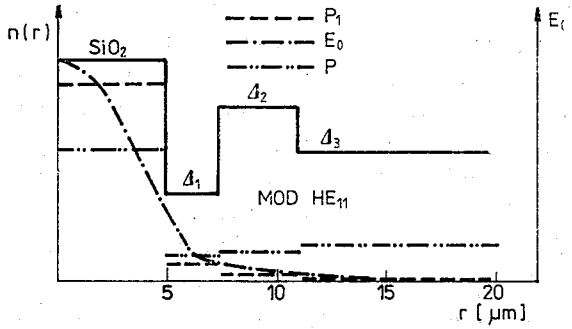


Fig. 2. Relative field E_0 and power distribution P_1 of the fundamental mode at 1.3 μm and power distribution P at 1.6 μm

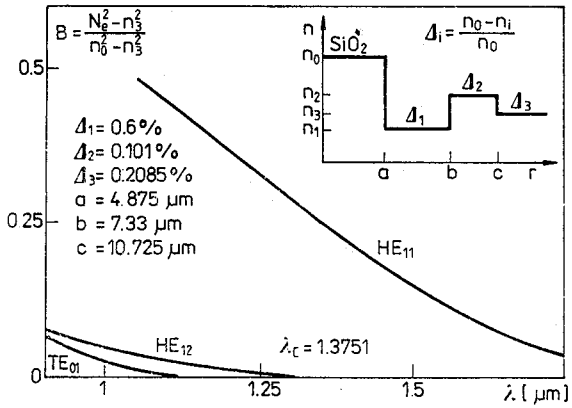


Fig. 3. Normalized phase constant B vs wavelength λ for the HE_{11}. HE_{12} and TE_{01} mode for the TC-fiber

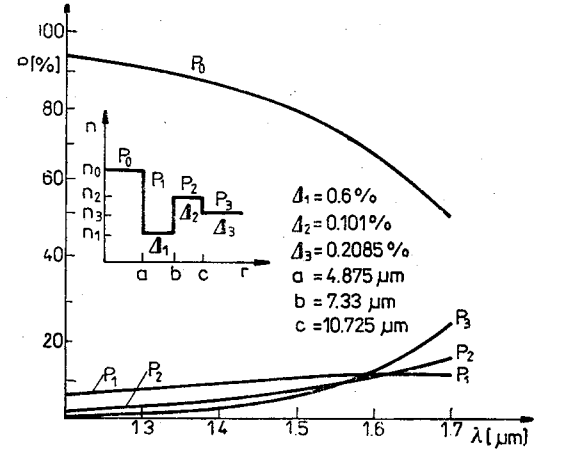


Fig. 4. Fractional power distribution propagated in each layer (P_0 denotes the power in the core)

wavelength of the HE_{12}-mode (Fig. 3) that has been based on the fact to fully utilize the maximum of the index depression in the second layer. In fact, the accurately calculated value $\lambda_c = 1.3751$ μm for the HE_{12}-mode seems to be much lower in practical measurements due to the existence of the nonvanished losses of the material itself.

For the fundamental mode the moderate value of the phase constant influences the fractional power distribution in each layer (Fig. 4) decreasing its value in the core such as the ratio of the power in the core and in the clads falls down for longer wavelengths. At 1.55 μm 73.5% of the total power propagates in the core (P_0 in Fig. 4) and in the second layer P_1 only 10.2% of the power is transmitted. The effects of the low index difference for the second layer and the optimized dispersion curve over a wide range affect also the value of the spot size. This performance is remarkably illustrated in Fig. 5. where the variation

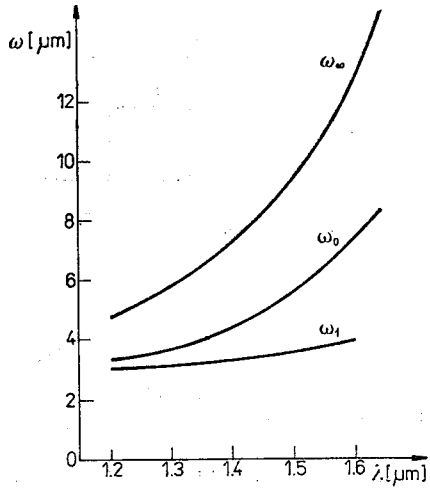


Fig. 5. The spot size ω_0, ω_1 and ω_∞ as a function of the wavelength λ

of the three spot sizes, based on the fundamental definitions as in [20, 17, 6] is shown as a function of wavelength for the considered fiber design. The increasing excess between lower and upper estimation of the spot size should be considered as the indication of microbending losses in the fiber. Too larger spot size implies poor mode confinement, hence susceptibility to bending losses.

The propagation characteristic optimization in such kind of the fibers with the propagation constant of the order of 10% can hardly be obtained when it is based on all the desired purposes. Complexity of the fiber parameters in the dispersion performance and the cutoff wavelength (see Fig. 6) as well as in microbending loss component appeared for the fiber in question is no well-defined directly. Additionally this relation is implicitly complicated, what should be pointed out, by the nonlinear characteristics of the material dispersion included in the method. Therefore, it is necessary to distinguish among different goals of optimization for propagation characteristics and the only approach to the problem is to choose carefully step-by-step a proper set of the parameters, basing on the understanding of the nature of the fiber with a low-propagation constant. Generally,

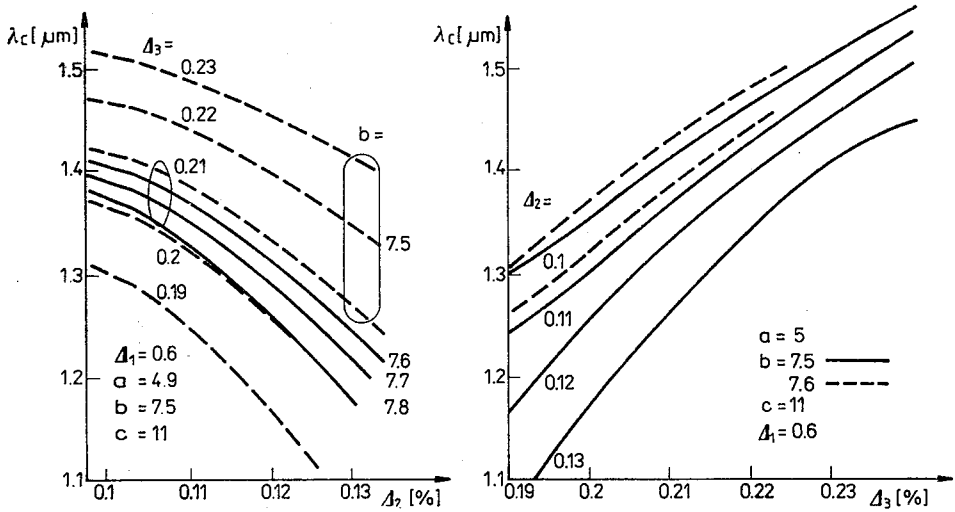


Fig. 6. Influence of parameter variations on the cutoff wavelength λ_c for the TC-fiber

the flattening of the dispersion curve increases the cutoff wavelength not affecting the single-mode range of the low-loss window at 1.55 μm and having small improvement in performance as provides added resistance to microbendings.

The other question of practical importance is that the fiber design obtained should not have been considerably affected by variations of the parameters. This kind of currently important reproducibility aspect of multiclad designs has been also critically examined. The extreme sensitivity of the dispersion characteristics to small changes in the fiber design parameters is illustrated in Fig. 7 and Fig. 8, when the variation of the dispersion with wavelength is shown as a function of the fiber design parameters. With the radius a = = 4.875 μm the dispersion zeros occur at 1.408, 1.536 and 1.578 μm. Keeping all the parameters constant, the only 1.% change in the radius a results in the lack of the zero dispersion near 1.55 μm. During the process of drawing the fiber from the preform one can observe the same relative change of each radius (the radial scale factor F). The influence of the F-factor on dispersion spectra is shown in Fig. 9. From this consideration we see that the required control of the scale factor F should be better than 0.2% in order to keep the change in the dispersion coefficient below 0.24 ps/nm/km which is difficult to match in a technological process of drawing.

The dimensional control will be always a key factor for a such optimized dispersion-flattened fibers in making the fibers reproducibly. Also, a doping control is required to be done but not only at so high level as that of the dimensional one. The rate of change of the dispersion D, with the radius r of the layer, $S = \mathrm{d}D/\mathrm{d}r$ (where r denotes a, b or c), and the sensitivity of the dispersion to small changes in the index differences, $S = \mathrm{d}D/\mathrm{d}\Delta$ and the scale factor F, $S = \mathrm{d}D/\mathrm{d}F$, are given in Fig. 10.

From these considerations, we see that much more additional care should be paid during the design process of a fiber than it could seem while searching for the optimization purposes and also the tolerance requirements have to be incorporated into complexity of the fiber design parameters for the propagation characteristics.

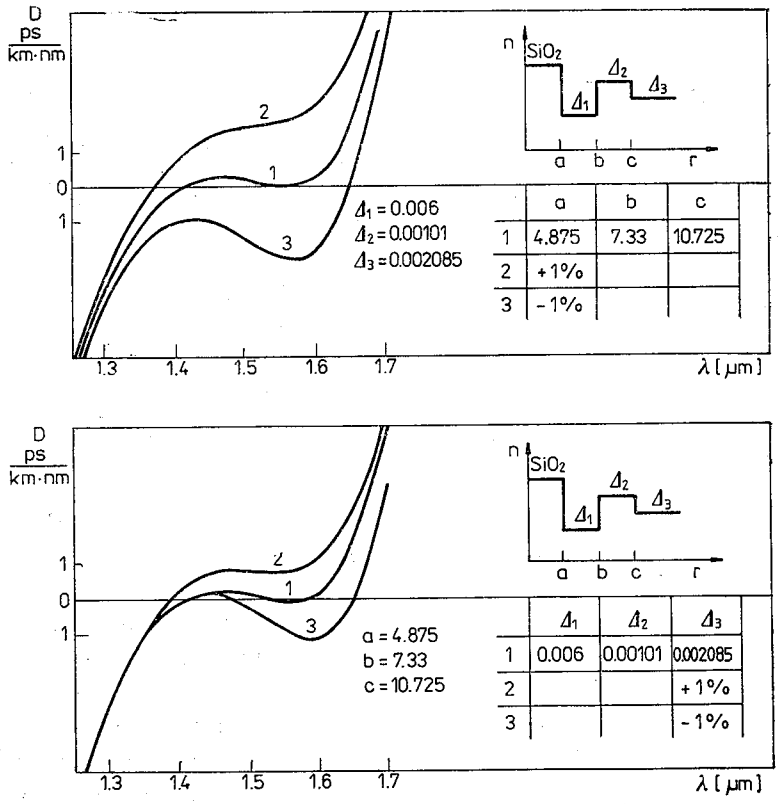


Fig. 7. The effect of small change of 1% in radius a and in concentration Δ_3 on dispersion spectra

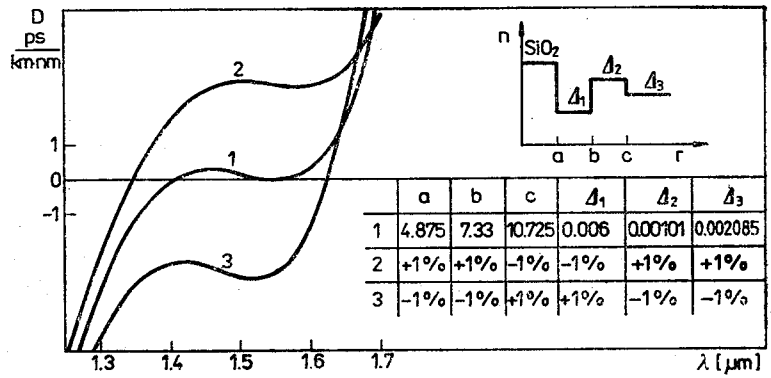


Fig. 8. The effect of small change of 1% in design parameters (the 1%-worst case design) on dispersion spectra

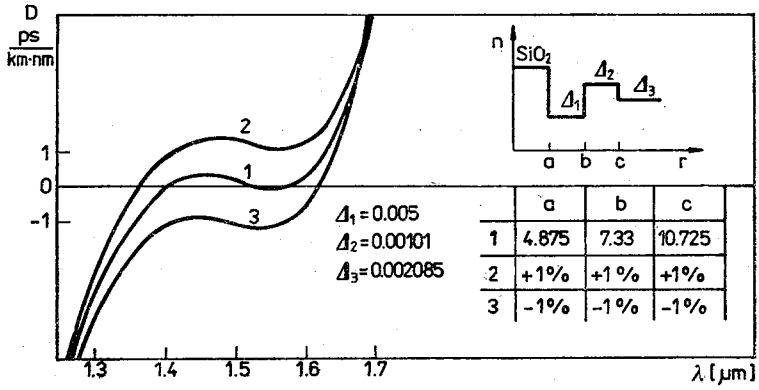


Fig. 9. The effect of the scale factor F on dispersion spectra

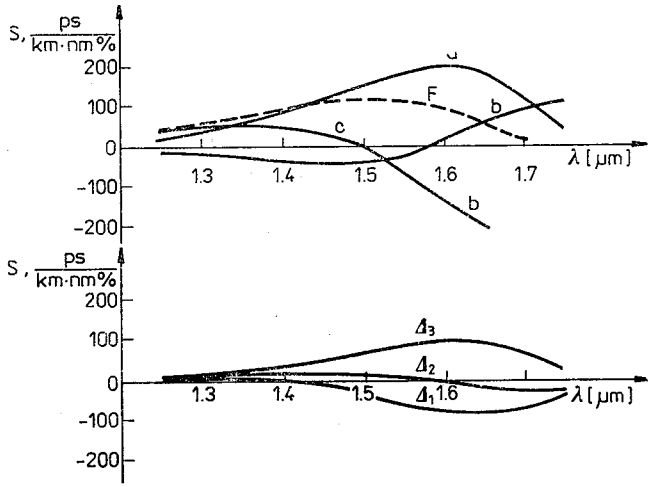


Fig. 10. Sensitivity of dispersion spectra to each of the fiber design parameters and the scale factor F

CONCLUSIONS

It is possible to design multiclad single-mode fibers only with moderate doping levels of fluorine and realize most of the benefits of dispersion-flattened fibers. Although similar results can be realized with GeO_2- and F-dopants in pure silica, the doping only by the use of fluorine seems to be attractive from the lower attenuation obtained, and lower sensitivity achieved to hydrogen, and to radiation, comparing to that of the formers. The zero of the dispersion for the F-doped silica fibers is easily shifted to the 1.55 μm-range even though the highest level of F-doping in the clads has rather moderate value in comparison to those that are achievable by the use of GeO_2- and F-dopants.

A good understanding of the nature of the single-mode fiber with a low refractive-index difference, based on accurately carried out computer analysis, and a very sensitive

character of the flattened dispersion fibers to changes of their parameters indicate the possible approach to designing a truly single-mode *F*-doped silica fibers.

Such a kind of the single-mode fiber with a low-dispersion over a wide wavelength range near the 1.55 μm window and a long cutoff wavelength $\lambda_c > 1.3$ μm seems to be especially necessary for the wavelength division multiplexing (WDM) when installing an integrated optical broadband network (e.g. ISDN) in the very near future.

The evidence is presented that the well-compensated dispersion fibers are very sensitive to small variations of the index profile and imperfections that occurr during the technological processes. In optimization of the propagation characteristics, the dispersion performance, the cutoff wavelength, microbending losses, the spot size and tolerance requirements should be considered in details although the relative importance of each of these effects depends on the design objectives and priorities. It seems that restricting the flattening of the dispersion to a more modest range allows a wider range of designs to be achieved by the systematic optimization but the dimensional control remains an important issue.

REFERENCES

1. L. G. Cohen, W. L. Mammel, S. J. Jang: *Low-Loss Quadruple-Clad Single-Mode Lightguides with Dispersion below 2 ps/km/nm over the 1.28—1.65 μm Wavelength Range.* Electr. Lett., 1982, v. 18, No 24, pp. 1023—1024
2. V. A. Bhagavatula, M. S. Spotz, W. F. Love, D. B. Keck: *Segmented-Core Single-Mode Fibers with Low Loss and Low Dispersion.* Electr. Lett., 1983, v. 19, No 9, pp. 317—318
3. A. Majewski, H. G. Unger: *Single-Mode W-Fibers with Profile-Threshold for Low Dispersion over a Wide Spectral Range.* AEU, 1983, B. 37, H. 9/10, pp. 336—338
4. A. Majewski: *Single-Mode Optical Fibres with very Low Dispersion.* Research Reports, Warsaw University of Technology, 1985, z. 80, (in Polish)
5. A. Majewski, Z. Wawrzyniak: *Propagation Characteristics of Single-Mode Optical Fibers.* MIOP'86 Conf., Wiesbaden, 1986
6. R. Kuhne, K. Peterman, A. Majewski: *Limits for Microbending Loss in Multiple-Clad Fibers.* Proc. IOOC-ECOC'86, Venice, 1985, pp. 313—316
7. P. K. Bachmann, P. Geittner, D. Leers, H. Wilson: *Loss Reduction in Fluorine-Doped SM- and High N.A.-PCVD Fibers.* J. Light. Techn., 1986, v. LT-4, No 7, pp. 813—817
8. P. K. Bachman, D. Leers, H. Wehr, F. Weirich, D. V. Wiechert: PCVD-*DFSM Fibers*: *Performance, Limitations, Design Optimization.* Proc. IOOC-ECOC'86, Venice, 1985, pp. 197—200
9. A. Majewski: *Analiza numeryczna światłowodów wieloskokowych.* Arch. Elektr., 1984, v. XXXIII, z. 3/4
10. J. W. Fleming, D. L. Wood: *Refractive-Index Dispersion and Related Properties in Fluorine Doped Silica.* Appl. Opt., 1983, v. 22, pp. 3102—3104
11. A. Majewski: *Optical Waveguides-Theory and Design.* (in Polish), WPW, in print.
12. A. Majewski: *Numerical Analysis of Real Single-Mode Fibers,* Proc. Conf. Fibers and their Applications, Warszawa, 1986
13. H. U. Bonewitz, A. Muhlich, T. Müller, W. Lieber, W. E. Heinlein: *Quadruple-Clad Pure-Silica Core Single-Mode Fiber for Low Dispersion over Extended Spectral Range.* Proc. 9thECOC, Geneva, 1983, pp. 361—364
14. H. Etzkorn, W. E. Heinlein: *Dispersion Single-Mode Silica Fiber with Undoped Core and Three F-Doped Claddings.* Electr. Lett., 1984, v. 20, No 10, pp. 423—424

15. H. Etzkorn, W. Lieber, M. Loch, W.E. Heinlein, H.U. Bonewitz, K.F. Klein, A. Muhlich: *Dispersion Characterisrics of Broad-Band only F-Doped Single-Mode Silica Fibers*. Proc. 10th ECOC, Stuttgart, 1984, pp. 278—279
16. W. Lieber, M. Loch, H. Etzkorn, W.E. Heinlein, K.F. Klein, H.U. Bonewitz, A. Muhlich: *Three-Step-Index Strictly Single-Mode only F-Doped Silica Fibers for Broadband Low Dispersion*. Proc. IOOC-ECOC'85, Venice, 1985, pp. 201—204
17. K. Petermann: *Contrains for Fundamental-Mode Spot-Size for Broad Band Dispersion Compensated Single Mode Fibers*. Electr. Lett., 1983, v. 19, pp. 712—714
18. R. Setaka, H. Takanashi, S. Yoshida: *Fluorine Doping Levels in the Low-Pressure PCVD Process*. Tech. Dig. OFC'84, New Orleans, 1984, paper TUM2
19. G. E. Berkey: *Fluorine Fibres by the Outside Vapour Deposition Process*. Tech. Dig. OFC'84, New Orleans, 1984, paper MG3
20. K. Petermann: *Theory of Microbending Loss in Monomode Fiber with Arbitrary Refractive Index Profile*. AEU, 1976, B. 30, pp. 337—342

A. MAJEWSKI, Z. WAWRZYNIAK

CHARAKTERYSTYKI PROPAGACYJNE WIELOPŁASZCZOWYCH ŚWIATŁOWODÓW DOMIESZKOWANYCH FLUOREM

Streszczenie

W pracy rozważono jednomodowe światłowody z rdzeniem krzemowym i trzema płaszczami domieszkowanymi fluorem, przeznaczone do pracy szerokopasmowej na zakresy tzw. trzeciego okna. Wykazano, że można zaprojektować wielopłaszczowy światłowód jednomodowy przy niewielkich domieszkach fluoru i uzyskać większość z tych zalet, które wykazują światłowody z płaską charakterystyką dyspersyjną w otoczeniu 1.55 μm przy długości fali odcięcia modu HE_{12} poniżej 1.4 μ,m. Tego typu światłowody są szczególnie potrzebne do systemów z podziałem w dziedzinie długości fali WDM. gdy wykorzystuje się je w cyfrowych sieciach zintegrowanych ISDN.

A. MAJEWSKI, Z. WAWRZYNIAK

CARACTÉRISTIQUES DE PROPAGATION DES FIBRES OPTIQUES À GAINES MULTIPLES DOPÉES EN FLUOR

Résumé

Dans l'étude on considère les fibres optiques unimodales à coeur en silicium et à triples gaines avec dopage en fluor, qui sont déstinées à utiliser à bande large dans le régime de la troisième "fenêtre". On a montré, qu'il est possible de concevoir une fibre optique unimodale à gaines multiples avec dopage en fluor modéré et obtenir la majorité de ces avantages, que présentent les fibres optiques à dispersion plate centrée sur 1.55 μm pour la longuer d'onde de la coupure du mode HE_{12} au-dessus de 1.3 μm. Les fibres optiques de ce type sont nécessaires surtout pour les systèmes à partage de la longuer d'onde (WDM), lorsqu'ils sont utilisés dans les réseaux digitaux integrés (ISDN).

A. MAJEWSKI, Z. WAWRZYNIAK

CHARAKTERISTIKEN DER F-DOTIERTEN MEHRMANTEL LICHTWELLENLEITER

Zusammenfassung

Im Beitrag werden Monomodenfaser mit Quarzglasskern und drei fluordotierten Mänteln erwogen, die als Breitband LWL im Bereich des sog. dritten Fensters bestimmt sind. Es wurde nachgewiesen,

daß man ein mehrmantel LWL mit geringen Zugabe von Fluor projektieren kann, und dabei die meisten derjenigen Vorteile erreichen, die bei den LWL mit flacher Disperssionscharakteristik im Bereich von 1,55 μm und cutoff des HE_{12} — modus unter 1,4 μm gewonnen werden. LWL dieser Art eigen sich besonders für Wellelängermultiplëxsysteme (WDM), wenn sie für integrierte Digitalnetze (ISDM) genutzt werden.

А. МАЕВСКИ, З. ВАВЖИНЯК

ХАРАКТЕРИСТИКИ ПРОПАГАЦИИ МНОГООБОЛОЧЕЧНЫХ СВЕТОВОДОВ С ПРИМЕСЯМИ ФЛУОРА

Р е з ю м е

Рассмотрены одномодовые световоды с кремниевым сердечником и тремя оболочками легированными флуором предназначенные для широкодиапазонной работы т.н. третьего окна. Доказано, что можно спроектировать одномодовый многооболочечный световод при незначительном легировании флуором получая большинство положительных качеств характеризующих световоды с плоской дисперсионной характеристикой в окрестности 1,55 мкм при длине волны отсечения основного мода свыше 1,3 мкм. Световоды этого типа особенно нужны для систем с разделом в области длины волны (WDM) при использовании в цифровых интегрированных сетях (ISDN).

ROZPRAWY ELEKTROTECHNICZNE, 1988 ,34, z. 4, ss. 1159-1175

535.2

Laser Measurement 1968—1987 and beyond

WIESŁAW WOLIŃSKI

Instytut Mikroelektroniki i Optoelektroniki, Politechnika Warszawska

Received 1987.09.16

Authorized 1988.04.20

The report[1], being a general introduction to laser metrology, presents examples of laser application to tracing straight lines, to determining angular deviations and the reference plane, in measurements of distance, displacement and distortion, to measurements of atom and particle spectra, and for the construction of length and frequency standards. These examples display not only the achieved level of measuring methods but also the progress reached during the last 20 years.

INTRODUCTION

In 1968 the URSI Conference on Laser Measurement was held in Warsaw. Even at this time, not so long after the invention of laser radiation, (Maiman 1960), the far-reaching role to be played by the latter in different field of measurement was fully appreciated. When looking over the proceedings of this Conference (Smoliński, Hahn, 1969) we find many very interesting research results on: the development of optical frequency standards, the application of lasers in interferometry at a large difference of the optical path and of optical radar, laser measurements of plasma parameters and the measurement of element size by means of diffraction picture analysis as well as much work on the properties of lasers themselves. Neverheless, the 20 years which have elapsed since, raised the level, expanded the range and, in the first place, spread the application of laser metrological equipment.

Now, passing to the survey illustrating the progress in this field, I would like to recall briefly the properties of laser radiation from which the possibilities of its application in metrology result directly.

Laser radiation is emitted in a single direction, determined by the optical resonator, the laser beam having a very small divergence angle (of the order of several milliradians). The whole power, emitted within this small solid angle, gives a high luminance of the source (W/m^2sr) higher, by some orders of magnitude, than that of sources emitting

[1] Tutorial lecture presented during the XXII URSI General Assembly, Tel Aviv, Israel, August 24 — September 2, 1987.

light in conformity with Lambert's law. For the fundamental TEMooq mode, the beam intensity distribution is displayed by the Gaussian curve. This beam, by means of very simple optic systems, can be collimated still better and focused (on small surfaces of a diameter of several to several tens of micrometers) reaching large power and energy densities, and, this, evoke nonlinear phenomena such as light harmonics, autocollimation and bistability, not known before the invention of lasers.

Laser-radiation is, to a high degree, monochromatic, having a small spectral width. According to the well known equation of Schawlow and Townes, the limiting spectral width of the operation of a laser in a single mode can be determined from:

$$\Delta\nu_L \simeq \frac{2\pi h\nu(\Delta\nu_{\text{pass}})^2}{P}$$

where: $h\nu$-quantum energy, $\Delta\nu_{\text{pass}}$ — half width of the resonance curve of a passive resonator, P — laser output power.

For instance, for He-Ne lasers: $\lambda = 632.8$ nm, $\nu = 4.73\ 10^{14}$ Hz, $\Delta\nu_{\text{pass}} = 0.5$ MHz, $P = 1$ nW, the value of determined from this equation mounts to $5 \cdot 10^{-4}$ Hz. Actually, different destabilising factors extend this width by several orders of magnitude. In specially designed and electronically controlled lasers, a short-term line-width of about 1 Hz is reached. Solid lasers, as compared with gas ones, have an inferior monochromaticism.

The laser radiation is coherent in time and space, i.e. it shows a perfect interference capability. The two beams into which a single beam is divided and which pass two different optical paths interfere with one another. The difference of these optical paths or the so-called coherence path $L = c\Delta t$, (where c — light velocity and the coherence time $\Delta t \simeq 1/\Delta\nu_L$) may reach many kilometers. Similarly two beams emitted from different points of the source and having passed different optical paths will interfere. From among all classical sources, only low pressure and small current discharge tubes, emitting spectra line groups of widths resulting from the Doppler effect, show a time coherence, reduced to the difference of optical paths of about 80 cm. The above mentioned features as well as the possibility of choosing lasers operating in the range from soft X-rays to far infrared (several hundred micrometers) and with continuous duty or pulse duty even down to femtosecond ranges, the tunability of the wavelength of some lasers over a wide spectral range and the possibility to obtain large powers and energies, means that it is difficult, today, to find a field of science and technology where lasers are not applied, in particular, in metrology. Lasers are applied in the following main fields of metrology:

- geodesic measurements — of straight line, angle deviation, reference plane and distance;
- interference measurements — of quality of transparent media, displacement, distortion of small surfaces, dynamic distortion halography, seismic disturbance, size of elements by analysing diffraction pictures and surface state, etc.
- spectroscopic measurements — by absorption methods or by methods based on nonlinear phenomena and the application of spectroscopic methods for: construction of length and frequency standards, detection and discrimination of chemical compounds (methane, HpD, isotopes), analysis of the surface state and

contamination of media, measurement of the flow of gases and liquids and of rotational speed, etc.

Each of the above mentioned fields covers many different procedures and many different designs of the measuring equipment. In the following part of this report, I will limit myself to a brief consideration of topics which, from the aspect of wide application and progress of science and technology, contributed new values and played an essential role.

GEODESIC MEASUREMENTS

The application of lasers in geodesic measurement allowed for:

- carrying out measurements in conditions of limited visibility and at points with difficult access
- reduction of the duration of measuring procedures accompanied by a raised measuring accuracy;
- automation of geodesic works such as:
  - testing of the distortion of hydrotechnical structures, bridges, viaducts, crane tracks, crane guides, shaft hoists, etc.
  - levelling of longitudinal and transverse profiles,
  - carrying out works by means of the tunnel method and in trenches as well as laying of pipelines,
  - multi-storey building, high accuracy alignment of components of large machines,
  - control of the operation of all kinds of working machines (dredgers, bulldozers, etc),
  - standardization of geodesic optical instruments,
  - satellite bearing.

In all the above mentioned applications, excepting the last one, laser levelling instruments, theodolites, interscans, plumbers, and rotational equipment, are applied. In these equipments, most often provided with automatic levelling and plumbing, the measuring axis or plane is determined by a laser radiation beam.

The equipment operates with optoelectronic auxiliary devices determining precisely the centre of the incident beam (eg. quadrant detectors placed on levelling staffs) or maintaining their position within the beam or measuring plane, recording simultaneously deviations appearing in time.

All the above mentioned equipment are provided with integrated He-Ne lasers emitting a wavelength $\lambda = 632.8$ nm of a power ranging from 0.5 to 5 mW in the TEMooq fundamental mode.

When setting the spot position on the staff by eye at a distance of 100 m a precision from 1.3 to 1.8 mm is obtained for a beam focused to 12 mm dia. When quadrant detectors are used, this precision of the determination of the axis rises to 0.2 mm. At distances up to 20 m with good focusing of the beam and using a CCD mosaic detector a measuring precision of the position of the beam axis reaches about 0.05 mm.

The determination of a plane by means of a rotating beam does not afford good

precision. The latter is ± 20 mm at a distance of 250 m. This equipment, however, allows for the simultaneous levelling at many points of an area of this radius or a simple control of a bulldozer levelling the area.

The determination of the plumb-line (perpendicular) is carried out by means of a compensator diffracting the horizontal laser beam through a right angle and maintaining it in this direction either by the coincidence of the output beam with that reflected from the mercury surface mirror or by a design of the equipment, using the gravitation force to maintain the equipment and the beam perpendicular. The latter method allows for determining the beam axis with a precission ± 10 mm at a distance of 250 m.

In geodesy engineering practice the necessity arises to determine the space coordinates of structures, e.g. of a tunnel. Tachymeters are applied for this purpose. After projecting the laser point upon the structure, the horizontal and the vertical angles are measured with a precision of 1 angular minute. The distance is measured by the principle of a variable base of the instrument and a constant parallactic angle at the aim point. These instruments, intended for the measurement of small distances up to 60 m, assure a precision of distance measurement of ± 1 cm.

The development of geodesic methods and equipment was aimed, not only at raising their accuracy, but at providing them with microprocessor data-processing systems, the extension of their measuring capability and the construction of specialized systems. For instance, the equipment for determining planes, provided with detectors and a timer measuring the time interval between pulses obtained from the reflection of the rotating beam from appropriately set reflectors, also measures simultaneously the angle with an accuracy up to several angular seconds (the head turning at a speed of 600—60,000 rpm controlled with a quartz generator). Another specialized equipment controls a tunnel cutting disk. This equipment not only leads the disk along a straight line but is also able to deviate it in a programmed manner from the latter.

Distance measurement from some hundred meters up to some scores of kilometers (80 km at earth level and pure air) is carried out using two methods. The first consists in the modulation of the beam of a continuous duty He-Ne laser ($\lambda = 632.8$ nm) with a frequency 0.1 — 40 MHz and after its reflection either from the retroreflectors or from the natural object, the precise measurement of the phase shift $2\pi\Phi$ The approximate expressions for the distance and the measuring error are as follows (Holejko, 1981):

$$L = \frac{\lambda_s}{2}(n+\Phi) = \frac{v}{2f_s}(n+\Phi),$$

$$\Delta L = \left[\frac{\Delta v}{v} + \frac{\Delta f_s}{f_s}\right]L + \frac{\lambda_s}{2}\Delta\Phi,$$

where: L — measured distance: λ_s, f_s = length and frequency of standard wave; n = = total number of waves λ_s in $2L$; v = wave velocity.

It can be seen that one component of the error depends upon the distance and the second is constant, depending on the length of the standard wave and the measuring accuracy of the phase shift. Carrying out measurements of the same distance at different frequencies of the standard wave, the absolute value of this distance can be determined. The manufacturers of this kind of equipment assess the measuring error either as $\Delta L =$

$= \pm 1$ mm or $10^{-6} \times L$ for reflection from a retroreflector or as $L = \pm 3$ mm or $10^{-4} \times L$ for a diffused reflection. The error adopted is the higher value.

The second method is analogous to that used in microwave radars. The emitted light pulse, after its reflection, returns to the receiver stopping the timer measuring the time needed for passing the double path. For this purpose, NdYAG pulsed lasers with Q-switched optical resonators are used most often. For distances of the order of 1.5 km, semiconductor lasers are used. In special equipment, because of the atmospheric window, pulsed molecular CO_2 lasers emitting a wave of 10.6 μm, are applied. As an example, an anti-tank laser telemeter can have a 5 km range and a measuring accuracy of 3 m, where the source consists of a pulsed CO_2 laser TEA type emitting 60 ns pulses. Two pulses assure the measurement. Ranges of 1500 m at a visibility of 500 m have been demonstrated with the use of this equipment.

Another group comprises equipment intended for geodesic and geophysical satellite research. By measuring the distances earth-moon and earth-artificial earth satellites, not only can the position of the station be determined, but also data on the earth's rotation, the gravity field, shifts of poles and the dynamics of oceans can be obtained. The eqiupment operators on the same principles as those described above but for the fact that the pulse is reflected from a set of retroreflectors provided on the satellite comprising from ten to several hundred reflecting elements. The satellites are placed on different orbits of an apogee from 500 to 2000 km. The measuring accuracy is determined by the standard deviation. For symmetrical pulses this standard deviation of the distance measurement is determined by the expression (Szydlak 1985).

$$\sigma_R = A \cdot \frac{T}{\sqrt{n \cdot M}}$$

where: A = proportionality factor, T = duration of generated pulse, n = number of received photons, N = number of measurements per second.

Hence results the requirement for laser emitters used in this kind of equipment. They are to generate pulses of large energy, short duration and a (possibly) high recurrence rate. In order to ensure that a maximum number of photons reach the satellite, a beam with a minimum deviation angle (10^{-5} rd) is to be produced. At first ruby-crystal lasers (wavelength 694.3 nm) with a Q-switched optical reconator producing pulses of energy of about 1J, width 20—30 ns, repeated at a rate of 0.1 Hz were used. In spite of large diameter emitter and receiver devices, more than 2.5 m, the accuracy was only 1 m. The second generation of this equipment used pulsed ruby-crystal lasers of pulse width from 2 to 5 ns, with a Gaussian radiation distribution and a repetition rate of 1 Hz. A measuring accuracy, eg of the distance to the satelilte LAGEOS (apogee 5,900 km) of about 10 cm was reached. Equipment of the third generation used NdYAG lasers emitting waves of length $\lambda = 1{,}06$ μm converted to its 2nd harmonic $\lambda = 530$ nm to which the detectors used are very sensible. The lasers operate, mostly, in a mode synchronization system with selection and amplification of pulses of width 20 to 200 ps. These lasers may also be ones with active modulation of the Q-factor and time-variable losses of the resonator. Besides the considerable shortening of pulses, their repetition rate can be raised even up to 20 Hz. The main phenomena, rendering measurement difficult and bringing

about the necessity of introducing many corrections, are: attenuation, dispersion and refraction of the radiation, curving the beam path. Equipments similar to that described above are the socalled lidars of range 300 m to 15 km used for measuring the dynamics of clouds, smoke, dust, etc.

INTERFERENCE MEASUREMENTS

Within the group of interference measurements a very important role is to be attributed to equipment allowing for the determination of displacements with an accuracy up to a fraction of the light wavelength (Dukes, Gordon, 1970), (Baldwin et al. 1971), (Holejko 1981). Two measuring methods are used. One is based on the direct counting of interference striae accompanying the displacement of the mirror constituting one of the arms of a Michelson interferometer (Fig. 1). The radiation source used is a He-Ne

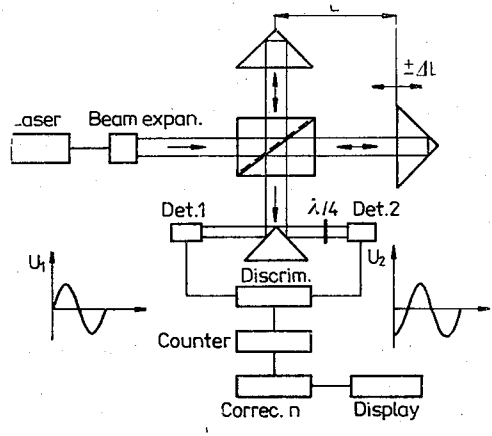


Fig. 1. Layout of a single-frequency interferometer for measuring displacement

laser operating at a single frequency maintained with a stability better than 5 MHz per day (frequency instability to about 10^{-8}). To determine the direction of changes, the image of the striae is divided into two beams, with a quarter-wave plate inserted in the path of one of them. The signals obtained from the detectors have phases shifted by 90°, the discriminator comparing both pulse series brings about the addition or subtraction of the number of pulses displayed on the counter. In most cases this equipment affords an accuracy of $\lambda/4$, i.e. for the laser used, about 0.15 μm. An equipment based on the use of the Doppler mirror effect (Fig. 2) is also used. The He-Ne laser of frequency instability 10^{-9} generates, due to the Zeeman effect in the transition $3s_2 \rightarrow 2p_4$, two frequencies f_1 and f_2. In order to facilitate their discrimination, these beams have mutually perpendicular linear polarisation. The reference is provided by the frequency difference $(f_2 - f_1)$ and the signal, taking into account the Doppler effect produced by the mirror moving with a velocity v, is the frequency difference $f_2 - (f_1 \pm \Delta f)$. The difference of these two signals gives the Doppler frequency $\Delta f = \pm 2f_1(v/v_L)$ where v_L = light velocity in the

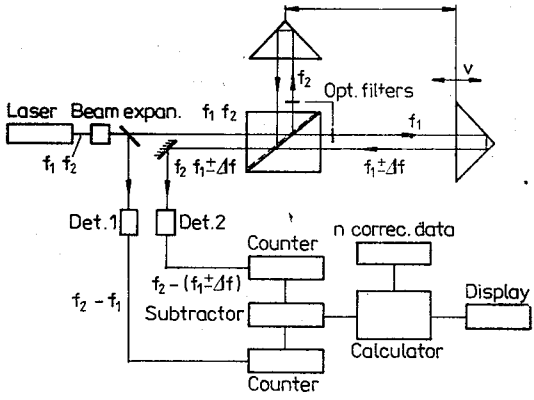


Fig. 2. Layout of a double-frequency interferometer for measuring displacement

measuring conditions. The displacement can be determined from the following expression:

$$\Delta L = \int_0^T v\,dt = \pm\frac{\lambda}{2}\int_0^T \Delta f\,dt$$

To increase the resolution, resolution expanders multiplying Δf are applied additionally behind the detectors. A measuring resolution of the order 10^{-2} μm is obtained for a distance up to 60 m. The series connection of expanders allows for raising the resolving power by one order. The accuracy of the displacement measurement is limited by the accuracy of the frequency measurement and amounts to 10^{-7}. These interferometers are controlled automatically by microprocessors, taking into account corrections due to the variation of humidity, pressure and temperature of the air.

Equipment used for the precise measurement of displacement allowing for the automation and increased precision of many manufacturing processes, play an outstanding role in technology today. For instance, according to Laser Focus (p. 21 March 1986), equipment for displacement measurements constitutes 57.3% of the total number of laser equipment sold for microelectronics for a sum of $1255 million in 1985. A rise in sales from $719 million up to $1220 million is foreseen in 1990.

Another group of measurements developed since 1978 concerns holographic interferometry. It allows for touchless nondestructive quality control of products and measurement of the shape of objects, measurement of small displacements, deformations and vibrations, detection of noise sources in equipment and of the paths along which noise is transmitted. Holographic interferometry of applied in tests of cool and hot plasma, research on diffusion, crystallization, shock-waves, flow patterns, etc. The gratest advantage of these tests is the fact that they can cover the whole of a large object, their accuracy being of the order of a fraction of a micrometer. This measuring technique is of particular importance in the case of extreme reactions, e.g. in rocket, aircraft and motor industries. The measurements are carried out by taking a hologram of the tested object and then subjecting it to the action of specified forces, a second exposure is made

on the same photographic plate. Interference of the two images displays all deformations and defects. This method allows, moreover, to test the dynamics of stresses by taking a hologram with two high-power laser pulses quickly following one another. For this kind of measurement generally two-pulse ruby lasers (λ = 693.3 nm) generating pulses of an energy of the order of one tenth Joule and length from 5 to 50 ns with a precisely controlled interval between them ranging from 100 to 400 ms, are used. These laser heads require correction of the radiation coherence. Also NdYAG pulsed lasers with 2-nd harmonic transformation (λ = 530 nm) can be used for this purpose.

How surprising and novel the applications of this technique may be illustrated by the holographic interferogram of the vibration of the human throat and face when singing (Fig. 3) (Pawluczyk 1980). It can be used for the assessment of the professional

Fig. 3. Holographic interferogram of vibrations of the surface of throat and face

ability of vocalists. Today holographic interferograms are used for qualitative tests. Nevertheless technical literature provides some information on the computer processing of holographic interferograms, furnishing data on the mechanical and operational features of the objects tested.

One of the main phenomena limiting resolution power and simulation contrast is the unwanted diffraction and interference effect, known as coherent noise, appearing in the laser light. An interesting and very efficient method for the elimination of this noise is suggested in the reports (Mroz et al 1971), (Pawluczyk Mroz 1973). It consists in changing the direction of incidence of the beam lighting up to object by means of a rotating glass cube whose axis of rotation is perpendicular to the beam axis. Above a certain velocity

of rotation of this cube, the grain images shifted along the interference striae, and those whose sources lie outside the plane of the image, are blurred. This is the so-called single-direction time-averaging method of coherent noise. This method has been successfully applied in the development of holographical and holographic-interference microscopes. Fig 4 represents two holograms of a biological preparation without and with attenuation of the coherent noise (Pluta, 1980).

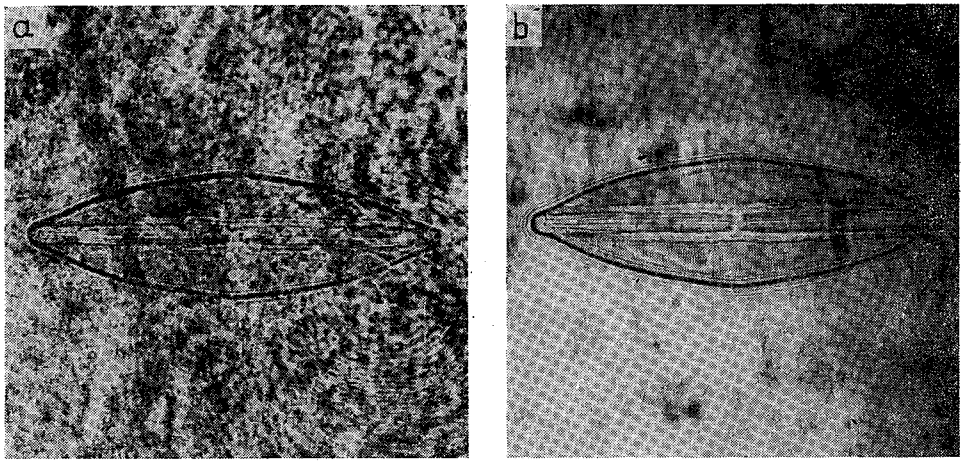


Fig. 4. Holographic pictures of a diatom (bacillariophycea) recorded by microscopy: a) without attenuation of coherent noise, b) with attenuation of coherent noise

SPECTROSCOPIC MEASUREMENTS

Spectroscopy is a field whose development is to be particularly attributed to the application of lasers, mainly of tune lasers among which the first place is occupied by dye lasers ($\lambda = 0.3 \ldots 1$ μm) of continuous and pulsed duty. Besides the latter, tuned excimer lasers ($\lambda = 0.12 \ldots 0.32$ μm), lasers with coloured centres ($\lambda = 0.7 \ldots 3$ μm), Raman spin-flip type lasers ($\lambda = 4 \ldots 5.5$ μm and $\lambda = 8 \ldots 12$ μm), solid-state laser ($\lambda = 0.5 \ldots 30$ μm), parametric oscillators ($\lambda = 0.5 \ldots 4$ μm) and nonlinear frequency mixing systems ($\lambda = 35$ nm ... 100 μm) are also used. The application of lasers to the well-known absorption methods has permitted simplification of the measuring systems and also raised the resolution and sensitivity of methods. However, a qualitative jump in this field is due to nonlinear spectroscopy characterized, in the first place, by the elimination of Doppler line broadening. Saturation spectroscopy (Letochow, Czebotajew, 1982), (Gawlik 1985) is the basic method. Fig 5 provides a diagram of the measuring system. Two beams, the saturating one and the sampling one of pulsation ω are propagated in the test medium in opposite directions. An atom moving with a velocity component V_z along the propagation direction of the radiation sees a beam of pulsation $\omega(1-V_z/c)$ due to the Doppler effect. The laser beam reacts selectively with atoms of such V_z that a resonance occurs between the atomic transition, pulsation ω_0 and the beam pulsation in the moving system, i.e., $\omega(1-V_{z/c}) = \omega_0$.

Hence, $V_z = V_{res} = c(\omega - \omega_0)/\omega_0$ and on the curve of the population of the upper level, a peak appears at the beam's V_z. The sampling beam of a low radiation intensity, not changing the distribution of the population of levels $N(V_z)$ and reacting with the group of atoms of $(-V_{res})$, undergoes attenuation after passing through the medium. The mea-

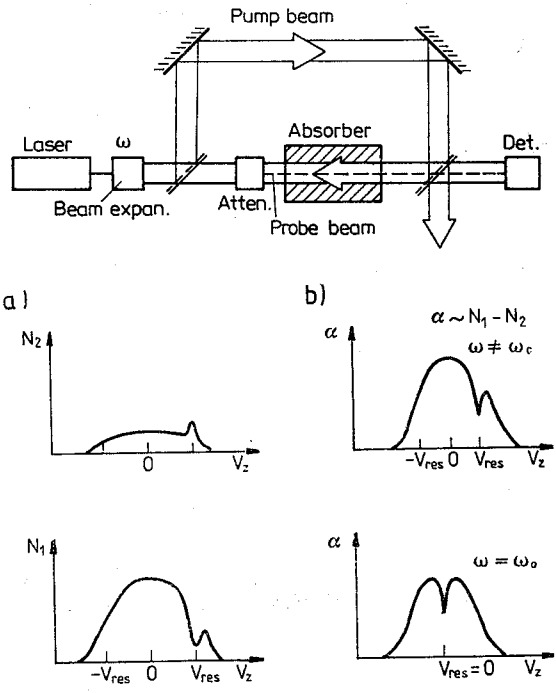


Fig. 5. Layout of equipment for saturation spectroscopy: a) population of lower (N_1) and upper (N_2) energy levels of atoms versus their velocity component conform with the direction of radiation propagation v, under reaction of a strong laser beam of close to $_0$; b) absorption factor versus v_z for close to $_0$ and $= _0$

sured radiation intensity is a measure of the absorption factor of the medium. When the laser is tuned strictly to resonate with the tested transition $\omega = \omega_0$ and when $V_{res} = 0$ both beams react with the same group of atoms. The feeble beam samples the medium saturated already by the strong beam and is less attenuated after passing then beyond resonance, displaying a narrow peak. Doppler broadening is eliminated by the selection of a narrow velocity class.

During the last decade polarization spectroscopy has been developed, which is certain kind of saturation spectroscopy. The difference consists in the fact that the sampling beam detects not the change of the saturated absorption factor but the change of the dispersion properties of the medium caused by the saturating beam (Wieman, Hansch, 1976), (Gawlik, Series, 1979). This allows for the elimination of the Doppler-broadened background and only narrow lines are recorded. The amplitude of the signal can be increased by beating it with the sampling beam. This brings about, at the same time, a change from the Lorentz shape to a dispersion one which can be used, for instance, as a signal for a very precise stabilization of the laser. The progress achieved in measuring resolution is illustrated by the curves reported (Hansch et al ... 1979) for a line structure α, a wave-

length $\lambda = 656$ nm, Balmer's series of a hydrogen atom. The resolution power of these spectra is so high that it displays the hyperfine structure of lines, allowing for such precise measurement of Rydberg's constant ($R = 2\pi^2 me^4/h^2$) that it is one of the most accurately determined physical constants known ($R = 10973731.521$ (ll) m^{-1}), (Amin et al 1981).

Besides the above mentioned measurements, many other spectroscopic methods raising the resolution and sensitivity and bringing about the possibility of investigating selected atoms or levels have been developed. These methods comprise: two-photon spectroscopy, velocity-selective optical pumping spectroscopy (V.S.O.P.), fast ion beam spectroscopy, labelling spectroscopy, spectroscopy with ultrasonic molecular beam, etc. I want only to call your attention to the two topics which will influence the future development in this field. The first is the application as detectors of lines of photodiodes or CCD matrices allowing for a spatial reproduction of the whole spectrum image. Matrices are often coupled with opto-electric image converters and amplifiers of the microchannel plate type, allowing to obtain a signal in each channel of a sensitivity corresponding to the detection of single photons. This renders possible a multichannel reproduction of the spectrum which is essential for the investigation of dynamic processes so short that, until now, they were impossible to record. The second topic concerns the study in the range of subnatural spetroscopy reaching beyond the natural line width resulting from

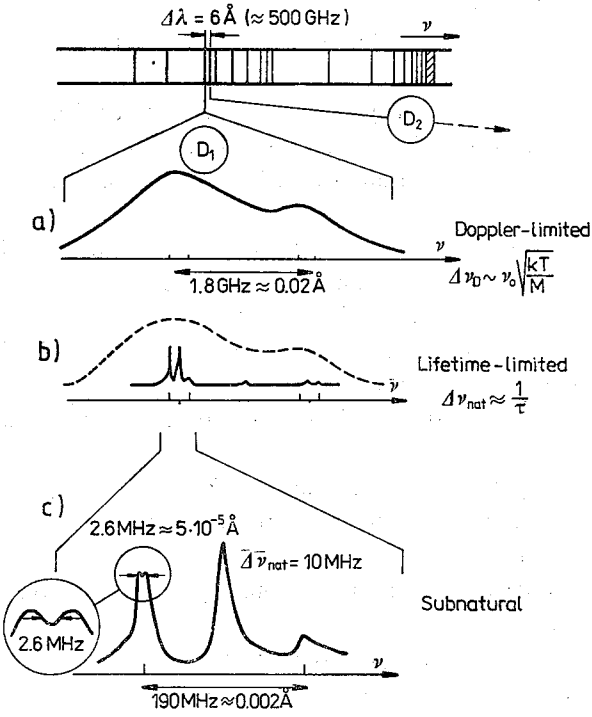


Fig. 6. Set of measurements of line D_1 sodium Na: a) spectral profil of line D_1 with account taken of the hypersubtle structure and the Doppler broadening; b) the „Doppler-less" method of polarization spectroscopy; c) fragment of the spectrum recorded by new method

spontaneous emission energy levels. From among the methods quoted in literature, I would like to call attention to the method using nonstationary optical pumping effects (Gawlik et al ... 1982), (Gawlik, 1986). Firstly, the narrow structures could be attributed to Zeeman coherences which are induced by the probe beam revealed by differential light shifts that remove the degeneracy of the magnetic sublevels. Secondly, nonstationary optical pumping into a level that cannot be excited further by any of the two laser beams. This method allows to investigate atoms and molecules in the gaseous phase. Some of these molecules, being subject to optical pumping, i.e. those having a suitable structure and a long enough lifetime of their lower energy levels. Measurements are made by means of equipment similar to that used in polarization spectroscopy but provided with additional adjustable diaphragms and attenuators. The difference of these measurements as compared with the polarization method consists in the fact that the probe beam has a high radiation intensity and, similarly to the pumping beam, strongly disturbs the medium. Fig. 6 represents test results for the line D1 of Na atoms. Curve (a) is obtained by the polarization method. Curve (b) represenets a fragment of the spectrum obtained using the new method, representing the hyperfine components $3S_{1/2}$ $(F = 2) \rightarrow 3P_{1/2}(F = 1$ and $F = 2)$ with an additional resonance of cross-over type in the centre between them. An indentation of 2.6 MHz half-wave width and much narrower than the natural width of the line D_1 amounting to 10 MHz, is visible in the centre of one of the component lines. The simplicity of the system together with its subnatural accuracy and high sensitivity onstitutes the undoubted advantage of the method.

Spectroscopy is responsible for important achievements in the field of fundamental research, such as:

- — statement of maintained parity in atomic spectra; verification of the theory of the unification of reactions (Forston, Cets 1980), (Stacey, 1984);
- — direct observation of quantum transition (Aspect et al, 1982), (Perrie et al ... 1985);
- — determination of new value of the electron radius (Dehmelt 1985);
- — observation of the modification of spontaneous emission through a resonant cavity (Gabrielse, Dehmelt, 1985), (Hulet et al ... 1985);
- — observation of new states of the electromagnetic field of minimum fluctuations, known as squeezed states (Wu et al ... 1986);

and many many others. Also the application of spectroscopy in different fields is very wide, for instance, for:

- — analysis and diagnostics of media and surfaces;
- — verification and control of chemical reactions (photochemistry, photomedicine, separation of isotopes)
- — ultra-sensitive detection of radiation, mainly of microwave type (by means of Rydberg atoms);
- — construction of gyroscopes and flowmeters;
- — construction of new length and time standards.

Let us look closer at the last group of applications mentioned.

Gas lasers, such as He-Ne and CO_2 ones, because of the small spectral width of their radiation ($\Delta\nu$), are much better as length standards than the standard established in 1960

in the form of a Kr^{86} discharge tube emitting for the transition $2p_{10} \rightarrow 5d_5$, a wavelength 605.7802105 nm. Moreover, the coherent laser radiation, easily treated by heterodyne procedures, allows for realization of a frequency standard. A laser suitably stabilized by means of external negative feedback systems may, at the same time, be used as a length and frequency standard. Dip Lamba stabilized He-Ne lasers, or lasers using for stabilization the Zeeman splitting effect of the line $\lambda = 632.8$ nm, assure an instability of frequency of the order 5×10^{-8} and serve as length standards in interference measurements. A quantum standard comparable with the caesium (Cs^{133}) standard (frequency instability $(1 \ldots 2) \times 10^{-13}$, $\nu = 9192631770$ Hz) have been developed stabilizing the laser radiation frequency by means of the saturable absorption phenomenon. Reports on this topic were presented at the URSI conference of 1968. At this time a frequency instability of $10^{-9} \ldots 10^{-11}$ was obtained, i.e. from four to two orders of magnitude worse than that reached today. An atomic or molecular standard in the form of an absorption cell placed, generally, inside the laser resonator must fulfil several conditions. The absorption line of the standard must coincide with the laser emission line and the transition frequency must be constant and reproducible at best with an accuracy specified for a quantum standard. The latter condition is fulfilled by many different transitions between undisturbed electric and magnetic fields and collisions, energy levels of quantum systems. Moreover, taking into account the possibility of a servosystem used for stabilization with reference to the centre of the resonance curve, the relative width of the resonance curve of the standard must not exceed the desired magnitude of the frequency instability more than $10^3 \ldots 10^4$ times. At 10^{-13} this relative width should be of the order $10^{-9} \ldots 10^{-10}$, i.e., several orders of magnitude narrower than brought about by Doppler broadening ($10^{-5} \ldots 10^{-6}$). A resonance minimum appears in the case of an absorption cell placed inside the laser resonator when the generated frequency is tuned to the frequency of the absorption curve centre. In consequence, a resonance maximum used for stabilization is in the middle of the laser output power curve plotted versus frequency. A certain fre-

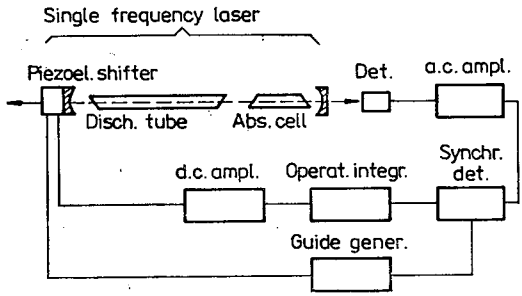


Fig. 7. Frequency stabilization system of laser with internal absorption cell

quency shift occurs between the maximum power peak and the middle of the absorption curve due to the phenomenon of frequency pulling. This is the method used to stabilize a He-Ne laser generating a radiation of wavelength $\lambda = 3.39$ μm and provided with a CH_4 absorption cell. In the spectral line P7 of the oscillatory-rotating band ν_3 of methane having a lifetime of about 0.01 s, a narrow resonance minimum is obtained. The peak

width on the laser output power curve amounts to 300 kHz. Using synchronous detection, a discrimination curve of a very steep slope and a high signal to noise ratio is obtained. The application of a conventional stabilization loop keeps the laser frequency within the peak centre (Fig. 7). A frequency instability of 1×10^{-13} for averaging times from several seconds and longer and a reproductibility of 10^{-11} have been obtained in this system (Barger, Hall, 1969). Because the wavelength lies in the range of visible light, the application of a He-Ne laser $\lambda = 632{,}8$ nm is attractive. In this case iodine J^{127} or J^{129} absorption cells are used. The iodine lines coincide precisely with the centre of the Doppler curve of the line $\lambda = 632.8$ nm of the laser, requiring more complex stabilization systems consisting in the synchronous detection of the third derivative of the auxiliary modulation signal. In this way, a quantum standard of frequency instability from 3×10^{-12} to 1×10^{-11} in an averaging time 1 ... 1000 s and a reproductibility of one order worse, i.e., 4×10^{-11} to 1×10^{-10} (Quenelle, Wuerz, 1983) is obtained.

For the stabilization of the operating frequency of lasers, there is also a method according to which the absorption cell is places outside the laser resonator. This is typical for saturation spectroscopy. A strong beam saturates the medium and the feeble probe beam is reflected from the mirror and propagated in the opposite direction. This kind of equipment needs a high laser output power and high absorption by the cell. However, it has the advantage that the frequency pulling phenomenon does not appear. A typical equipment comprises a CO_2 laser with a SF_6 cell. In this case, a power peak having a width close to the half-width of the natural absorption line is obtained. The above equipment gives a frequency instability of 10^{-12} (Ouhayoun, Borde, 1977) to be reached.

The above brief survey of laser applications, selected from the three main groups of laser applications in metrology, shows that the laser has become an indispensable tool in all fields of technology and science. It not only facilitates and increases the accuracy of methods using well-known principles, but gives access to quite novel possibilities. Metrology is developing along new lines and now applies new methods exerting an essential influence on the worldwide level of science and technology.

ACKNOWLEDGEMENTS

Last but not least I would like to thank Mr W. Gawlik as well as Mr. S. Pachuta and Mr R. Nowicki for the materials sent to me which helped, to a large extent, in the preparation of this report.

REFERENCES

1. S. R. Amin, C. D. Caldwell, W. Lichten: *Crossed-beam spectroscopy: A new value for the Rydberg constant.* 1981. Phys. Rev. Lett. 47, 1234—1238
2. A. Aspect, P. Grangier, G. Roger,: *Experimental test of realistic local theories via Bell's theorem.* Phys. Rev. Lett. 1981, 47, 460—463
3. A. Aspect, J. Dalibard, G. Roger: *Experimental tests of Bell's inequalities using time-varying analyzers.* Phys. Rev. Lett. 1982, 49, 1804—1807

4. R. R. Baldwin, G. B. Gordon, A. F. Rude: *Remote laser interferometry.* Hewlett-Packard Journal, 1971, 12, 14—20
5. R. L. Berger, J. I. Hall: *Pressure shift and broadening of methane line at 3.39 m studied by laser saturated molecular absorption.* Phys. Rev. Lett. 1969, 22, No 1, 4—8
6. H. Dehmelt: *Single atomic particle at rest in free space: new value for electron radius.* Ann. Phys. Fr. 1985, 10, 777—795
7. J. N. Dukes, G. B. Gordon: *A two hundred foot yardstick with graduation every microinch.* Hewlett-Packard Journal, 1970, 8, 2—8
8. E. N. Forston, L. Wilets: Adv. At. Molec. Phys. 1980, 16, 319
9. G. Gabrielse, H. Dehmelt: *Observation of inhibited spontaneous emission.* Phys. Rev. Lett. 1985, 55, 67—70
10. W. Gawlik: *Laser spectroscopy.* I-st Symposium of Laser Technology (eds W. Woliński et al 267—286, ZGPW, Warsaw (in Polish), 1985
11. W. Hansch, A. L. Schawlow, G. W. Series: *The spectrum of atomic hydrogen.* Scientific American, 1979, 240, 72—86.
12. K. Holejko: *Precision Electronic Distance and Angle Measuring Instruments,* 43, 168—171. Technical Publishing House, Warsaw (in Polish), 1981
13. R. G. Hulet, E. S. Hilfer, D. Kleppner: *Inhibited spontaneous emission by a Rydberg atom.* Phys. Rev. Lett. 1985, 15, 2137—2140
14. W. S. Letochov, W. P. Czebotaev: *Nonlinear Laser Spectroscopy 26—39.* Technical Publishing House, Warsaw, (in Polish), 1982
15. E. Mroz, R. Pawluczyk, M. Pluta: *A method for coherent optical noise elimination in optical systems with laser illumination.* Optica Applicata. 1971, 1, No 2, 9—15
16. O. Ouhayoun, C. J. Borde: *Frequency stabilitzaion of CO_2 lasers through saturated absorption in SF_6 .* Metrology 1977, 13, 149—150
17. R. Pawluczyk, E. Mroz: *Unindirectional optical coherent noise elimination in optical systems with laser illumination.* Optica Acta 1937, 20, 379—386
18. R. Pawluczyk: *Holographic interferometry of opaque object.* In Optical Holographv (ed. M. Pluta) 205. Polish Scientific Publishers, Warsaw (in Polish), 1980
19. W. Pierrie, A. J. Duncan, H. J. Beyer, H. Kleinpoppen : *Polarization correlation of the two photons emitted by metastable atomic deuterium: a test of Bell's inequality.* Phys. Rev. Lett. 1985, 54, 1790—1793, 2647
20. M. Pluta: *Holographic microscopy and interferometry.* Optical Holography, (ed. M. Pluta) 515. Polish Scientific Publishers. Warsaw (in Polish), 1980
21. R. C. Quenelle, L. J. Wuerz: *A new microcomputer-controlled laser dimensional measurement analysis system.* Hewlett-Packard Journal 1983, 4, 3—7
22. A. Smoliński, S. Hahn, (inv. eds.): *Proceedings of the URSI Conference on Laser Measurement.* Electron Technology (ed. B. Paszkowski) 2, 2/3. Polish Scientific Publishers, Warsaw, 1969
23. D. N. Stacey: *Experiments on parity non-conservation in atoms.* Acta. Phys. Polonica, 1984, A66. 377—397
24. J. Szydlak: *Application of lasers to satelitte ranging.* 1st Symposium of Laser Technology (eds W. Woliński et al) 287—308, ZGPW, Warsaw (in Polish), 1985
25. C. Wieman, T. W. Hansch: *Doppler-free laser polarization spectroscopy.* Phys. Rev. Lett. 1976, 36, 1170—1173
26. L. A. Wu, H. J. Kimble, J. L. Hall, H. Wu: *Generation of squeezed states by parametric down conversion.* Phys. Rev. Lett. 1986, 57, 2520—2523

W. WOLIŃSKI

POMIARY LASEROWE W LATACH 1968—1987 I DALSZE PERSPEKTYWY

Streszczenie

W artykule przedstawiono grupy zastosowań wykorzystujące takie właściwości promieniowania laserowego jak wysoką monochromatyczność, zdolność do interferencji, przy dużych różnicach dróg optycznych, kierunkowość i mały kąt rozbieżności wiązki oraz możliwości uzyskania pracy ciągłej, impulsowej a także impulsowej o dużej częstotliwości powtarzania impulsów i przestrajania długości fali. Omówiono zastosowania laserów do wytyczania linii prostej, odchyleń kątowych, płaszczyzny odniesienia, pomiarów odległości, przemieszczeń i odkształceń, pomiarów widm atomowych i cząsteczkowych oraz budowy wzorców długości i częstotliwości. Przykłady prezentują nie tylko poziom metod pomiarowych ale również postęp jaki dokonał się na przestrzeni ostatnich 20 lat.

W. WOLIŃSKI

MESURES DES LASERS PRISES DE 1968 A 1987 ET REALISATIONS EN PERSPECTIVES

Résumé

Dans l'article on a présenté les groupes des applications mettent à profit de telles propriétés du rayonnement de laser, que haute monochrominance, pouvoir interférentiel pour les grandes différences des chemins optiques, directivité et petit angle de divergence d'un faisceau ainsi que possibilité d'ontenir un travail continu à impulsions, et aussi à impulsions à haute fréquence de répétitions des impulsions et de transpositions de la longueur d'onde. On a considéré l'application des lasers à la désignation de la ligne droite, des écarts angulaires, du plan de référence, des mesures de distances, des dislocations et des déformations, des mesures des epctres atomiques et moléculaires, ainsi que de la construction des étalons de longueur et de fréquence. Les exemples illustrent non soulement le niveau des méthodes de mesures mais aussi le progrès obtenu au cours de la dernière vingtaine d'années.

W. WOLIŃSKI

LASSERMESSUNGEN IN DEN JAHREN 1968—1987 UND WEITERE ENTWICKLUNGSAUSSICHTEN

Zusammenfassung

Im Artikel wurden Anwendungsgruppen geschildert unter Nutzung derartiger Eigenschaften der Laserstrahlung, wie: hohe Monochromatizität, Interferenzvermögen bei großen Unterschieden der optischen Wege, Richtungsrelation und geringer Aufspreizungswinkel des Bündels sowie Gewinnungsmöglichkeiten kontinuierlicher Leistung, Impulsleistung, auch einer Impulsleistung mit großer Wiederholungsfrequenz der Impulse und Wellenlängenverstimmung. Erwogen wurden Anwendungen von Lasern zur Absteckung von: geraden Linien, Winkelabweichungen, Beziehungsflächen, Entfernungsmessungen, Verlagerungen und Verformungen, Messungen der Atom- und Molekülspektren sowie Aufbau von Längen- und Frequenzmustern. Die Beispiele lassen nicht nur des Niveau der Meßmethoden erkennen, sondern auch auf den Entwicklungsfortschritt in den Letzten 20 Jahren schließen.

В. ВОЛИНЬСКИ

ЛАЗЕРНЫЕ ИЗМЕРЕНИЯ В 1968—1987 ГОДАХ И ДАЛЬНЕЙШИЕ ПЕРСПЕКТИВЫ

Резюме

Представлены группы примений использующих такие свойства лазерного излучения, как высокая монохроматичность, способность интерференции при значительных разницах протяженности оптических трасс, направленность и незначительное рассеяние луча, возможность работы

постоянной, импульсной при чем также выскочастотной и возможность перенастройски длины волны. Обсуждены применения лазеров при транссировнии прямой линии, угловых смещений, базисной плоскости, измерений расстояния, перемещений и деформации, измерений атомных и молекулярных спектров, конструирования эталонов длины и частоты. Примеры показывают не только высокий уровень методов измерений но также и прогресс достичнутый в течение последних 20 лет.

ROZPRAWY ELEKTROTECHNICZNE, 1988, 33, z. 4, ss. 1177–1189

621.3.082.52:535,317

Optyczna interpolacja sygnałów w optoelektronicznych przyrządach opartych na wzorcach inkrementalnych

LESZEK WRONKOWSKI

Centrum Uczelniano-Przemysłowe Metrologii i Systemów Pomiarowych, Politechnika Warszawska

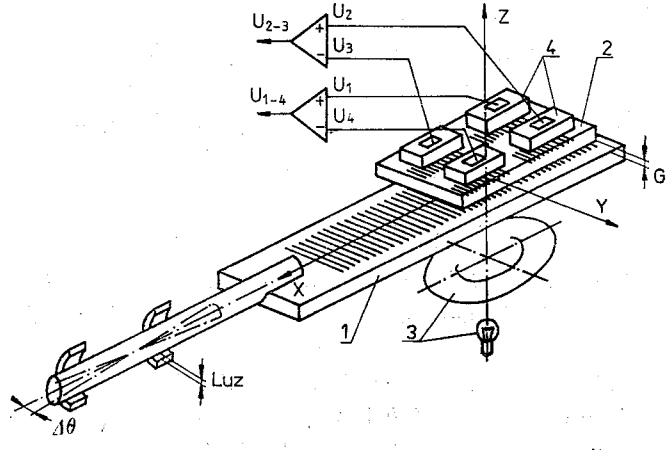
Otrzymano 1986.05.23

Autoryzowano do druku 1987.01.15

W pracy przedstawiono metody powiększenia rozdzielczości pomiarowej optoelektronicznych przyrządów za pomocą optycznej interpolacji. Na podstawie teoretycznych rozważań wykazano, że interpolacja optyczna umożliwia uzyskanie mniejszych błędów wskazań przyrządów niż w przypadku wykorzystania w nich interpolacji elektronicznej.

1. WPROWADZENIE

W dotychczasowych rozwiązaniach konstrukcyjnych optoelektronicznych przyrządów np. służących do pomiaru długości można wyróżnić dwa podstawowe zespoły: głowicę pomiarową i elektroniczny układ interpolatora, w którym jest również blok zliczający. W głowicy pomiarowej (rys. 1) znajduje się przetwornik przetwarzający przemieszczenie pomiarowe Δx na okresowe sygnały napięciowe lub prądowe. Przetwarzanie to jest realizowane za pomocą dwóch wzorców inkrementalnych (1, 2), które przesuwając się wzglę-



Rys. 1. Uproszczony konstrukcyjny schemat inkrementalnego przetwornika optoelektrycznego

dem siebie w wiązce promieni oświetlacza (3) modulują strumień świetlny odbierany przez fotoelementy (4). Obydwa wzorce inkrementalne są binarnymi siatkami dyfrakcyjnymi. Najczęściej jeden z nich skonstruowany jest w postaci optycznego przesuwnika fazy [12]. Optyczny przesuwnik fazy jest to n-elementowy zespół dyfrakcyjnych siatek binarnych naniesionych na jednej płaszczyźnie np. szklanej płytce w ten sposób, że kolejne elementy (sektory) siatek są wzajemnie przesunięte w fazie. Dotychczas stosowano w konstrukcjach dwu i czteroelementowe przesiuwniki fazy. Dla rewersyjnego zliczania impulsów niezbędne są dwa sygnały przesunięte w fazie o 90°. W przypadku dwuelementowego przesuwnika sektory są wzajemnie przesunięte o $d/4$ (d — stała, okres wzorca). Wartość działki elementarnej przetwornika z tym przesuwnikiem wyniesie $w_e = d/4$, przesuwniki te stosowane są najczęściej w przetwornikach do pomiaru kąta. Czteroelementowy przesuwnik fazy ma przesunięte sektory odpowiednio o 0, 1/4 d, 1/2 d i 3/4 d. Fotoelementy umieszszone za płaszczyzną przesuwnika przetwarzają zmienny strumień świetlny na sygnały napięciowe lub prądowe. Sygnały te najczęściej sumuje się różnicowo i uzyskuje się na wyjściu przetwornika dwa sygnały przesunięte o 90° z wyeliminowaną składową stałą. W rezultacie wartość działki elementarnej wyniesie również $d/4$. Dokładny opis różnych konstrukcji optycznych przesuwników fazy wraz z analizą układów detekcyjnych, z którymi są związane przedstawiono w pracach [7, 12].

W celu zwiększenia rozdzielczości pomiarowej stosuje się elektroniczną interpolację uzyskanych z przetworników sygnałów różnicowych. Uzyskane w bloku interpolatora zwielokrotnienione elektronicznie sygnały są dalej formowane na przebiegi prostokątne, różniczkowane i w postaci impulsów odbierane przez licznik. Wskazanie cyfrowe liczby impulsów jest miarą przesunięcia wzorca podstawowego o wartość Δx [1, 8].

W artykule tym zostały wykorzystane wyniki analiz i obliczeń różnych typów optycznych przesuwników fazy podanych w pracy [12].

2. BŁĘDY PRZETWORNIKA A BŁĘDY ELEKTRONICZNEJ INTERPOLACJI

Na początku rozważmy przypadek optoelektronicznego przyrządu pomiarowego opartego wyłącznie na przetworniku, a zatem bez bloku układu interpolującego.

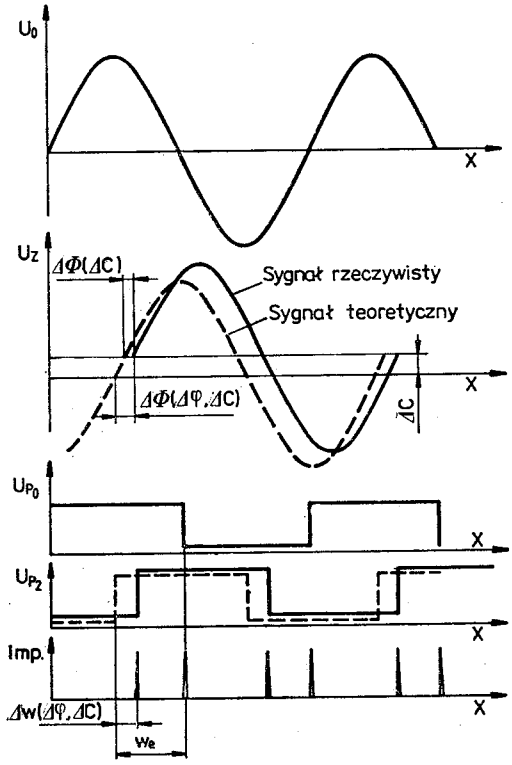
Na wyjściu przetwornika otrzymujemy dwa sygnały różnicowe, z których jeden przyjmujemy za odniesieniowy (1), a drugi przesunięty w fazie o 90° zawiera wszystkie składniki błędów: błąd fazy $\Delta\Phi$, błąd amplitudy ΔA i błąd składowej stałej ΔC

$$U_0 = \sin(360°x/d) \tag{1}$$

$$U_z = (1+\Delta A)\sin(360°x/d \pm 90° + \Delta\Phi \pm \arcsin(\Delta C)). \tag{2}$$

Wartość błędu składowej stałej ΔC jest unormowana do wartości amplitudy, której wartość nominalną przyjęto: $A = 1$. Zakłada się naturalnie, że przebiegi te mają charakter sinusoidalny.

Przyrównując zależności 1 i 2 do zera i rozwiązując je otrzymamy rozkład błędów równomierności podziału w funkcji $\Delta\Phi$ i ΔC, jakie powstaną w procesie formowania i różniczkowania tych sygnałów (rys. 2). Przez błąd równomierności podziału rozumie



Rys. 2. Schemat ilustrujący powstanie błędów równomierności podziału

się tutaj zależność

$$\Delta w = w - w_e, \tag{3}$$

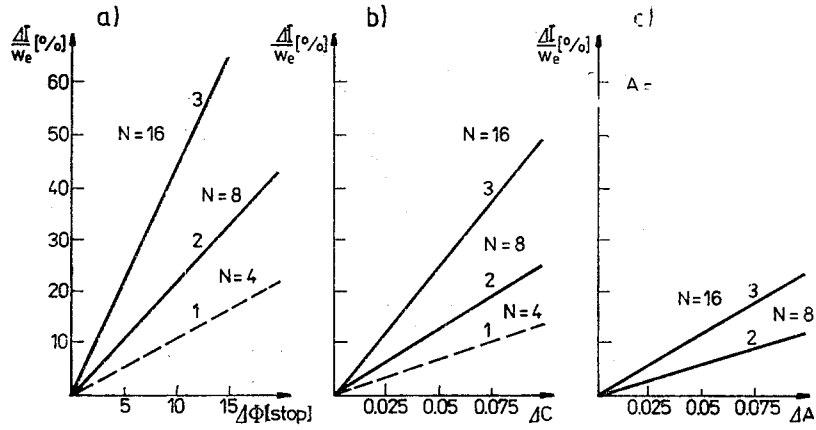
gdzie:

w_e — wartość nominalnej działki elementarnej,

w — wartość rzeczywista działki elementarnej.

Na rys. 3ab przedstawiono graficznie zależność względnego błędu równomierności podziału w funkcji błędów $\Delta\Phi$ i ΔC (linia przerywana). W badanym przedziale można przyjąć, że mają one charakter liniowy. Błędy amplitudy w tym przypadku nie wpływają na wartość błędu wskazania przyrządu. Fakt ten będzie wyjaśniony niżej.

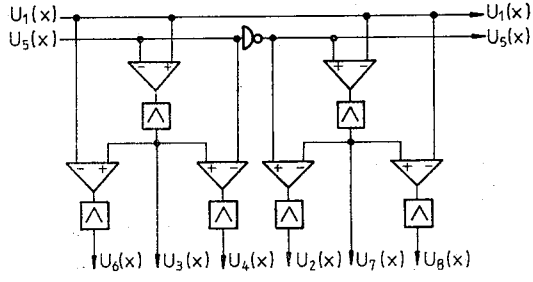
Rozpatrzmy drugi przypadek, kiedy w przyrządzie wykorzystano elektroniczny układ interpolatora. Jedna z metod realizacji interpolacji przedstawiona jest na rys. 4 — z dwóch przebiegów wyjściowych z przetwornika tworzy się za pomocą wzmacniaczy różnicowych dodatkowo sześć sygnałów. W rezultacie na wyjściu interpolatora uzyskuje się 8 sygnałów poprzesuwanych teoretycznie w fazie o 22,5° i wartość działki elementarnej na wyjściu przyrządu $d/16$. Dla wyznaczenia błędu wskazania przyrządu można dalej operować pojęciem błędu równomierności podziału (przykład rozwiązania tego problemu podany jest w [8]). Jednak bardziej precyzyjne kryterium oceny błędów przyrządów pracujących w systemach inkrementalnych zaproponował T. Karczmarczyk w pracy [10]. Jest to maksymalny błąd interpolacji $\pm\Delta I$, który wystąpi dla przypadku gdy dowolna (co do



Rys. 3. Wykresy maksymalnych błędów interpolacji sygnałów w funkcji a) błędów fazy $\Delta\Phi$, b) błędów składowej stałej ΔC, c) błędów amplitudy ΔA

liczby składników) suma kolejnych wartości działek rzeczywistych osiągnie największą różnicę w stosunku do sumy nominalnych wartości tych działek.

Nietrudno jest zauważyć, że względny błąd równomierności podziału, przyjęty dla obliczenia błędów samego przetwornika jest graniczną wartością maksymalnego błędu interpolacji, gdyż liczba składników sumy dającej największą różnicę wartości działek rzeczywistych w stosunku do nominalnych równa jest 1. Pozwala to na bezpośrednie porównanie błędów samego przetwornika z błędami uzyskanymi po interpolacji.



Rys. 4. Schemat realizacji elektronicznej interpolacji sygnałów

Wykorzystując powyższe stwierdzenie można na podstawie prac [8] i 10] sporządzić wspólne wykresy błędów $\Delta I = f(\Delta\Phi)$, $\Delta I = f(\Delta C)$ i $\Delta I = f(\Delta A)$ dla optoelektronicznych przyrządów, w których wykorzystuje się jedynie przetwornik jak i dla przyrządów, w których zastosowano również elektroniczne układy interpolujące. Na rys. 3abc przedstawiono owe wykresy, przy czym linią przerywaną (jak podano wyżej) zaznaczono charakterystykę błędów dla przyrządów opartych wyłącznie na przetworniku. Dla przyjętej zasady formowania impulsów prostokątnych — komparacji sygnałów na poziomie 0 — wpływ błędów amplitudy na błąd interpolacji sygnałów jest praktycznie wyeliminowany (rys. 3c). Liczba N oznacza krotność interpolacji, przy czym dla przetwornika wykorzystywana jest wyłącznie interpolacja optyczna a dla pozostałych przypadków optyczna i elektroniczna.

Analizując powyższe wykresy łatwo jest zauważyć, że maksymalne błędy interpolacji pochodzące od poszczególnych składników błędów sygnałów ($\Delta\Phi$, ΔC, ΔA) będą proporcjonalne do liczby krotności interpolacji N. Natomiast bezwzględne wartości błędów na wyjściu przyrządów, w których wykorzystano taki sam wzorzec a zwiększenie ich rozdzielczości osiągnięto przez wprowadzenie elektronicznej interpolacji będą w przybliżeniu stałe. Jedynie przy wykorzystaniu wyłącznie optycznej interpolacji błąd ten będzie mniejszy o składnik związany z błędem amplitudy. Powyższy wniosek skłania do poszukiwań nowych rozwiązań układów detekcyjnych przetworników, których rozdzielczość możnaby zwiększyć na drodze zwiększenia krotności interpolacji optycznej.

Powiększanie rozdzielczości pomiarowej poprzez wykorzystanie wzorców inkrementalnych o mniejszej stałej nastręcza wiele trudności konstrukcyjnych i technologicznych — zastosowanie wzorców o dwukrotnie większej rozdzielczości, przy warunku utrzymania stałego błędu względnego, powoduje konieczność uzyskania w przybliżeniu czterokrotnie mniejszych błędów wywołanych luzem bądź błędami kształtu prowadnic wzorców.

3. PROPOZYCJA POWIĘKSZENIA ROZDZIELCZOŚCI PRZETWORNIKA POPRZEZ ZWIELOKROTNIENIE INTERPOLACJI OPTYCZNEJ

W punkcie 3 pracy [12] przedstawiono nowy typ optycznego czteroelementowego przesuwnika fazy. Analiza rozkładów przesunięć fazowych tego przesuwnika oraz stosowanego dotychczas przesuwnika, którego sektory są rozmieszczone w konfiguracji kwadratowej [3, 12] umożliwiła opracowanie teoretyczne nowego układu detekcyjnego z ośmioelementowym optycznym przesuwnikiem fazy. Układ ten umożliwia ośmiokrotną optyczną interpolację i jednocześnie zapewnia autokompensację głównych składników błędów fazy, wywołanych luzami prowadnicy wzorca podstawowego przetwornika.

Schemat powyższego układu przedstawiony jest na rys. 5. Pokazano na nim ośmioelementowy optyczny przesuwnik fazy z naniesionym rozkładem przesunięć fazowych poszczególnych jego sektorów wraz z systemem połączeń ośmiu fotoelementów. Rozmieszczenie poszczególnych fotoelementów zwymiarowano, przyjmując za bazę oś optyczną układu.

Poszczególne sygnały uzyskane z fotoelementów można przedstawić w sposób następujący:

$$U_1 = \sin(x-45°-(b+\Delta b_{1,0})\Delta\varphi_\psi+(a+\Delta a_{1,0})\Delta\varphi^*+\varphi_{s1}) \tag{4}$$

$$U_2 = \sin(x+45°+(b+\Delta b_{0,2})\Delta\varphi_\psi+(a+\Delta a_{0,2})\Delta\varphi^*+\varphi_{s2}) \tag{5}$$

$$U_3 = \sin(x+225°-(b+\Delta b_{3,0})\Delta\varphi_\psi-(a+\Delta a_{3,0})\Delta\varphi^*+\varphi_{s3}) \tag{6}$$

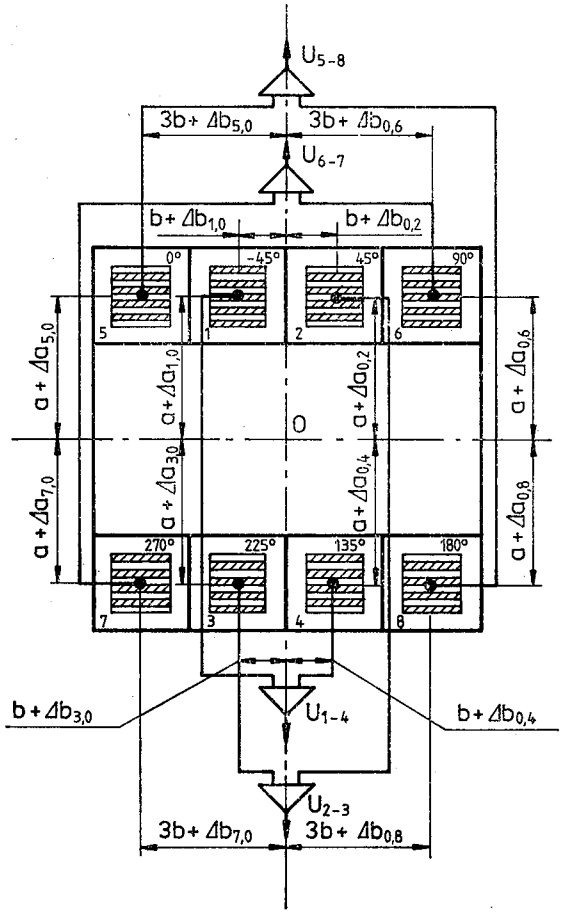
$$U_4 = \sin(x+135°+(b+\Delta b_{0,4})\Delta\varphi_\psi-(a+\Delta a_{0,4})\Delta\varphi^*+\varphi_{s4}) \tag{7}$$

$$U_5 = \sin(x-(3b+\Delta b_{5,0})\Delta\varphi_\psi+(a+\Delta a_{5,0})\Delta\varphi^*+\varphi_{s5}) \tag{8}$$

$$U_6 = \sin(x+90°+(3b+\Delta b_{0,6})\Delta\varphi_\psi+(a+\Delta a_{0,6})\Delta\varphi^*+\varphi_{s6}) \tag{9}$$

$$U_7 = \sin(x+270°-(3b+\Delta b_{7,0})\Delta\varphi_\psi-(a+\Delta a_{7,0})\Delta\varphi^*+\varphi_{s7}) \tag{10}$$

$$U_8 = \sin(x+180°+(3b+\Delta b_{0,8})\Delta\varphi_\psi-(a+\Delta a_{0,8})\Delta\varphi^*+\varphi_{s8}), \tag{11}$$



Rys. 5. Schemat optycznego ośmioelementowego przesuwnika fazy wraz z systemem połączeń fotoelementów

gdzie:

$$\Delta\varphi_{\psi} = \frac{\sin\Delta\Theta}{(1+\delta)d}\,2\pi$$

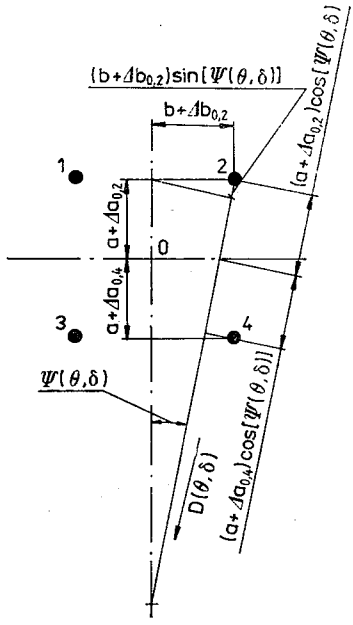
$$\Delta\varphi^{*} = \frac{(1+\delta-\cos\Delta\Theta)}{(1+\delta)d}\,2\pi.$$

Obydwa powyższe wyrażenia w połączeniu z odpowiednią odległością pomiędzy fotoelemtami są składnikami błędu fazy powstałych:

— pierwszy, na skutek zmiany kąta Ψ (Θ, δ), kąta pomiędzy kierunkiem przemieszczania się wzorca Δx, a kierunkiem prostopadłym do linii prążków moire $D(\Theta, \delta)$ (δ jest parametrem charakteryzującym rozbieżność wiązki oświetlacza układu detekcyjnego przetwornika, patrz praca [7]);

— drugi w wyniku pojawienia się prążka moire o skończonej długości (rys. 6).

Prążki moire wywołane są pojawiającymi się szczątkowymi wartościami kąta $\Delta\Theta$



Rys. 6. Schemat ilustrujący sposób powstawania składników błędów fazy: $(a+\Delta a)\varphi^*$, $(b+\Delta b)\varphi_\psi$

(teoretycznie obydwa wzorce nie powinny być wzajemne skręcone $\Delta\Theta = 0$, wówczas $D \to \infty$ a $\Psi \to 0$) oraz rozbieżnością wiązki oświetlacza [7]. Szczątkowe wartości wywołane są luzem łożyskowym prowadnic przetworników. φ_s — jest to składnik błędu fazy pochodzący od niedokładności naniesienia poszczególnych sektorów optycznego przesuwnika fazy, b, $2b$ ... oraz a są nominalnymi odległościami pomiędzy osiami fotoelementów i osią optyczną układu detekcyjnego, $\Delta b_{0,1}$, $\Delta b_{0,2}$ ... oraz $\Delta a_{0,1}$, $\Delta a_{2,0}$ ... są błędami geometrycznymi rozmieszczenia osi optycznych fotoelementów.

Postępując podobnie jak w pracy [12] sumując różnicowo poszczególne sygnały z fotoelementów wg schematu na rys. 5 uzyska się dwie pary sygnałów, których różnica faz będzie określona zależnościami 12 i 13:

$$\varphi_{\text{I}} = \varphi(U_1-U_4)-\varphi(U_2-U_3) = -90+\frac{\Delta b_{3,0}-\Delta b_{1,0}+\Delta b_{0,4}-\Delta b_{0,2}}{2}\Delta\varphi_\psi+$$

$$+\frac{\Delta a_{1,0}-\Delta a_{2,0}+\Delta a_{3,0}-\Delta a_{0,4}}{2}+\frac{\varphi_{s1}-\varphi_{s2}+\varphi_{s4}-\varphi_{s3}}{2} \tag{12}$$

$$\varphi_{\text{II}} = \varphi(U_5-U_8)-\varphi(U_6-U_7) = -90°+\frac{\Delta b_{0,6}-\Delta b_{5,0}+\Delta b_{0,8}-\Delta b_{7,0}}{2}\Delta\varphi_\psi+$$

$$+\frac{\Delta a_{5,0}-\Delta a_{7,0}+\Delta a_{0,6}-\Delta a_{0,8}}{2}\Delta\varphi^*+\frac{\varphi_{s5}-\varphi_{s6}+\varphi_{s8}-\varphi_{s7}}{2}. \tag{13}$$

Obydwie pary przebiegów spełniają warunki rewersji — przesunięte są w fazie o 90°. Błędy fazy $\Delta\Phi_{\text{I}}$ i $\Delta\Phi_{\text{II}}$ opisują pozostałe człony zależności 12 i 13. Błędy amplitudy natomiast można przedstawić w formie następującej:

$$\Delta A_{\mathrm{I}} = 4\sin\left(\left(\frac{\Delta b_{0,2}+\Delta b_{3,0}-\Delta b_{1,0}-\Delta b_{0,4}}{4}\right)\Delta\varphi_{\psi}+\right.$$

$$+\left(a+\frac{a_{1,0}+a_{0,2}+a_{3,0}+a_{0,4}}{4}\right)\Delta\varphi^{*}+\frac{\varphi_{s1}+\varphi_{s3}-\varphi_{s2}-\varphi_{s4}}{4}\Bigg)$$

$$\sin\left(\left(b+\frac{\Delta b_{1,0}+\Delta b_{0,2}+\Delta b_{3,0}+\Delta b_{0,4}}{4}\right)\Delta\varphi_{\psi}+\right.$$

$$\left.-\frac{\Delta a_{1,0}+\Delta a_{0,4}-\Delta a_{0,2}-\Delta a_{3,0}}{4}\Delta\varphi^{*}-\frac{\varphi_{s1}+\varphi_{s3}-\varphi_{s2}-\varphi_{s4}}{4}\right) \tag{14}$$

$$\Delta A_{\mathrm{II}} = 4\sin\left(\frac{\Delta b_{0,6}+\Delta b_{0,8}-\Delta b_{5,0}-\Delta b_{0,6}}{4}\Delta\varphi_{\psi}+\right.$$

$$+\left(a+\frac{\Delta a_{5,0}+\Delta a_{7,0}+\Delta a_{0,8}-\Delta a_{7,0}}{4}\right)\Delta\varphi^{*}+\frac{\varphi_{s7}+\varphi_{s5}-\varphi_{s6}-\varphi_{s7}}{4}\Bigg)$$

$$\sin\left(\left(3b+\frac{b_{5,0}+b_{0,6}+b_{7,0}+b_{0,8}}{4}\right)\Delta\varphi_{\psi}+\right.$$

$$\left.-\frac{a_{5,0}+a_{0,6}+a_{0,8}-a_{7,0}}{4}\Delta\varphi^{*}-\frac{\varphi_{s5}+\varphi_{s6}-\varphi_{s7}-\varphi_{s8}}{4}\right) \tag{15}$$

Jak widać z powyższego zapisu błędy amplitudy bardzo silnie są związane z przyjmowanymi odległościami a i b pomiędzy fotodetektorami. W związku z tym różnice pomiędzy amplitudami uzyskanych sygnałów różnicowych — przy stosowanych odległościach a i b w obecnych konstrukcjach przetworników — mogą wynosić nawet kilkanaście procent [12]. Fakt ten jednak dla rozważanego przypadku jest nieistotny, natomiast ogranicza on możliwość wprowadzenia dodatkowo interpolacji elektronicznej.

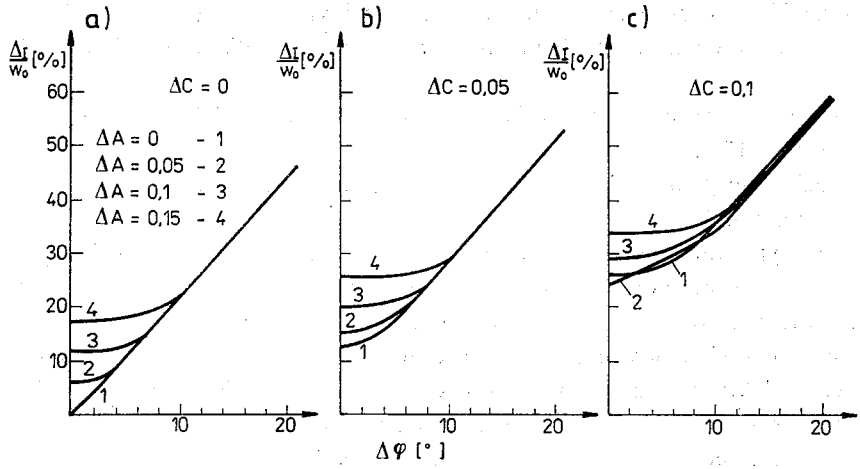
Znacznie korzystniej przedstawia się sytuacja w zakresie błędów fazy. Jak wynika z zależności (12) i (13), błędy fazy zależą od różnic pomiędzy odległościami poszczególnych par fotodetektorów (Δa i Δb), których wartości mogą być o rząd, dwa lub nawet trzy rzędy mniejsze od a i b oraz od dokładności wykonania optycznego przesuwnika fazy φ_s. Analiza jakościowa i ilościowa wpływu tych parametrów na błędy fazy oraz błędy amplitudy została przedstawiona w pracy poprzedniej [12].

4. ANALIZA I WNIOSKI

Jak wynika z p. 3, zaproponowany układ detekcyjny zapewnia powstanie dwóch par sygnałów, które spełniają warunki rewersji. Z sygnałów tych, po ich uformowaniu na przebieg prostokątny i zróżniczkowaniu, można uzyskać ciąg impulsów. Teoretyczne odległości pomiędzy nimi powinny wynosić $we = d/8$ (ośmiokrotna interpolacja optyczna). W rzeczywistości wyżej omówione błędy fazy i błędy składowych stałych wywołują nie-

równomierność rozkładu wartości działek elementarnych, można ją ocenić za pomocą przyjętego kryterium — maksymalnego błędu interpolacji. W przypadku nie wprowadzania dalszej interpolacji elektronicznej i przyjęcia zerowego poziomu komparacji tych sygnałów, ich błędy amplitudy nie wpływają praktycznie na wartość błędu interpolacji.

W celu porównania błędów na wyjściu przyrządów wykorzystujących ośmiokrotną interpolację — częściowo optyczną i częściowo elektroniczną — z błędami przyrządów, w których zastosowano wyłącznie optyczną interpolację na rys. 7a, b, c przedstawiono, w postaci wykresów, globalny wpływ błędów $\Delta\Phi$, ΔA i ΔC na maksymalny błąd inter-



Rys. 7. Wykresy maksymalnych błędów interpolacji ($N = 8$) w funkcji a) błędów fazy i amplitudy dla $\Delta C = 0$, b) błędów fazy i amplitudy dla $\Delta C = 0{,}05$, c) błędów fazy i amplitidy dla $\Delta C = 0{,}1$

polacji [6]. Przy optycznej interpolacji wartość tego błędu można wyznaczyć na podstawie wykresu 3a uwzględniając składnik błędu ΔC w błędzie fazy (zależność 2 i rys. 2).

Z analizy przedstawionych wykresów wynika, że dla małych wartości błędów fazy ($\Delta\Phi < 10°$) znaczący wpływ na maksymalny błąd interpolacji mają błędy amplituudy i błędy składowej stałej. Natomiast przy zwiększaniu się błędu fazy przyjmuje on na siebie dominującą rolę we wzroście błędu interpolacji. Wynika stąd pierwszy istotny wniosek — zastosowanie optycznej interpolacji może spowodować zwiększenie dokładności wskazań optoelektronicznych przyrządów pomiarowych, ale tylko w przypadku, kiedy błędy fazy ich układów detekcyjnych są niewielkie, rzędu paru stopni.

Dla pełniejszego zilustrowania powyższego wniosku przedstawiony tu jest przykład.

W optoelektronicznym przetworniku przyjęto, że luz w jego prowadnicach powoduje powstanie szczątkowej wartości kąta Θ, która może się zmieniać w przedziale $\Delta\Theta = \pm 6'$ a parametr charakteryzujący rozbieżność oświetlacza $\delta = 5 \cdot 10^{-4}$. Natomiast błędy układu detekcyjnego przetwornika przyjmą następujące wartości:

— błędy rozmieszczenia fotoelementów $\Delta a = \Delta b = 0{,}01$ mm (tak małe błędy można osiągnąć wykorzystując opracowane przez autora zintegrowane układy fotodetektorów [13]),

— błędy optycznego przesuwnika faz, rozumianych jako różnica nominalnych i rze-

czywistych przesunięć fazowych jego sektorów, nie powinny przekroczyć wartości $\varphi_s = 0{,}1/8\pi$

— błąd składowych stałych ΔC obydwóch par sygnałów nie powinien być większy niż 0,05.

Założono również, że w przetworniku wykorzystano wzorzec o stałej $d = 16$ μm, który przy przyjętej krotności interpolacji $N = 8$ pozwoli uzyskać nominalną wartość działki przyrządu $we = 2$ μm.

Na podstawie powyższych danych, wykorzystując wykresy 4a, 7a i 7b z pracy [12] oraz wykresy 3a lub 7a i 7b przytoczone w artykule, można obliczyć kolejno składniki błędów sygnałów tworzonych w układzie detekcyjnym przetwornika ($\Delta\Phi$, ΔA) oraz wynikające z nich maksymalne błędy interpolacji ($\Delta I/we$) dla obydwóch typów interpolacji:

a) interpolacja optyczna: $\Delta\Phi(\Delta a, \Delta b, \varphi_s \Delta C) = 0{,}8° + 0{,}2° + 3{,}3° + 2{,}9° = 7{,}2°$
$\Delta I/we(\Delta\Phi) = 17\%$,

b) interpolacja optyczno-elektroniczna: $\Delta\Phi(\Delta b, \varphi_s) = 3{,}5°$, $\Delta A(\Delta a, \Delta b, \varphi_s) = 0{,}1$ a$\Delta I/we(\Delta\Phi, \Delta A, \Delta C) = 21\%$

Wyznaczonym wartościom maksymalnych błędów interpolacji odpowiadają wartości błędów wskazań przyrządu $\delta_e = 0{,}34$ μm w pierwszym przypadku i $\delta_e = 0{,}42$ μm w drugim. Błędy te oczywiście nie uwzględniają błędu dyskretyzacji.

Uzyskana niewielka przewaga w zakresie dokładności wskazań interpolacji optycznej nad optyczno-elektroniczną dla przyjętych do obliczeń wartości parametrów błędotwórczych przetwórnika, która będzie wzrastała wraz z maleniem błędów fazy, nie rozsądza jeszcze w pełni zalet i wad omawianego układu detekcyjnego.

Należy zauważyć, że pierwotne pary sygnałów: $U_1 - U_4$, $U_2 - U_3$, $U_5 - U_8$ i $U_6 - U_7$ są generowane z par fotodetektorów symetrycznie rozłożonych względem osi x. Natomiast wzorzec przesuwny przetwornika „*sprzęgający*" optycznie obie części przesuwnika fazowego, poza informacją właściwą — przemieszczenie pomiarowe — przekazuje również składnik błędu związany ze szczątkową wartością kąta Θ. Jeżeli założymy, że błędy geometryczne Δa, Δb oraz φ_s są równe zeru, to dla każdej wartości z przedziału $\pm\Delta\Theta$ fotodetektory w obu częściach przesuwnika wygenerują sygnały z błędami fazy o równych wartościach lecz przeciwnych znakach. Proponowany sposób łączenia fotodetektorów przy tworzeniu sygnałów różnicowych (rys. 5) powoduje, że błędy te we współpracujących z sobą parach sygnałów wzajemnie się znoszą. Na tym polega zresztą wspomniana wyżej autokompensacja błędów fazy optycznego przesuwnika fazy.

Rozważmy teraz dwa skrajne, acz mało prawdopodobne przypadki, kiedy błędy Δa, Δb i φ_s są tak dobrane co do wartości i znaków, że człony zależności (12) i (13) reprezentujące błędy fazy $\Delta\Phi_{\mathrm{I}}$ i $\Delta\Phi_{\mathrm{II}}$ będą miały takie same wartości i znaki; lub takie same, ale maksymalne wartości i przeciwne znaki. W pierwszym przypadku błąd wskazania ograniczyłby się tylko do członu $\Delta I/we(\Delta C)$ i wyniósłby 0,24 μm. W drugim, błąd fazy wzrósłby dwukrotnie, co spowodowałoby zwiększenie błędu $\Delta I/we(\Delta\Phi)$ do 11,5% a błędu wskazania, który w tym wypadku miałby charakter błędu maksymalnego = do 0,5 μm. Błąd ten byłby więc nieco większy od błędu obliczonego dla przyrządu z tradycyjnym układem detekcyjnym.

W rzeczywistości przyjęta technologia wykonania optycznego przesuwnika fazy eli-

minuje systematyczne błędy nanoszenia na szklanej płytce jego sektorów, dopuszcza jedynie błędy przypadkowe. Podobnie rozkład błędów rozmieszczenia osi zespołu fotodetektorów, konstruowanego w postaci układu zintegrowanego na jednej płytce kryształu półprzewodnikowego, ma charakter rozkładu normalnego [2]. Wychodząc z powyższych stwierdzeń, jak i faktu, że liczba składników błędów jest spora — 16 nastąpi dość znaczne uśrednienie tych błędów. A dodając jeszcze, że poszczególne te składniki, jak wynika z zależności (12) i (13) odejmują się parami, rzeczywisty błąd fazy omawianego układu nie powinien zbytnio przekroczyć połowy obliczonego wyżej błędu maksymalnego.

Ponadto w przetworniku zmiana odległości pomiędzy wzorcami jest źródłem trzech innych składników błędów amplitudy spowodowanych:

— bardzo silnym maleniem natężenia oświetlenia przy wzroście odległości pomiędzy wzorcami, wyniki doświadczalne w tym zakresie przedstawiono w pracy [11],

— rozrzutem charakterystyk amplitudowo-częstotliwościowych fotodetektorów, która powoduje, że różnice zmian amplitud poszczególnych sygnałów wywołane zmienną prędkością v_p przesuwu trzpienia pomiarowego przetwrnika są niejednakowe (v_p: 0—500 mm/s czemu odpowiada zmiana częstości optycznego sygnału od 0 do 50 kH dla stałej wzorca $d = 16$ μm),

— zmiana widma uzyskanych sygnałów wywołanych zjawiskiem filtracji optycznej [4]. Ostatnie zjawisko polega na tym, że w miarę oddalania się wzorców od siebie, profil sygnałów zmienia się — od kształtu trójkątnego (w momencie zetknięcia się wzorców) do quasisinusoidalnego. Powoduje ono zarówno bezpośrednio zmiany amplitidy sygnału, jak też dodatkowe błędy w procesie elektronicznej interpolacji sygnałów [6]. Powyższe czynniki błędotwórcze wpływają naturalnie na wzrost wartości błędów przyrządów, w których zastosowano interpolację elektroniczną, natomiast praktycznie nie wpłyną na końcową wartość błędu przyrządu, w którym wykorzystano proponowany układ detekcyjny.

Podsumowując powyższe rozważania można stwierdzić, że wykorzystując wyłącznie optyczną interpolację sygnałów w optoelektronicznych przyrządach pomiarowych można uzyskać korzystniejsze ich charakterystyki dokładnościowe — jest to istotna zaleta proponowanego układu detekcyjnego. Dodatkowym walorem takiego rozwiązania jest wyeliminowanie z układu elektronicznego przyrządu zespołu interpolatora. Tym samym głowica pomiarowa, w której umieszczonoby prosty układ formujący i różniczkujący, mogłaby współpracować z standardowym licznikiem impulsów.

Natomiast wadą tego układu jest ograniczenie maksymalnej rozdzielczości, jaką można uzyskać. Otóż, z dotychczasowych doświadczeń wynika, że wzorce o największej rozdzielczości, jakie można obecnie wykonać mają stałą $d = 8$ μm. Tak więc przyrządy oparte na omawianym układzie detekcyjnym mogą osiągnąć najmniejszą działkę elementarną $we = 1$ μm. Co prawda, przez analogię można sobie wyobrazić optyczny przesuwnik fazy z szesnastoma sektorami, ale trudności technologiczne jego wykonania przekraczają w chwili obecnej możliwości realizacji tego pomysłu. Natomiast stosując bardzo dokładne systemy łożyskowania wzorca w przetworniku i elektroniczną interpolację można uzyskać większą rozdzielczość. Np. w Centrum Uczelniano-Przemysłowym Metrologii i Systemów Pomiarowych Politechniki Warszawskiej opracowano na bazie wzorca o $d = 8$ μm optoelektroniczny długościomierz cyfrowy, o zakresie pomiarowym 30 mm

i wartości działki elementarnej $we = 0{,}5$ μm. Przyrząd ten jest obecnie wdrażany do produkcji seryjnej w OBRN VIS.

Na zakończenie należy wyjaśnić, że zastosowane przez autora pojęcie elektronicznej interpolacji odnosi się do powszechnie stosowanego dotychczas systemu powielania sygnałów — jedną z metod jej realizacji przedstawia rys. 4. Obszerniejsze informacje na ten temat można znaleźć w pracy [6]. Prowadzone są natomiast w tej chwili prace nad wykorzystaniem w tych przyrządach innego typu elektronicznej interpolacji za pomocą układów mikroprocesorowych.

BIBLIOGRAFIA

1. L. Wronkowski: *Theoretical of the length measuring impulse transducer, accompanied by the computer simulation.* Acta IMEKO VII 1979
2. L. Wronkowski: *Fehlerguelen zu Langmassung bestimmten optoelektronischen Wandlern.* Feingeratetechnik 1980, nr 9
3. T. Karczmarczyk, L. Wronkowski: *Analiz pogrsznosti fazy i amplitudy w optoelektronicznych preobrazowatielach.* Mat. Konferencji „Emiscon'81" Strebskie Pleso, 1981
4. L. Wronkowski: *An analysis of the effect of diaphragm shape on the contrast of filtered moire fringe.* Optica Acta 1983, vol. 30, no 1
5. L. Wronkowski: *Wlijanie difrakcji Fresnela na metrologiczeskuju charakteristiku optoelektronnych analogo-impulsnych preobrazowatielej,* Mat. Konferencji Emiscon'83. Tatrzańska Łomnica 1983
6. T. Karczmarczyk: *Wybrane zagadnienia interpolacji sygnałów w przyrządach z przetwornikami opartymi na wzorcach inkrementalnych.* Mat. Konferencji „Metrologia'83", Warszawa 1983
7. L. Wronkowski: *Analiza czułości układów detekcyjnych optoelektronicznego długościomierza cyfrowego na błędy technologiczne jego przetwornika.* Mat. Konferencji „Metrologia'83" Warszawa 1983
8. L. Wronkowski: *Optoelektroniczne długościomierze — analiza struktury błędów.* Mechanik 1984, nr 2
9. L. Wronkowski, T. Karczmarczyk, M. Pleskacz: *Optoelektroniczne długościomierze cyfrowe.* Aparatura naukowa i dydaktyczna 1984, nr 9
10. T. Karczmarczyk: *Błędy procesu interpolacji w przyrządach z przetwornikami opartymi na wzorcach inkrementalnych.* PAK 1985, nr 5
11. L. Wronkowski: *Opto-electronic analog-impulse transducer accuracy from the point of diffraction phenomenon.* Mat. X Kongresu IMEKO, Praha, 1985, vol. 10
12. L. Wronkowski: *Optyczne przesuwniki fazy i ich wpływ na dokładność pracy optoelektronicznych przetworników.* Rozprawy Elektrotechniczne. PAN, 1988 tom XXXIV zeszyt 1
13. L. Wronkowski: *Sprawozdanie z realizacji I etapu pracy pt. „Optoelektroniczne długościomierze cyfrowe" realizowanej w ramach CPBR 12.2.* Politechnika Warszawska, 1987

L. WRONKOWSKI

OPTICAL INTERPOLATION OF SIGNALS IN OPTO-ELEKTRONIC INSTRUMENTS BASED ON INCREMENTAL MASTERS

Summary

The paper presents methods of raising the resolution of opto-electronic instruments by means of an optical interpolation. Basing on theorectic considerations it was proven that the indication errors obtaine when the optical interpolation is applied are smaller than when the electronic interpolation is applied.

L. WRONKOWSKI

INTERPOLATION OPTIQUE DES SIGNAUX DANS LES INSTRUMENTS OPTOÉLECTRONIQUES BASÉE SUR LES REGLES INCREMENTALES

Résumé

Les méthodes d'augmentation du pouvoir résolvant des instruments optoélectroniques à l'aide d'interpolation optique sont présentées dans l'article. L'analyse théorique a prouvé que l'interpolation optique permet de réduire les erreurs des instruments de mesure plus effectivement, que l'interpolation électronique.

L. WRONKOWSKI

OPTISCHE SIGNALINTERPOLATION BEI OPTOELEKTRONISCHEN GERÄTEN MIT INKREMENTALEN GEBERN

Zusammenfassung

In diesem Beitrag werden Steigerungsmethoden für die Meßquantisierung von optoelektronischen Gärenten mittels optischer Interpolation geschildert. Anhand theoretischer Überlegungen wird gezeigt, daß die optische Interpolation der elektronischen Interpolation gegenüber eine deutliche Steigerung der Anzeigegenauigkeit von Meßgeräten gewährleistet.

Л. ВРОНКОВСКИ

ОПТИЧЕСКАЯ ИНТЕРПОЛЯЦИЯ СИГНАЛОВ В ОПТОЭЛЕКТРОННЫХ ПРИБОРАХ ИСПОЛЬЗУЮЩИХ ИНКРЕМЕНТАЛЬНЫЕ ЭТАЛОНЫ

Резюме

Представлен метод увеличения измерительной разрешающей способности оптоэлектронных приборов путем применения оптической интерполяции. На базе теоретических выводов доказано, что оптическая интерполяция позволяет получить уменшение погрешности измерений по отношению к случаю использования в них оптоэлектронной интерполяции.

ROZPRAWY ELEKTROTECHNICZNE, 1988, 34, z. 4, ss. 1191-1202

621.3.087:53.087.32

Wpływ struktury i sposobu polaryzacji na właściwości metrologiczne analogowego przetwornika ciśnienia z diodą tunelową

JÓZEF SZMITKOWSKI

Szkoła Główna Służby Pożarniczej, Warszawa

Otrzymano 1987.05.15

Autoryzowano do druku 1988.01.22

W artykule przedstawiono, opierając się na własnych badaniach eksperymentalnych i obliczeniach numerycznych, rozważania dotyczące zastosowania diody tunelowej w prostej strukturze układowej do pomiaru ciśnienia hydrostatycznego.

Zastosowanie diody tunelowej do pomiaru ciśnienia jest znane od dość dawna i omówione w szeregu publikacjach zarówno krajowych jak i zagranicznych [1, 2, 3, 5, 6]. Analiza materiału literaturowego wykazuje, że spośród różnych analogowych układów pracy przetwornika, na uwagę zasługuje układ różnicowy o polaryzacji prądowej. Podstawową częścią składową tego układu jest prosta struktura równoległa, tj.: połączenie równoległe diody tunelowej z rezystorem.

W artykule przedstawiono, opierając się na wynikach własnych badań eksperymentalnych i obliczeniach numerycznych, analizę pracy układu równoległego uwzględniającą dobór jego parametrów. Przyjęto następujące oznaczenia:

DT — dioda tunelowa z otwartą obudową,
I_b — prąd bocznika,
I_d — prąd diody tunelowej,
I_0, I_0' — prąd polaryzacji,
I_p — prąd szczytu diody tunelowej,
I_v — prąd doliny diody tunelowej,
I — prąd wypadkowy układu równoległego dioda tunelowa — rezystor,
p — ciśnienie działające na diodę tunelową,
p_0 — ciśnienie odniesienia,
$\Delta p = p - p_0$ — przyrost ciśnienia,
S_p — czułość ciśnieniowa przetwornika,
T — temperatura diody tunelowej,
T_0 — temperatura odniesienia,
U_d — napięcie na diodzie tunelowej,
U_p — napięcie szczytu diody tunelowej,

U_v — napięcie doliny diody tunelowej,

U_{wy} — napięcie wyjściowe z przetwornika przy ciśnieniu p,

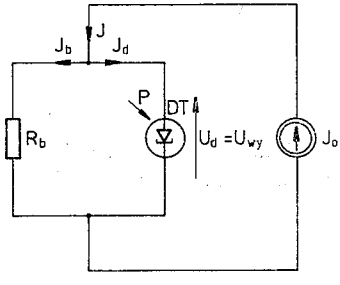
U_{wy_0} — napięcie wyjściowe z przetwornika przy ciśnieniu odniesienia p_0,

$\Delta U_{wy} = U_{wy} - U_{wy_0}$ — przyrost napięcia wyjściowego z przetwornika,

$\delta_T U_{wy}$ — względny błąd temperatury przetwarzania,

δ_{nl} — błąd nieliniowości charakterystyki przetwarzania.

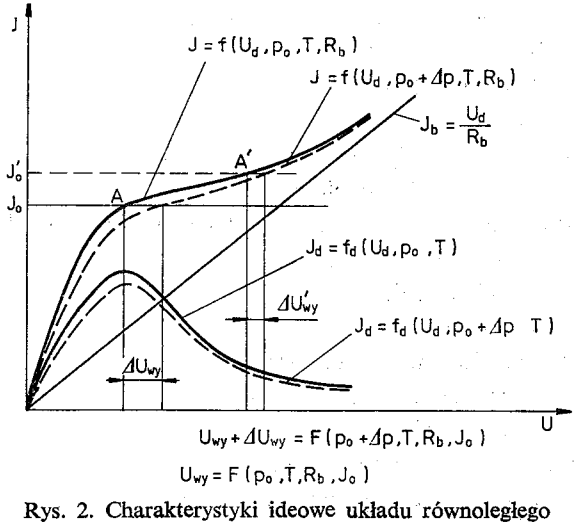
W układzie przedstawionym na rys. 1 dioda tunelowa DT jest połączona równolegle z rezystorem R_b, a następnie układ ten zasilany jest ze źródła prądowego o regulowanej wartości prądu I_0, który nazywany będzie dalej prądem polaryzacji. Napięcie na diodzie (z otwartą obudową) jest funkcją działającego na nią ciśnienia.



Rys. 1. Schemat układu równoległego o polaryzacji prądowej

Próba analitycznego opisu zależności napięcia diody w funkcji ciśnienia napotyka na trudności, ponieważ prąd diody jest złożoną funkcją napięcia, ciśnienia i temperatury, a podawane w literaturze wzory [5] do obliczania tego prądu wyprowadza się przy szeregu założeniach upraszczających i są one mało przydatne do obliczeń inżynierskich. Omawiane zależności, w postaci ideowych charakterystyk prądowo-napięciowych, podano na rys. 2.

Zależność prądu diody od napięcia na diodzie, ciśnienia i temperatury można zapisać



Rys. 2. Charakterystyki ideowe układu równoległego

w postaci ogólnej

$$I_d = f_d(U_d, p, T). \tag{1}$$

Z rys. 1 wynika, że

$$I = I_d + \frac{U_d}{R_b} \tag{2}$$

Zakładając określoną wartość prądu polaryzacji I_0 otrzymuje się wyrażenie

$$I_0 = f_d(U_d, p, T) + \frac{U_d}{R_b}, \tag{3}$$

z którego uzyskuje się, na przykład po rozwiązaniu graficznym — jak to pokazano na rys. 2 zależność napięcia wyjściowego przetwornika (napięcie na diodzie)

$$U_{wy} = F(p, T, I_0, R_b) \tag{4}$$

od ciśnienia, temperatury, prądu polaryzacji oraz rezystancji bocznika.
Przy zmianie ciśnienia o Δp, a przy niezmienionych wartościach wielkości wpływających, napięcie wyjściowe przetwornika zmieni się o ΔU_{wy}. Zmiany te w wybranych dwóch początkowych punktach pracy (A, A[1]) pokazano na rys. 2. Zależność między wartościami przyrostu ciśnienia Δp a odpowiadającymi im wartościami przyrostu napięcia ΔU_{wy} stanowi statyczną charakterystykę przetwarzania omawianego przetwornika

$$\Delta U_{wy} = f_p(\Delta p). \tag{5}$$

Ciśnieniowa czułość przetwornika, definiowana jako

$$S_p = \frac{\partial U_{wy}}{\partial p} \tag{6}$$

zależy, jak to wynika z rys. 2, od wyboru punktu pracy (A, A') na charakterystykach prądowo-napięciowych przetwornika. Z kolei punkt pracy zależy od wartości prądu polaryzacji I_0 i wartości rezystancji bocznika R_b.

Zmiana temperatury, podobnie jak zmiana ciśnienia, powoduje zmianę napięcia wyjściowego przetwornika, to z kolei powoduje powstanie błędu pomiaru, który nazywany będzie dalej temperaturowym błędem przetwarzania $\delta_T U_{wy}$. Zasadę doboru punktu pracy oraz ocenę wartości temperaturowego błędu przetwarzania omówiono w dalszej części artykułu.

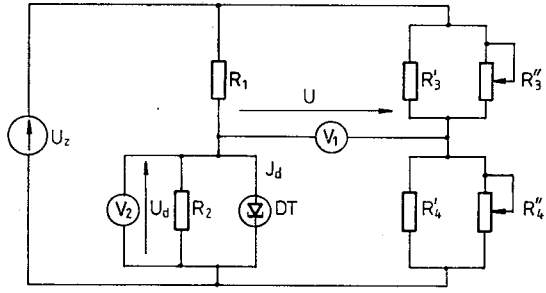
Do pomiaru charakterystyk diod tunelowych zastosowano układ mostkowy (rys. 3) z włączoną do jednej jego gałęzi głowicą pomiarową. W przedstawionym układzie napięcie na diodzie mierzy się bezpośrednim woltomierzem VC_2, natomiast prąd diody pośrednio — mierząc napięcie U na przekątnej mostka woltomierzem VC_1 i obliczając wartość prądu ze wzoru

$$I_d = \frac{U}{K} \tag{7}$$

gdzie:

$$K = \frac{R_1 R_2}{R_1 + R_2}.$$

Zależność (5) jest poprawna tylko w przypadku, gdy przed podłączeniem diody tunelowej układ mostkowy zostanie zrównoważony, tj. napięcie U zostanie sprowadzone do zera przez regulację wartości R_3'' i R_4''. Warunek ten można łatwo uzasadnić, wyprowadzając zależność wartości prądu I_d od wartości rezystancji gałęzi mostka i napięcia zasilającego.



Rys. 3. Układ mostkowy do pomiaru charakterystyk diod tunelowych

Na przykład, stosując twierdzenie Thevenina, prąd diody I_d można wyrazić wzorem

$$I_d = \frac{U_z \dfrac{R_2}{R_1+R_2}}{\dfrac{R_1 R_2}{R_1+R_2} + \dfrac{U_d}{I_d}} \tag{8}$$

a następnie po prostych przekształceniach otrzymuje się wyrażenie

$$I_d = \frac{U + U_z \dfrac{R_2 R_3 - R_1 R_4}{(R_1+R_2)(R_3+R_4)}}{\dfrac{R_1 R_2}{R_1+R_2}}, \tag{9}$$

gdzie:

$$R_3 = \frac{R_3' R_3''}{R_3'+R_3''}, \qquad R_4 = \frac{R_4' R_4''}{R_4'+R_4''},$$

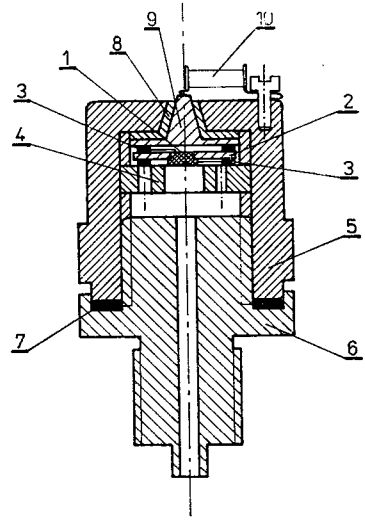
z którego wynika zależność (5), jeżeli spełniony będzie wymieniony wyżej warunek równowagi mostka: $R_2 R_3 = R_1 R_4$.

Konstrukcję głowicy pomiarowej, w której umieszczono badaną diodę, pokazano na rys. 4. Głowica ta składa się z dwóch części: w pierwszej (górnej) mocuje się diodę, a druga część służy do połączenia głowicy z prasą obciążnikowo-tłokową. Na zewnątrz głowicy między obudową a przepustem przyłączony jest rezystor ($R_2 = R_b$) bocznikujący diodę. Konstrukcja ta zapewnia łatwą i szybką wymianę badanych diod [8].

Z kolei głowicę pomiarową po połączeniu, za pomocą łącznika i odpowiedniej rurki, z prasą obciążnikowo-tłokową, umieszczano w termostacie. Przy nastawionej i ustalonej wartości temperatury, dla różnach wartości ciśnienia, mierzono „punkt po punkcie" współrzędne charakterystyk badanych diod.

W dalszej części artykułu przyjęto następujące określenia dotyczące omawianych dalej charakterystyk:

— podstawowa charakterystyka prądowo-napięciowa diody jest to zależność prądu diody I_d od napięcia na diodzie U_d w warunkach odniesienia, tj.: przy ciśnieniu p_0 (10^5 Pa) i temperaturze T_0 (20°C);
— charakterystyka prądowo-napięciowa z ciśnieniem jako parametrm jest to zależność przyrostu prądu diody $\Delta_p I_d$ od napięcia na diodzie U_d przy stałym ciśnieniu p i przy temperaturze odniesienia T_0;
— charakterystyka prądowo-napięciowa z temperaturą jako parametrem jest to zależność przyrostu prądu diody $\Delta_T I_d$ od napięcia na diodzie U_d przy stałej temperaturze T i przy ciśnieniu odniesienia p_0.

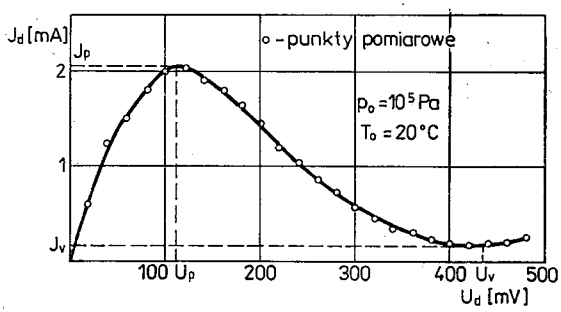


Rys. 4. Głowica pomiarowa
1 — dioda tunelowa, 2 — krążek izolacyjny, 3 — pierścienie mosiężne, do których przylutowane są wyprowadzenia diody, 4 — pierścień mocujący, 5 — górna część głowicy, w której umieszcza się diodę, 6 — dolna część głowicy, którą wkręca się do prasy obciążnikowo-tłokowej, 7 — pierścień uszczelniający, 8 — stożkowa tulejka izolacyjna, 9 — stożkowy przepust przewodzący, 10 — rezystor bocznikujący diodę

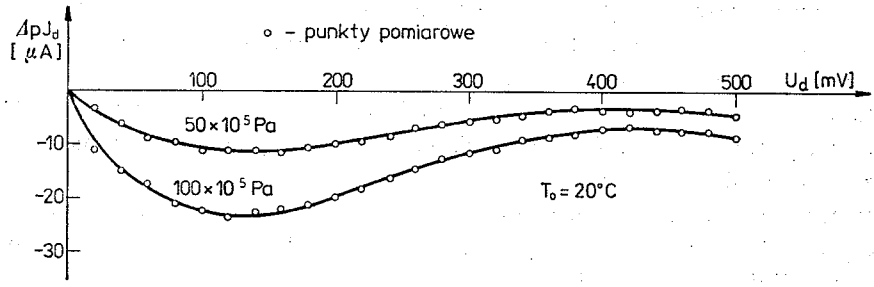
Przy czym: przyrost prądu $\Delta_p I_d$ jest różnicą między prądem diody przy ciśnieniu p i temperaturze T_0 a prądem diody w warunkach odniesienia; przyrost prądu $\Delta_T I_d$ jest różnicą między prądem diody przy temperaturze T ciśnieniu p_0 a prądem diody w warunkach odniesienia.

Do pomiarów charakterystyk wykorzystano, dostępne autorowi, diody tunelowe z arsenku galu produkcji radzieckiej typu AN — 301A.

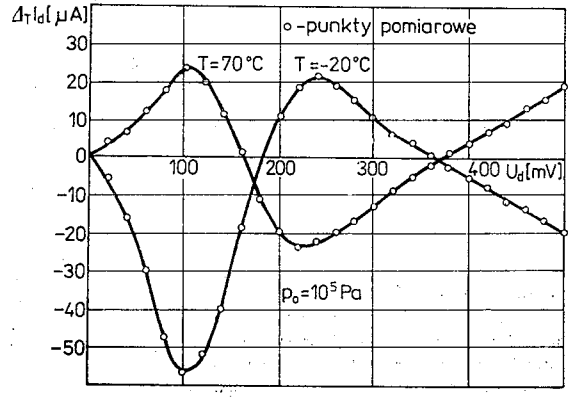
Przykładowe eksperymentalne charakterystyki diod, zmierzone za pomocą omówionego wyżej układu pomiarowego, przedstawiono na rysunkach: 5, 6, 7. Z wykreślonej



Rys. 5. Eksperymentalna podstawowa charakterystyka prądowo-napięciowa diody AN-301A



Rys. 6. Eksperymentalne charakterystyki prądowo-napięciowe z ciśnieniem jako parametrem diody AN-301A



Rys. 7. Eksperymentalne charakterystyki prądowo-napięciowe z temperaturą jako parametrem diody AN-301A

na rys. 5 charakterystyki prądowo-napięciowej można wyznaczyć punkty ekstremalne, których współrzędne wykorzystano w programie obliczeń numerycznych. Dla badanych diod z arsenku galu typu AN — 301A prąd szczytu I_p, którego wartość równa jest około 2 mA, występuje przy wartości napięcia U_p równej około 100 mV. Prąd doliny I_v, którego wartość równa jest około 0,2 mA, występuje przy wartości napięcia U_v równej około 450 mV. Rozrzut współrzędnych punktu szczytu badanych diod nie przekraczał wartości 3%, a punktu doliny — wartości 5%.

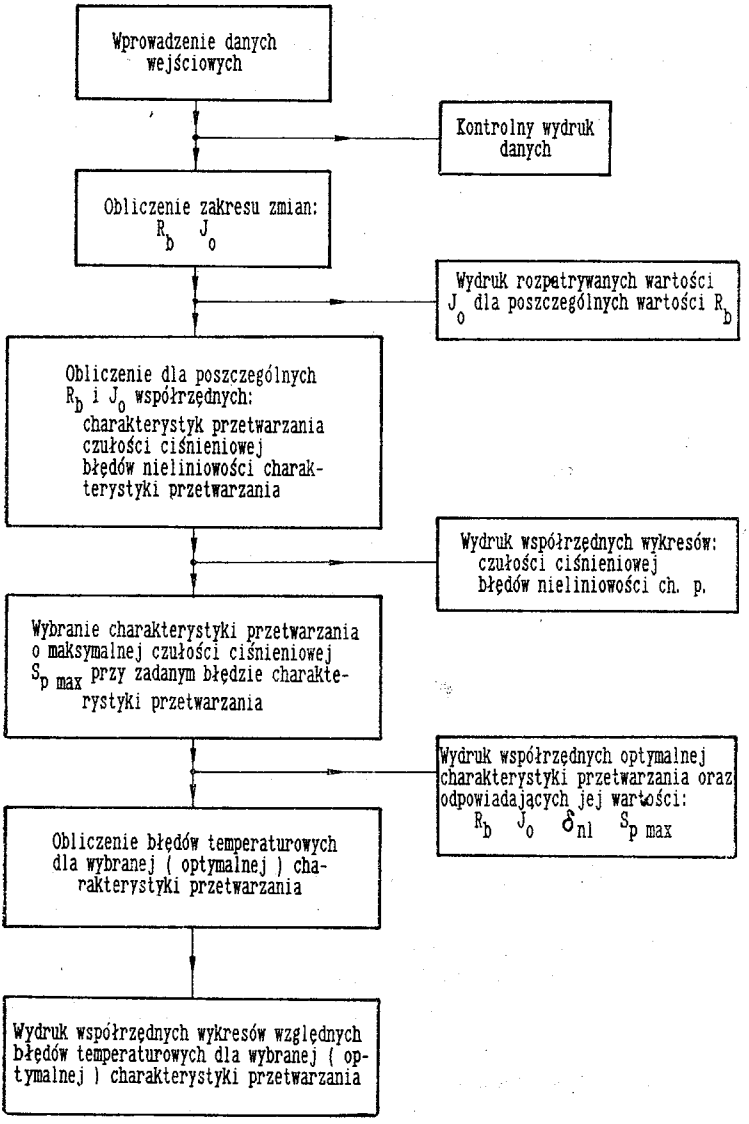
Eksperymentalne charakterystyki ciśnieniowe przedstawia rys. 6, z którego wynika, że dla badanych diod z arsenku galu ze wzrostem ciśnienia maleje prąd diody, a więc przyrosty prądu są ujemne. Ciśnieniowe względne zmiany prądu badanych diod w pobliżu punktu szczytu, który przyjęto jako punkt porównawczy, wynosiły około $1\%/10^7$ Pa.

Wpływ zmiany temperatury na zmianę prądu diody ma złożony charakter. Wartość i znak wartości przyrostu prądu pod wpływem zmiany temperatury zależy zarówno od napięcia diody, jak i od wartości temperatury. Wykresy pomiarowych charakterystyk temperaturowych pokazano na rys. 7.

Z rysunku tego wynika, że występują tu lokalne punkty ekstremalne, w których wartości prądu są funkcją napięcia i temperatury. Temperaturowe zmiany prądu diody w pobliżu punktu szczytu wynoszą około 0,17%/10°C w zakresie dodatnich przyrostów

temperatur, natomiast w zakresie ujemnych przyrostów temperatur — około 1,5%/10°C. Uzyskane wyniki z pomiarów kilkunastu badanych diod wykorzystano do obliczeń numerycznych.

Uproszczony schemat blokowy wykonywanych obliczeń przedstawiono na rys. 8. Program opracowano w języku FORTRAN. Jako główne dane wejściowe w programie obliczeń przyjęto: współrzędne eksperymentalnych charakterystyk prądowo-napięciowych (podstawowej oraz z ciśnieniem i temperaturą jako parametrem), współrzędne punktu szczytu podstawowej charakterystyki prądowo-napięciowej, zakres mierzonego ciśnienia, zakres temperatury pracy przetwornika.

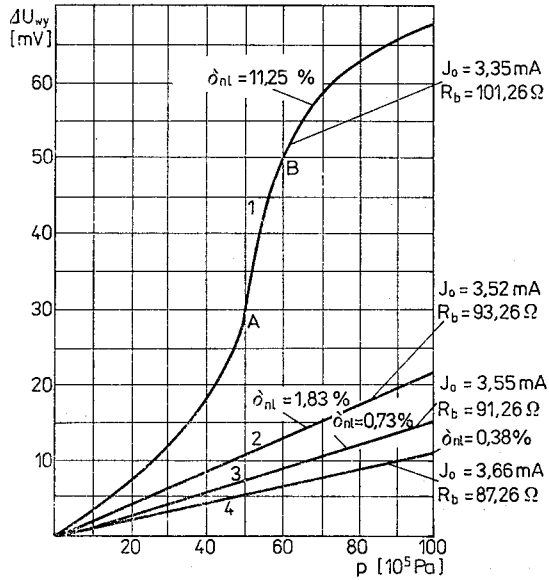


Rys. 8. Uproszczony schemat blokowy obliczeń numerycznych

Z zaprogramowanych obliczeń otrzymano m.in. wartość współrzędnych: charakterystyk przetwarzania dla różnych dopuszczalnych wartości błędów nieliniowości, czułości ciśnieniowej oraz błędu nieliniowości w funkcji prądu polaryzacji i rezystancji bocznika, względnego błędu temperaturowego w funkcji ciśnienia i temperatury.

Ponadto, przyjmując jako kryterium największą czułość ciśnieniową przy zadanej dopuszczalnej wartości błędu nieliniowości, zaprogramowano wydruki wartości współrzędnych charakterystyki przetwarzania (nazwano ją optymalną) oraz wartości rezystancji bocznika, prądu polaryzacji i czułości ciśnieniowej. Rezystancja bocznika i prąd polaryzacji są parametrami projektowymi, natomiast błąd nieliniowości i czułość ciśnieniowa określają właściwości metrologiczne przetwornika. Podane wyżej kryterium może stanowić zasadę doboru punktu pracy przetwornika.

Graficzną ilustrację obliczeń numerycznych przedstawiono na rysunkach: 9, 10, 11, 12. Na rys. 9 pokazano charakterystyki przetwarzania odpowiadające różnym punktom pracy (I_0, R_b). Z przedstawionych wykresów wynika, że gdy nie stawia się wymagań

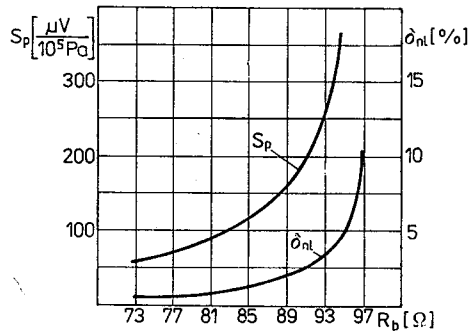


Rys. 9. Obliczeniowe charakterystyki przetwarzania o różnych błędach nieliniowości

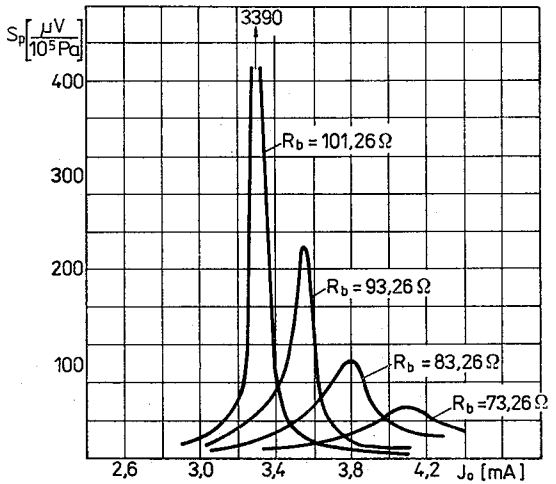
dotyczących błędów nieliniowości, to można uzyskać bardzo dużą czułość ciśnieniową przetwornika. Duża czułość występuje na odcinku charakterystyki od punktu A do punktu B. Zależność tę przedstawia krzywa 1. Kształt tej charakterystyki czyni ją mało użyteczną w założonym zakresie pomiarowym ciśnienia.

Przyjmując dopuszczalną wartość błędu nieliniowości (w programie obliczeń założono: 2; 1; 0,5%) oraz dobierając punkt pracy tak, aby uzyskać maksymalną czułość ciśnieniową — otrzymano charakterystyki oznaczone na rys. 9 odpowiednio: 2, 3, 4. Przedstawione wykresy charakterystyk pokazują, że czułość ciśnieniowa zależy od dopuszczalnego błędu nieliniowości. Im większa jest jego wartość, tym większa jest możliwa do uzyskania czułość ciśnieniowa.

Wpływ rezystancji bocznika na czułość ciśnieniową przedstawiony jest na rys. 10, z którego wynika, że wartość czułości ciśnieniowej rośnie wraz ze wzrostem wartości rezystancji bocznika. Osiąga ona bardzo duże wartości, gdy wartość rezystancji bocznika zbliża się do minimalnej wartości rezystancji dynamicznej diody $r_{d\,\min}$. Brana jest przy tym pod uwagę wartość bezwzględna rezystancji $|r_{d\,\min}|$ należącej do obszaru, w którym funkcja $I_d = f_d\,(U_d, T_0, p_0)$ jest malejąca. Na tle charakterystyki czułości ciśnieniowej wykreślona jest na rys. 10 zależność wartości błędu nieliniowości od wartości rezystancji bocznika. Z wykresów tych wynika podany już wcześniej wniosek, że wartość czułości ciśnieniowej rośnie wraz ze wzrostem wartości błędu nieliniowości charakterystyki przetwarzania.



Rys. 10. Obliczeniowe wykresy zależności czułości ciśnieniowej i błędu nieliniowości od rezystancji bocznika

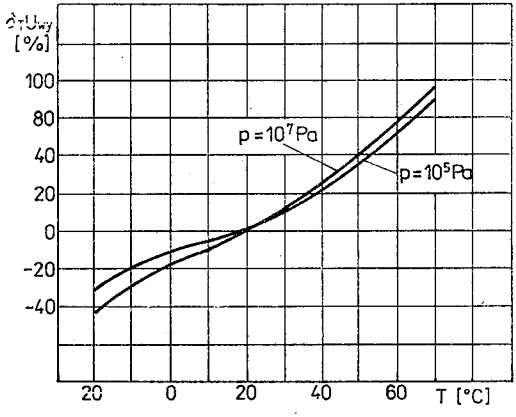


Rys. 11. Obliczeniowe wykresy zależności czułości ciśnieniowej od prądu polaryzacji i rezystancji bocznika

Interesujące, z punktu widzenia wymagań dotyczących stabilizacji wartości prądu polaryzacji, są wykresy pokazane na rys. 11. Zależność czułości ciśnieniowej od prądu polaryzacji, przy stałej wartości rezystancji bocznika, ma maksymalną wartość przy określ-

lonej wartości prądu polaryzacji. Wykresy te mają dużą stromość w pobliżu punktu ekstremalnego — tym większą, im większa jest maksymalna wartość czułości ciśnieniowej. Wybranie tego punktu jako punktu pracy wymagałoby zastosowania źródła prądu polaryzacji o dużym współczynniku stałości wartości prądu. Przy mniejszych ekstremalnych wartościach czułości ciśnieniowej wymóg ten jest łagodniejszy.

Zależność względnego błędu temperaturowego od temperatury przedstawia rys. 12. Wykresy te otrzymano przyjmując (w programie obliczeń) jako kryterium wyboru punktu pracy: maksymalną czułość ciśnieniową przy dopuszczalnej wartości błędu nieliniowości charakterystyki przetwarzania $\delta_{nl} \leqslant 0{,}5\%$ (wartość czułości ciśnieniowej wynosiła $S_p = 65\ \mu V/10^5$ Pa).



Rys. 12. Obliczeniowe wykresy względnych błędów temperaturowych

Przedstawione zależności wykazują, że prosty układ równoległy o polaryzacji prądowej obarczony jest (w rozpatrywanym zakresie pomiarowym ciśnienia od 10^5 do $100 \cdot 10^5$ Pa i przyjętym zakresie temperaturowy pracy od $-40°C$ do $+70°C$) dużymi błędami temperaturowymi przybierającymi wartości dochodzące do 100%. Wynika stąd wniosek, że układ ten nie może być wykorzystany do pomiaru ciśnienia rzędu 10^7 Pa. W tym przypadku należy zastosować bardziej rozbudowany układ przetwornika — np. układ różnicowy.

Przedstawione rozważania oparte na wynikach obliczeń komputerowych były weryfikowane eksperymentalnie na modelu przetwornika. Przeprowadzone eksperymenty potwierdziły zgodność wyników obliczeń z wynikami pomiarów.

Z przedstawionych rozważań wynikają następujące główne wnioski:

1. Właściwości metrologiczne (czułość ciśnieniowa, błąd nieliniowości charakterystyki przetwarzania) przetwornika ciśnienia z diodą tunelową zależą od wyboru punktu pracy na charakterystyce prądowo-napięciowej.

2. Wyboru punktu pracy nie można dokonać na drodze analitycznej, ponieważ wzory opisujące charakterystyki diody są złożone i mają przybliżony charakter.

3. W celu wyznaczenia optymalnego położenia punktu pracy można korzystać z obliczeń komputerowych opartych na charakterystykach eksperymentalnych.