ROZPRAWY ELEKTROTECHNICZNE 1982, 28, z. 1-2, ss. 201-217

621.317.31.014.33

Zastosowanie udarów termicznych do przyspieszonego oszacowania charakterystyk niezawodności mikroelektronicznych hybrydowych układów grubowarstwowych

ZBIGNIEW WĄSOWICZ (KRAKÓW), PIOTR WŁOKA (KRAKÓW)

Instytut Elektroniki Akademii Górniczo Hutniczej

Otrzymano 2.6.1980

Traktując mikroukład hybrydowy jako element i określając uszkodzenie jako wyjście jego parametrów ze zbioru stanów zdatności, pokazano realizację programu przyspieszonych badań trwałości z zastosowaniem udarów termicznych. Wprowadzono przekształcenie oparte na zmodyfikowanym równaniu Arrheniusa, które pozwala na oszacowanie realizacji statystyk pozycyjnych trwałości w warunkach rzeczywistych. Zaproponowano jednocześnie odpowiednią metodę analizy statystycznej, dającą się prowadzić przy użyciu ETO.

1. WSTĘP

Idea opracowanej metody zrodziła się na tle współpracy Instytutu Elektroniki AGH z jednym z zakładów produkujących podzespoły elektroniczne. Przedstawiona metoda daje możliwość oszacowania charakterystyk niezawodności $\lambda(t)$ w czasie $4 \div 8$ tygodni, potrzebnych na przygotowanie badań ($3 \div 4$ tygodni), przeprowadzenie badań przyspieszonych ($1 \div 2$ tygodni), dokonanie statystycznej analizy wyników i diagnostyki ($0{,}1 \div 2$ tygodni). Metoda może być również użyteczna dla placówek naukowo-badawczych oraz dla odbiorców mikroukładów celem sprawdzenia trwałości zakupionej partii. Jej stosowanie umożliwia szybkie otrzymanie potrzebnej informacji; jest więc praktyczna i ekonomiczna.

2. UZASADNIENIE I CEL BADAŃ

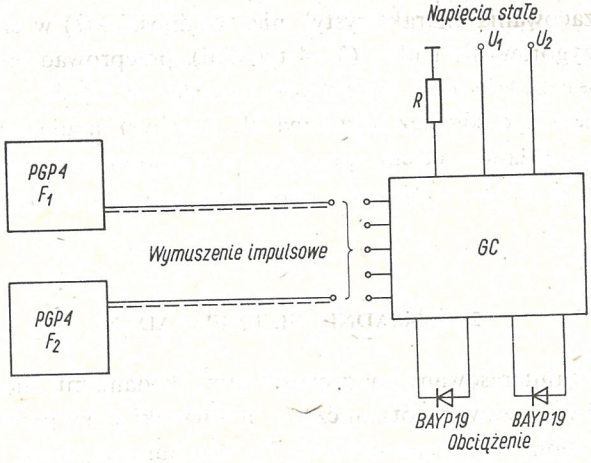
Producent jest zainteresowany przyspieszonymi badaniami niezawodności, które dają się zrealizować w możliwie krótkim czasie. Ich wyniki są co prawda jedynie przybliżone, jednak w określonym zakresie obrazują mechanizm uszkodzeń oraz trwałość układów elektronicznych eksploatowanych zgodnie z przeznaczeniem. Informacja o zależności intensywności uszkodzeń $\lambda(t)$ od czasu daje możliwość oceny niezawodnościowych własności badanych układów, pozwala na wyciągnięcie wniosków dotyczących ich konstrukcji i technologii oraz orientuje o niezawodności systemu zbudowanego z tych elementów. I tak na przykład układy charakteryzujące się okresem adaptacji [1] należy poddać

selekcji (starzeniu adaptacyjnemu) lub korekcji konstrukcyjno-technologicznej. Proces badań daje również możliwość obserwacji diagnostycznych uszkodzonych układów, zwłaszcza w przypadku badań układów niehermetyzowanych.

Poniżej przedstawiono metodę transformacji umożliwiającą przeliczenie wyników badań przyspieszonych na wartości zbioru w warunkach rzeczywistych. Oparto się więc na odpowiedniej metodyce badań i analizy statystycznej, korzystając z wyników własnych badań teoretycznych i empirycznych oraz z literatury [1], [2], [3].

3. CHARAKTERYSTYKA BADAŃ

W badaniach przeprowadzonych na hybrydowych mikroukładach *GC*, uwzględniając empiryczną zależność prędkości zużywania się elementów elektronicznych od temperatury, w postaci równania Arrheniusa, przyjęto ten czynnik jako wymuszenie zużycia mikroukładów. Ze względu na hermetyzację indywidualną elementów składowych i całościową mikroukładu uznano wilgotność jako czynnik odgrywający w tym przypadku mniejszą rolę. Natomiast temperatura jest również czynnikiem wymuszającym w trakcie eksploatacji mikroukładów zgodnie z ich przeznaczeniem. Stąd więc do przyspieszonych badań trwałości wymienionych układów zastosowano metodę szoków temperaturowych jako stosunkowo szybką i efektywną [5], [6]. Stwierdzono również, na podstawie badań własnych i obserwacji użytkowników układów scalonych, że ich magazynowanie powoduje pogorszenie charakterystyk niezawodnościowych [4], [6], w stosunku do charakterystyk układów wyprodukowanych w tym samym okresie, lecz zasilanych i pracujących zgodnie z przeznaczeniem. W trakcie badań mikroukładów *GC* zastosowano ich zasilanie, obciążenie wyjścia zgodnie z warunkami technicznymi (*WT*) oraz podawano wymuszenie impulsowe (rys. 1).



Rys. 1. Schemat blokowy ramy trwałości

Takie potraktowanie procesu badań narzuca modyfikację przyjętego w transformacji równania Arrheniusa [2], co przeprowadzono w p-cie 8. Dokonane przekształcenie, dotychczas nieznane w literaturze, pozwoliło pomyślnie rozwiązać postawione zadanie. Na podsta-

wie badań własnych przyjęto tu, że o niezawodności układów elektronicznych decydują głównie elementy półprzewodnikowe. Założono następnie, że rozkład trwałości badanych układów jest rozkładem potęgowym, co potwierdziły badania. Rozkład ten uwzględnia procesy adaptacji i zużycia, przy równoczesnym ograniczeniu zbioru realizacji statystyk pozycyjnych trwałości w przedziale $(0, b)$, gdzie b jest parametrem skali.

4. ZAKRES BADAŃ

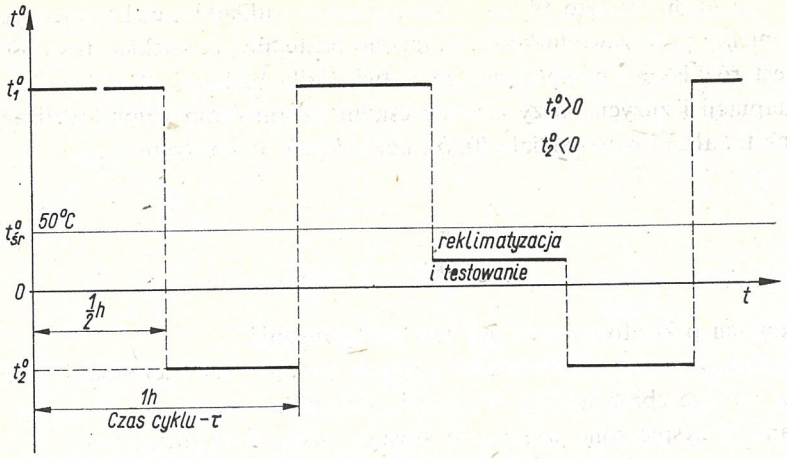
W badaniach przyjęto następujący tok postępowania.

- Ustalono warunki badań poszczególnych próbek, ich liczność i badane wersje wykonania, a także zbiór stanów zdatności mikroukładów.
- Badanie przyspieszone wykonano stosując szoki temperaturowe.
- Dokonano transformacji wyników badań przyspieszonych na wartości zbioru w warunkach rzeczywistych.
- Przeprowadzono analizę statystyczną uzyskanych wyników badań.
- Wykonano badania diagnostyczne uszkodzonych mikroukładów wyciągając odpowiednie wnioski.
- Sprecyzowano zalecenia dotyczące korekty konstrukcyjno-technologicznej badanych układów.

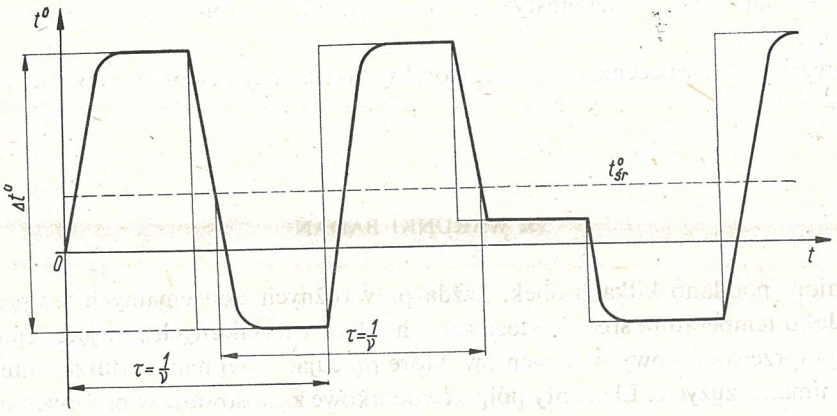
5. WARUNKI BADAŃ

Badaniom poddano kilka próbek, każdą przy różnych ekstremalnych temperaturach szoków. Jako temperaturę średnią stosowanych udarów termicznych przyjęto temperaturę złącza t_z półprzewodnikowych elementów, które pracując w tej temperaturze winny wykazywać minimalne zużycie. Elementy półprzewodnikowe zastosowane w badanych układach osiągają w czasie pracy w warunkach znamionowych temperaturę złącza $t_z = (40 \div 50)°C$. Stąd dla wszystkich badanych próbek średnią temperaturę szoków przyjęto $t_{śr} = 50°C$. Prędkość zużycia mikroukładów rośnie z różnicą ekstremalnych temperatur szoku oraz ich częstotliwością. Dla zachowania w czasie badań identycznego mechanizmu uszkodzeń elementów wszystkich próbek określono ekstremalne temperatury oraz czas trwania cykli. W przypadku badanych układów temperaturę maksymalną t_1 przyjęto niższą od temperatury topliwości (183°C) lutowia standardowego Sn-Pb, zastosowanego w procesie technologii. Natomiast temperatura minimalna t_2 szoku winna być wyższa od minimalnej temperatury (−75°C), w której element półprzewodnikowy jeszcze pracuje poprawnie. Półokres szoku (czas nagrzewania lub chłodzenia układu) winien być dłuższy od czasu uzyskania przez próbkę temperatury otoczenia, co zależne jest od mocy stosowanego urządzenia klimatycznego. Praktycznie czas ten wynosił około 6 minut.

Celem zachowania powtarzalności badań i zmniejszenia wpływu stanów przejściowych, przyjęto czas półokresu równy 30 minut. Stąd czas trwania cyklu równy jest 60 minut, co upraszcza proces badań i obliczeń (rys. 2, rys. 3).



Rys. 2. Zmiany temperatury otoczenia w czasie badań próbek o numerach $l = 1, 2, 3, 4, 5$



Rys. 3. Zależność zmian temperatury wnętrza badanych mikroukładów od czasu

6. DOBÓR PRÓBEK

Spośród wyodrębnionych losowo 300 szt. układów, wyprodukowanych przez obydwie zmiany robocze w ciągu tygodnia, wybrano losowo trzy próbki układów hermetyzowanych, każda o liczności $n = 18$ szt. i dwie próbki układów niehermetyzowanych o liczności $n = 20$ szt. Charakterystyki badań próbek zawiera tabl. 1.

Hipotezę o jednorodności zbioru próbek zweryfikowano testem statystycznym Kruskala-Wallisa [1], zastosowanego dla zbioru wyników badań przeliczonych na wartości rzeczywiste. W wyniku przeprowadzonego testu na poziomie istotności $\alpha = 0{,}1$ stwierdzono brak podstaw do odrzucenia wysuniętej hipotezy. W trakcie badań przeprowadzono testowanie kontrolne na stanowisku diagnostycznym, gdzie mikroukład, którego parametry nie mieściły się w zbiorze wymagań ustalonych przez *WT*, traktowano jako uszkodzony.

Tablica 1

Charakterystyka badanych próbek mikroukładów *GC*

l — numer próbki	Liczność próbki (sztuk)	Badane mikroukłady	Warunki badań		Uwagi
			Granice szoku [°C]	$\Delta t°$ szoku [°C]	
$l = 1$	$n_1 = 18$	hermetyzowane	+170, −70	240	zasilane
$l = 2$	$n_2 = 18$	,,	+150, −50	200	,,
$l = 3$	$n_3 = 18$	,,	+130, −30	160	,,
$l = 4$	$n_4 = 20$	odkryte	+150, −50	200	,,
$l = 5$	$n_5 = 20$	,,	+130, −30	160	,,

7. TOK BADAŃ METODĄ SZOKÓW TEMPERATUROWYCH

Zmieniając dla każdej z próbek temperatury ekstremalne (tabl. 1), przeprowadzono badania w następującej kolejności.

1. Próbkę mikroukładów zamocowanych w ramie trwałości oraz zasilanych umieszczono w komorze o temperaturze t_1 na 30 minut.

Tablica 2

Realizacje statystyk pozycyjnych trwałości próbek mikroukładów *GC* badanych przy zastosowaniu szoków temperaturowych o parametrach jak w tabl. 1

	Czasy do kolejnych uszkodzeń $t_{(m)}$ [godz.]				
$i = m$ \ l	1	2	3	5	6
1	15	32	31	4	3
2	15	50	31	24	59
3	17	52	32	34	63
4	36	52	51	42	77
5	36	53	75	49	77
6	46	53	81	56	82
7	46	59	83	62	106
8	51	60	106	67	108
9	51	64	114	67	109
10	53	64	118	69	111
11	53	65	120	73	114
12	53	71	128	73	125
13	55	83	130	75	125
14	61	87	133	77	126
15	66	92	137	79	126
16	—	—	—	81	134
17	—	—	—	—	138
18	—	—	—	—	—
19				—	—
20				—	—

„—" oznacza, że nie nastąpiło następne uszkodzenie w próbce w trakcie badań,
i — numer kolejnego uszkodzenia *GC* w próbce o numerze l,
m — liczba układów uszkodzonych w przedziale czasu $(0, t_{(m)})$

2. Nagrzaną próbkę przemieszczono do komory o temperaturze t_2 na okres 30 minut.

3. Badaną próbkę umieszczono ponownie w komorze o temperaturze t_1 na 30 minut.

4. Próbkę o temperaturze t_1 poddano reklimatyzacji (25 minut), po czym przeprowadzono testowanie podanych układów, usuwając uszkodzone z próbki.

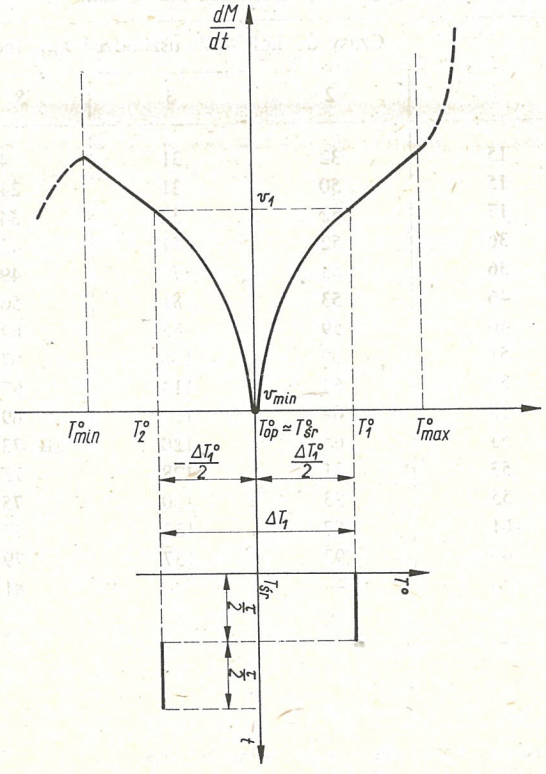
5. Czynności 2 do 4 powtarzano aż do uszkodzenia się około 80% układów badanej próbki.

W oparciu o wyniki testowania uzyskano realizacje statystyk pozycyjnych trwałości (tabl. 2).

8. TRANSFORMACJA WYNIKÓW BADAŃ PRZYSPIESZONYCH NA WARTOŚCI ZBIORU W WARUNKACH RZECZYWISTYCH

Równanie Arrheniusa [2] przedstawia prędkość starzenia się mikroukładów magazynowanych w zależności od temperatury. W prowadzonych badaniach układy były jednocześnie zasilane, a więc w zaistniałych warunkach badań należy wziąć pod uwagę pewne nowe aspekty fizyczne, wymienione niżej.

1. Istotnie optymalnej temperatury pracy t_{op} mikroukładów, w której ich prędkość zużycia jest najmniejsza. Można przypuszczać (patrz punkt 4), że temperatura t_{op} będzie



Rys. 4. Zależność prędkości zużycia mikroukładów zasilanych od temperatury otoczenia

zbliżona do temperatury pracy t_z złączy elementów półprzewodnikowych. Stąd przyjęto $t_{op} = t_{śr} = t_z$.

2. Eksploatacja mikroukładów w temperaturze odbiegającej od optymalnej zwiększa prędkość ich zużycia.

3. Funkcja opisująca procesy zużycia układów zasilanych może być potraktowana z pewnym przybliżeniem jako parzysta w przedziale temperatur (t_1, t_2). Ułatwia to operację transformacji.

4. Oś symetrii funkcji jest prostą równoległą do osi rzędnych i przechodzącą przez punkt odciętej $t_{op} = t_{śr}$ (patrz rys. 4).

5. Przyjmując, że oscylacje temperatur przebiegają wokół $t_{śr} = t_{op}$ z amplitudą $\frac{\Delta T}{2}$ i okresem τ (rys. 2), równanie Arrheniusa przyjmuje postać

$$\frac{dM}{dt} = A\exp\left[-\frac{B}{t_{śr} \pm \frac{\Delta T}{2}}\right], \tag{1}$$

gdzie A, B — stałe.

Dla $\Delta T = 0$

$$\frac{dM}{dt} = v_{min} = A\exp\left[-\frac{B}{t_{śr}}\right].$$

6. Przyjmując początek układu współrzędnych w punkcie $(t_{śr}, v_{min})$, równanie (1) można zapisać (rys. 4) w postaci

$$\frac{dM}{dt} = A\exp\left[\mp\frac{2B}{\Delta T}\right]. \tag{2}$$

Całka funkcji (2) w przedziale $(0, \tau]$ przedstawia liczbę uszkodzonych układów w tym przedziale

$$\Delta M_\tau = A\exp\left[\pm\frac{2B}{\Delta T}\right]\tau. \tag{3}$$

W dowolnym czasie $t = n\tau$ ilość uszkodzonych układów

$$\Delta M_t = m = A\exp\left[\pm\frac{2B}{\Delta T}\right]t. \tag{4}$$

Po obustronnym zlogarytmowaniu równania (4) otrzymuje się

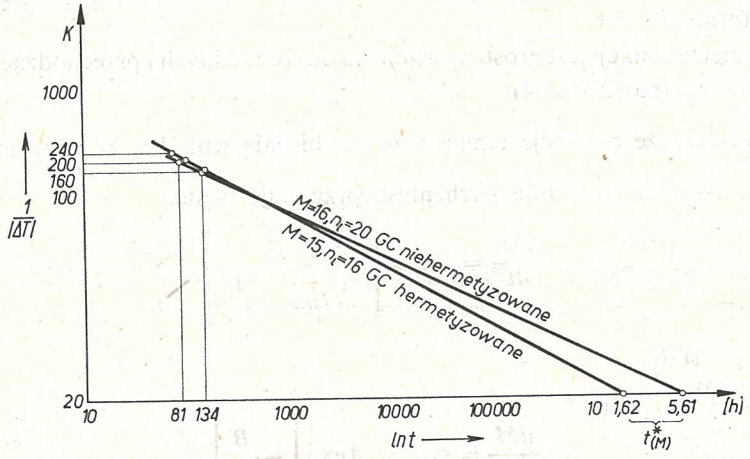
$$\frac{1}{\Delta T} = A_1 \ln t + B_1, \tag{5}$$

gdzie $A_1 = \frac{1}{2B}$ oraz $B_1 = \frac{\ln A - \ln m}{2B}$.

Zależność (5) jest liniowa w siatce funkcyjnej o odciętej $\xi = \ln t$ i rzędnej $\eta = 1/\Delta T$. W przyjętym układzie współrzędnych, dla zaobserwowanej ilości m uszkodzeń mikroukładów, otrzymuje się rodzinę prostych równoległych o równaniu

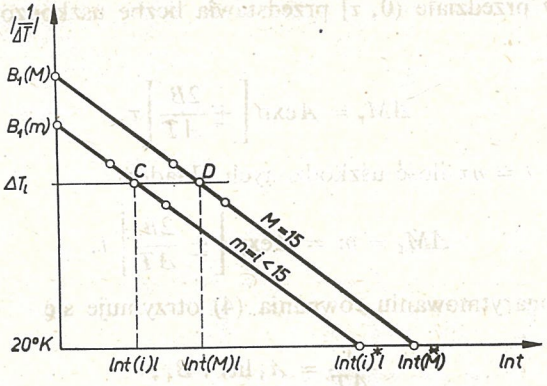
$$\eta = A_1\xi + B_1(m). \tag{6}$$

Proste te ilustrują w jakim czasie przy danej różnicy temperatur ΔT ulegnie uszkodzeniu określona liczba układów m. Stosując siatkę i znając wyniki badań kilku próbek (minimum 2), można oszacować graficznie realizacje statystyk pozycyjnych trwałości mikroukładów w warunkach rzeczywistych (rys. 5). Realizacje te można również wyznaczyć analitycznie, jeśli oszacuje się wartość przelicznika p^*.



Rys. 5. Oszacowanie graficzne nieznanej trwałości rzeczywistej ok. 80% w próbce mikroukładów *GC*

W tym celu wykreślono prostą wyrażoną równaniem (6) dla realizacji statystyk pozycyjnych trwałości o najwyższej randze $t_{(M)}$ próbek mikroukładów. Dla $\Delta T = 20°K$ odczytuje się szacunkową wartość transformowaną $t^*_{(M)}$ — rys. 6. Korzystając z równości trój-



Rys. 6. Transformacja realizacji statystyk pozycyjnych trwałości mikroukładów

kątów: C, $\ln t_{(i),t}$, $\ln t^*_{(i),t}$ i D, $\ln t_{(M),t}$, $\ln t^*_{(M)}$ otrzymuje się

$$t^*_{(i),t} = t_{(i),t} \frac{t^*_{(M)}}{t_{(M),t}}$$

oraz

$$t^*_{(i),t} = p^* t_{(i),t}, \tag{7}$$

gdzie:

$t_{(i),l}$ — realizacja statystyki pozycyjnej trwałości o randze (i) próbki l,
$t^*_{(i),l}$ — transformacja realizacji statystyki j.w.

9. ANALIZA STATYSTYCZNA UZYSKANYCH WYNIKÓW BADAŃ

Analizę można przeprowadzić dwiema metodami [1]:

1 — wykorzystując możliwości graficznego oszacowania parametrów hipotetycznego rozkładu trwałości,
2 — wysuwając hipotezę o potęgowym rozkładzie trwałości i stosując zależności analityczne dla tegoż rozkładu.

Metoda pierwsza wymaga przeliczeń wszystkich, uzyskanych realizacji statystyk pozycyjnych trwałości na warunki rzeczywiste. Metoda druga daje pożądane efekty wykorzystując jedynie transformowaną realizację najwyższej rangą statystyki pozycyjnej trwałości $t^*_{(M)}$. Ponadto metoda ta jest mniej czasochłonna i daje możliwość stosowania elektronicznej techniki obliczeniowej. Zastosowano zatem metodę drugą.

Matematyczny model rozkładu trwałości [1]

1) Zakładając potęgowy rozkład trwałości dokonano estymacji parametrów kształtu δ rozkładów trwałości próbek $l = 1, 2, 3, 4, 5$ korzystając z zależności

$$\delta^*_{M,l} = \left[\ln t_{(M),l} - \frac{1}{M}\sum_{i=1}^{M} \ln t_{(i),l}\right]^{-1}. \tag{8}$$

Jak widać — do obliczenia parametru kształtu potęgowego rozkładu trwałości dowolnej próbki l nie są potrzebne wartości transformowane $t^*_{(i),l}$. Wynika to z liniowego charakteru zależności (6) i (7) oraz własności estymatora $\hat{\delta}_M$. Dla badanych próbek $l = 1, 2, 3$ układów hermetyzowanych uzyskano następujące wartości parametrów kształtu:

$$\delta^*_{M,1} = \delta^*_{15,1} = 1{,}9535,$$
$$\delta^*_{M,2} = \delta^*_{15,2} = 2{,}398,$$
$$\delta^*_{M,3} = \delta^*_{15,3} = 1{,}9389.$$

Dla próbek niezhermetyzowanych o numerach $l = 4, 5$, których badania przeprowadzono dla celów porównawczych i diagnostycznych, uzyskano następujące wartości oszacowań parametrów kształtu:

$$\delta^*_{M,4} = \delta^*_{16,4} = 4{,}418651,$$
$$\delta^*_{M,5} = \delta^*_{16,5} = 4{,}418522.$$

2) Dla założonego rozkładu trwałości wyznaczono oszacowania parametrów skali b próbek o numerach $l = 1, 2, 3, 4, 5$. Korzystając ze wzoru:

$$b^*_{M,l} = t^*_{(M)} \left(\frac{n}{M}\right)^{1/\delta^*_{M,l}}, \tag{9}$$

uzyskano wartości:

$$b^*_{M,1} = b^*_{15,1} = 1{,}76256 \cdot 10^6 \text{ godz.},$$
$$b^*_{M,2} = b^*_{15,2} = 1{,}74798 \cdot 10^6 \text{ godz.},$$
$$b^*_{M,3} = b^*_{15,3} = 1{,}7739 \cdot 10^6 \text{ godz.},$$
$$b^*_{M,4} = b^*_{16,4} = 15{,}0348 \cdot 10^6 \text{ godz.},$$
$$b^*_{M,5} = b^*_{16,5} = 15{,}0347 \cdot 10^6 \text{ godz.}$$

Z uzyskanych rezultatów wynika, że układy niehermetyzowane cechuje lepsza charakterystyka $\lambda(t)$.

3) Zweryfikowano hipotezę o potęgowym rozkładzie trwałości dla wspomnianych próbek. W rozpatrywanym przypadku ilość uszkodzonych układów $M > 12$, zatem wymieniony test wykorzystujący własności funkcji wiodącej intensywności $\Lambda(t)$ opierać się będzie na statystyce Cochrana. Stwierdzono brak podstaw do odrzucenia hipotezy o potęgowym rozkładzie trwałości.

Poniżej przytoczono postępowanie weryfikacyjne wysuniętej hipotezy dla próbki $l = 1$. Hipotezę o rodzaju rozkładu trwałości zweryfikowano korzystając z uzyskanych realizacji statystyk pozycyjnych trwałości, co wynika z własności transformaty wyznaczonej w punkcie 8. Parametry skali rozkładu potęgowego wyznaczono korzystając ze wzoru (9). Podstawiając w miejsce $t^*_{(M)}$ realizacje $t_{(M)}$ uzyskano wartości $b_{M,l} = b_l$: $b_1 = 72$ h, $b_2 = 99$ h, $b_3 = 150$ h, $b_4 = 86$ h, $b_5 = 142$ h.

Dla rozkładu potęgowego słuszne są zależności:

funkcja wiodąca

$$\Lambda(t_{(m)}) = -\ln R(t_{(m)}) = -\ln\left[1-\left(\frac{t_{(m)}}{b}\right)^{\delta}\right], \tag{10}$$

gdzie:

δ — parametr kształtu,

b — parametr skali,

oraz statystyka

$$U_m = (n_1 - m + 1)\left\{-\ln\left[1-\left(\frac{T_{(m)}}{b_l}\right)^{\delta}\right] + \ln\left[1-\left(\frac{T_{(m-1)}}{b_l}\right)^{\delta}\right]\right\} \tag{11}$$

dla: $m = 1, 2, 3, 4, \ldots, 15$,

$M_1 = 15$ i $n_1 = 18$.

Znając wartości $\delta^*_{15,1} = \delta^*_1$ i $b_{15,1} = b_1$ obliczono poszczególne wartości u_m statystyki U_m. Wyniki obliczeń u_m ujęto w tablicy 3, z której odczytujemy:

$$\max(u_1, u_2, u_3, \ldots, u_{15}) = 3{,}387$$

oraz

$$\sum_{m=1}^{15} u_m = 14{,}9604;$$

stąd

$$g = \frac{\max(u_1, u_2, u_3, \ldots, u_{15})}{\sum_{m=1}^{15} u_m} = 0{,}2264. \tag{12}$$

Tablica 3

Tablica pomocnicza do weryfikacji hipotezy o potęgowym rozkładzie trwałości próbki $l = 1$

m	$\frac{t_{(m),1}}{b_1}$	A	B	u_m
1	0,2083	0,0408		0,7344
2	0,2083	0,0408	0	0
3	0,2361	0,0619	0,0211	0,3376
4	0,5	0,2877	0,2258	3,387
5	0,5	0,2877	0	0
6	0,6389	0,5276	0,2399	3,1187
7	0,6389	0,5276	0	0
8	0,7083	0,6931	0,1655	1,8202
9	0,7083	0,6931	0	0
10	0,7361	0,7765	0,0834	0,7506
11	0,7361	0,7765	0	0
12	0,7361	0,7765	0	0
13	0,7639	0,8675	0,0910	0,5460
14	0,8472	1,2730	0,4055	2,0275
15	0,9167	1,8320	0,5596	2,2384
				$\sum_{m=1}^{15} u_m = 14{,}9604$

$$A = -\ln\left[1-\left(\frac{t_{(m)}}{b_1}\right)^{\delta_1}\right],$$

$$B = \ln\left[1-\left(\frac{t_{(m-1)}}{b_1}\right)^{\delta_1}\right]-\ln\left[1-\left(\frac{t_{(m),1}}{b_1}\right)^{\delta_1}\right]$$

Przyjmując poziom istotności $\alpha = 0{,}05$ z tablicy XXIII [1], odczytano wartość kwantyla statystyki Cochrana o parze $(M, 2)$ stopni swobody, przy czym $g_{M,2,1-\alpha} = g_{15,2,0,95} = 0{,}3346$. Na podstawie wyników obliczeń stwierdzono, że nierówność

$$\frac{\max(u_1, u_2, \ldots, u_M)}{u_1+u_2+\ldots+u_M} \geqslant g_{M,2,1-\alpha} \tag{13}$$

nie jest spełniona i wobec tego nie ma podstaw do kwestionowania słuszności hipotezy, że trwałość mikroukładów GC w próbce $l = 1$ ma rozkład potęgowy.

Do takich samych stwierdzeń dochodzi się stosując test Firkowicza [1] dla pozostałych próbek o numerach $l = 2, 3, 4, 5$.

4) Po udowodnieniu, że rozkład trwałości mikroukładów jest rozkładem potęgowym, dokonano dokładniejszej oceny nieznanych parametrów δ i b z próbki łącznej, oddzielnie dla układów hermetyzowanych i niehermetyzowanych. Wymaga to uprzedniego zweryfikowania hipotezy o jednakowym potęgowym rozkładzie trwałości w rozważanych próbkach, w oparciu o równość parametrów kształtu i parametrów skali dla próbek układów hermetyzowanych:

$$\begin{aligned} \delta_1 &= \delta_2 = \delta_3, \\ b_1 &= b_2 = b_3, \end{aligned} \tag{14}$$

oraz dla próbek układów odkrytych:

$$\delta_4 = \delta_4, \qquad b_4 = b_5. \tag{15}$$

Skorzystano tu z kryterium Firkowicza—Hartley'a [1], przy czym hipotezę tę należy odrzucić jeśli jest spełniona nierówność

$$\frac{\max(\delta_1^*, \delta_2^*, \delta_3^*)}{\min(\delta_1^*, \delta_2^*, \delta_3^*)} \geqslant \eta_{K,2(n-1),1-\alpha}, \tag{16}$$

w której $\eta_{K,2(n-1),1-\alpha}$ jest kwantylem rzędu $1-\alpha$ statystyki Hartley'a η o parze $[K,2(n-1)]$ stopni swobody, zaś K oznacza liczbę próbek.

W rozważanym przypadku uzyskano

$$\frac{\max(\delta_{M,1}^*, \delta_{M,2}^*, \delta_{M,3}^*)}{\min(\delta_{M,1}^*, \delta_{M,2}^*, \delta_{M,3}^*)} = \frac{2{,}398}{1{,}9389} = 1{,}23678.$$

Przyjmując poziom istotności $\alpha = 0{,}1$ odczytano z tablicy XXII [1] wartość kwantyla

$$\eta_{K,2(n-1),1-\alpha} = \eta_{3,34,0,9} \approx 2.$$

Zatem nierówność (16) nie jest spełniona, przeto brak podstaw do odrzucenia weryfikowanej hipotezy.

Sprawdzenie się tej hipotezy nasuwa wniosek, że w czasie badań próbek o numerach $l = 1, 2, 3$ istniał jednakowy mechanizm uszkodzeń [7]. Ponieważ hipoteza, że $\delta_1 = \delta_2 = \delta_3$ nie została zakwestionowa, zweryfikowano hipotezę dotyczącą równości parametrów skali $b_1 = b_2 = b_3$ pod warunkiem, że parametr δ jest we wszystkich zbiorach wyników badań próbek $l = 1, 2, 3$ taki sam. Skorzystano z kryterium Firkowicza opartego o statystykę Hartleya [1], przy czym weryfikowaną hipotezę $[(b_1 = b_2 = b_3)|(\delta_1 = \delta_2 = \delta_3)]$ na danym poziomie istotności α należy odrzucić, jeżeli jest spełniona nierówność

$$\eta \geqslant \eta_{K-1,2,1-\alpha}, \tag{17}$$

gdzie:

$$\eta = \frac{\max(u_1, u_2)}{\min(u_1, u_2)} \tag{18}$$

oraz

$$u_j = j \ln \frac{b_{(j+1)}^*}{b_j^*}, \qquad j = 1, 2. \tag{19}$$

Jednocześnie wartości u_j zostały obliczone dla ciągu rozdzielczego

$$b_1^* \leqslant b_2^* \leqslant b_3^*. \tag{20}$$

Konstruując ciąg rozdzielczy (20) z wartości $b_{M,l}^*$ wyznaczonych ze wzoru (9), obliczono wartości u_j

$$u_j = 0{,}008 \quad \text{i} \quad u_2 = 0{,}012 \quad \text{i} \quad \eta = 1{,}5.$$

Przyjmując poziom istotności $\alpha = 0{,}1$ z tabl. XXII [1] odczytano

$$\eta_{K-1,2,1-\alpha} = \eta_{2,2,0,9} = 19.$$

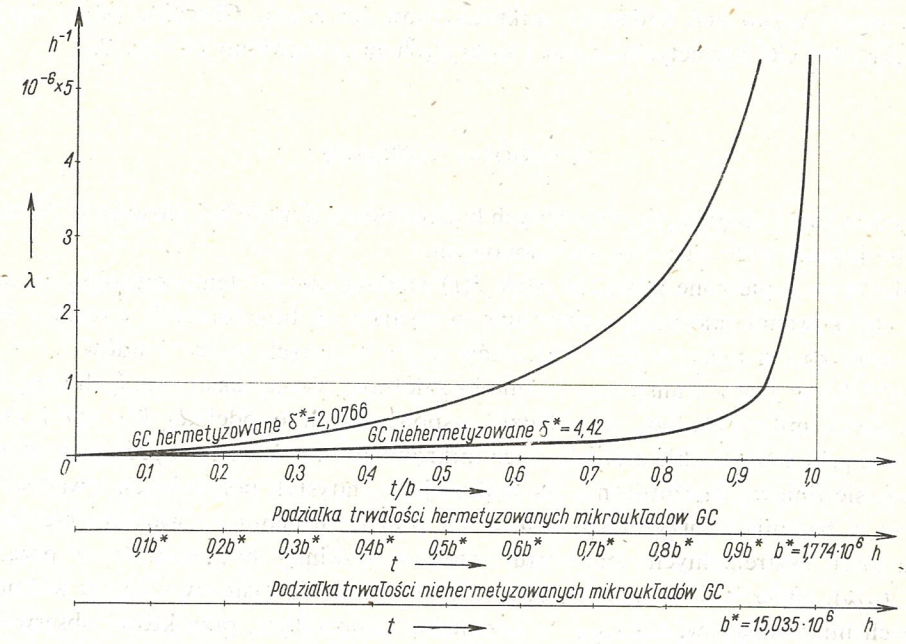
Nierówność (17) nie jest spełniona, zatem brak jest podstaw do odrzucenia wysuniętej hipotezy. Można więc wyznaczyć nieznane parametry δ i b rozkładu trwałości układów hermetyzowanych z próbki łącznej stosując wzory:

$$\frac{1}{\delta^*} = \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\delta^*_{M,l}}, \tag{21}$$

$$b^* = \max\{b^*_{M,l}\} = b^*_{M,3}, \tag{22}$$
$$1 \leqslant l \leqslant K.$$

Stąd dla hermetyzowanych mikroukładów uzyskano wartości oszacowań: $\delta^* = 2{,}0766$, $b^* = 1{,}7739 \cdot 10^6$ godz., zaś dla niehermetyzowanych: $\delta^* = 4{,}42$, $b^* = 15{,}0348 \cdot 10^6$ godz.

Chcąc ustalić zbiór wartości, które mogą przyjmować te parametry należy wyznaczyć ich dwustronne przedziały ufności na danym poziomie ufności, jednakże obliczenie tych wartości nie wnosi niczego nowego do meritum zagadnienia. Zatem informacje na ten temat zachowano na własny użytek.



Rys. 7. Charakterystyki niezawodnościowe $\lambda(t)$ mikroukładów *GC*

5) Znajomość modelu matematycznego rozkładu trwałości badanych mikroukładów pozwala prognozować zależność teoretyczną intensywności uszkodzeń od czacu — $\lambda(t)$. Przymiotnik „teoretyczna" jest po części pojęciem umownym, gdyż podstawowe wielkości rozkładu, takie jak parametr skali i parametr kształtu, zostały wyznaczone jako b i δ z danych empirycznych.

Przyjęty w pracy rozkład potęgowy określony jest gęstością

$$f(t) = \begin{cases} \frac{\delta}{b}\left(\frac{t}{b}\right)^{(\delta-1)} & \text{dla} \quad 0 < t < b, \\ 0 & \text{dla} \quad t \leqslant 0, \quad t \geqslant b. \end{cases} \tag{23}$$

Charakterystyki niezawodnościowe $R(t)$ i $\lambda(t)$ określone są wzorami:

$$R(t) = \begin{cases} 1-\left(\frac{t}{b}\right)^{\delta} & \text{dla} \quad 0 < t < b, \\ 0 & \text{dla} \quad t \geqslant b, \end{cases} \tag{24}$$

$$\lambda(t) = \frac{\delta\left(\frac{t}{b}\right)^{(\delta-1)}}{b\left[1-\left(\frac{t}{b}\right)^{\delta}\right]} \quad \text{dla} \quad 0 < t < b. \tag{25}$$

Funkcja $\lambda(t)$ posiada minimum $\lambda_M(t)$ w punkcie $t = t_M = b(1-\delta)^{1/\delta}$, gdy $\delta < 1$; przy $\delta \geqslant 1$ $t_M = 0$. Znając parametry δ^* i b^* można zatem wyznaczyć średnią zależność intensywności uszkodzeń badanych mikroukładów od czasu. Charakterystyki $\lambda(t)$ dla mikroukładów *GC* hermetyzowanych i odkrytych przedstawiono na rys. 7.

10. WNIOSKI KOŃCOWE

Przedstawiona metoda przyspieszonych badań trwałości może być stosowana do mikroukładów jedynie przy zasilaniu znamionowym.

Badania przyspieszone charakterystyk $\lambda(t)$ metodą szoków temperaturowych, oparte na zmodyfikowanej zależności Arrheniusa, pozwalają na transformację wyników badań na wartości zbioru rzeczywistych warunków pracy badanych mikroukładów.

Modyfikacja daje w transformacie liniową zależność różnicy ekstremalnych temperatur od trwałości mikroukładów, w układzie współrzędnych o odciętej $\xi = \ln t$ i rzędnej $\eta = 1/\Delta t$, przy stałej liczbie uszkodzonych mikroukładów w badanych próbkach. Ponadto uzyskuje się liniową zależność między realizacjami statystyk pozycyjnych trwałości uzyskanymi w badaniach przyspieszonych, a ich transformacjami na warunki rzeczywiste

Wartości ekstremalnych temperatur szoków powinny być zawarte w przedziale (t_{MAX}, t_{MIN}), gdzie: t_{MAX} — temperatura topliwości lutowia zastosowanego w technologii badanych mikroukładów, zaś t_{MIN} — minimalna temperatura, przy której obserwuje się eszcze poprawną pracę elementów półprzewodnikowych mikroukładów.

Celem przeprowadzenia badania w najkrótszym czasie, należy stosować ekstremalne temperatury szoków, niezbyt odbiegające od granic dozwolonego przedziału.

W celu korzystania z transformaty należy przebadać co najmnej dwie próbki przy różnych ekstremalnych temperaturach szoku i stałej temperaturze średniej, równej temperaturze złącza w warunkach znamionowych.

Hipotezę o jednorodności badanych próbek można zweryfikować stosując test Kruskalla-Wallisa.

Można przypuszczać, że mikroukłady hybrydowe grubowarstwowe posiadają potęgowy rozkład trwałości.

Metoda analizy statystycznej przedstawiona w punkcie 9 wymaga jedynie znajomości wartości transformowanej $t^*_{(M)}$. Metoda daje wystarczająco dokładne wyniki, jeśli ograniczyć się do niepełnej informacji z badanej próbki.

Zachowanie tego samego mechanizmu uszkodzeń w trakcie badań różnych próbek można sprawdzić weryfikując hipotezę o jednakowym rozkładzie trwałości (o tych samych parametrach) we wszystkich próbkach.

Dla typu *GC* stwierdzono konieczność dekompozycji konstrukcji mikroukładu i zmiany zastosowanych w serii *GC* tranzystorów na inne, wykazujące mniejszą zależność współczynnika wzmocnienia od temperatury otoczenia. Winno to polepszyć niezawodność układów *GC*.

Metoda pozwala na wyciągnięcie wniosków dotyczących własności niezawodnościowych badanych mikroukładów na etapie prototypu, serii próbnej lub produkcji. Ponadto umożliwia dokonanie obserwacji diagnostycznych uszkodzonych mikroukładów i porównanie wyników badań układów hermetyzowanych i odkrytych.

W przypadku zbadanych układów *GC* stwierdzono:

— Zależność intensywności uszkodzeń od czasu nie wykazuje istnienia okresu adaptacji, zatem układy *GC* nie wymagają selekcji (starzenia adaptacyjnego) dla poprawy niezawodności.

— Intensywność uszkodzeń jest funkcją monotonicznie rosnącą w czasie, zatem układy *GC* pracujące jako podzespoły maszyn cyfrowych (systemu) podlegają wymianom profilaktycznym.

— Intensywność uszkodzeń układów *GC* nie przekracza wartości $\lambda_0 = 10^{-6}$ 1/h w przedziale czasu eksploatacji (0 h, 10^6h).

— Hermetyzowane układy *GC* posiadają gorszą charakterystykę $\lambda(t)$ niż te układy przed hermetyzacją. Wynika stąd, że zastosowany do hermetyzacji epidian ma charakter destruktywny. Chcąc uzyskać lepszą niezawodność, należy hermetyzować układy *GC* zalewą termoutwardzalną posiadającą współczynnik rozszerzalności objętościowej bardziej zbliżony do rozszerzalności reszty bloku układu.

Reasumując można stwierdzić, że przedstawiona metoda badań przyspieszonych jest metodą prostą, ekonomiczną i szybką.

LITERATURA

1. S. Firkowicz, *Statystyczne badania wyrobów*, WNT, Warszawa 1970.
2. J. Kołodziejski, W. Rosiński, *Trwałość i stabilność przyrządów półprzewodnikowych*, Warszawa 1964.
3. E. F. Platz, *Reliability of Hybrid Microelectronics*, IEC. Pachagiging Symp., August 21÷22, 1968, San Francisco.
4. K. Durak, Z. Wąsowicz, *Wyniki badań wpływu czasu składowania mikroukładów hybrydowych na ich charakterystyki niezawodnościowe*, Zeszyty Naukowe AGH, 17/1977.
5. Ausgewählte Beiträge, Von einem Autorenkollektiv, *Zuverlässigkeit elektronischer Bauelemente*, VEB Deutscher Verlag für Grundstoffindustrie, Leipzig 1974.

6. Praca zbiorowa, *Niezawodność podzespołów i urządzeń elektronicznych*, Monografia ITR, Warszawa 1965.
7. Z. Wąsowicz, P. Włoka, *Przyspieszona estymacja ilościowych ocen niezawodności mikroelektronicznych układów hybrydowych przy zastosowaniu udarów termicznych*, Zeszyty Naukowe AGH, 1979, Automatyka, z. 20.

Z. WĄSOWICZ, P. WŁOKA

APPLICATION OF THERMAL SHOCKS FOR ACCELERATED ESTIMATION OF RELIABILITY CHARACTERISTICS OF THICK FILM MICROELECTRONIC HYBRID CIRCUITS

Summary

The program of the quick durability test of powered hybrid IC'S using thermal shocks is described. In the method the hybrid IC is treated as the basic element and its damage is defined as the departure from the performance area of its electrical parameters. The transformation based on the modified Arrhenius' equation enables one to estimate the realisations of the positional statistics of durability in real conditions. An appropriate method of statistical analysis for computer analysis is proposed.

Z. WĄSOWICZ, P. WŁOKA

UTILISATION DU CHOC THERMIQUE POUR ACCÉLÉRER L'ESTIMATION DES CARACTÉRISTIQUES DE FIABILITÉ DES CIRCUITS MICRO-ÉLECTRONIQUES HYBRIDES À COUCHE ÉPAISSE

Résumé

Dans l'article on a montré la méthode d'accélération de l'estimation numérique de la fiabilité des circuits hybrides par l'emploi d'un choc thermique. Dans la méthode, un circuit hybride est traité comme un élément et on suppose que ce circuit est endommagé quand ses paramètres électriques ne satisfont pas à la spécification technique. Ensuite on utilise l'équation modifiée d'Arrhenius, qui permet d'estimer la réalisation de la position statistique de la fiabilité dans une condition réelle. On présente aussi la méthode de l'analyse statistique à l'aide d'un ordinateur.

Z. WĄSOWICZ, P. WŁOKA

ANWENDUNG VON WÄRMESTÖßEN ZUR BESCHLEUNIGTEN BEWERTUNG DER ZUVERLÄSSIGKEITSCHARAKTERISTIKEN MIKROELEKTRONISCHER DICKSCHICHT-HYBRIDENSCHALTUNGEN

Zusammenfassung

Es wurde die Realisierung des Programms des beschleuningten Lebensdauerprüfung der gespeisten mikroelektronischen digitalen Schaltkreise unter Anwendung von Wärmestößen präsentiert. Dabei wurde die mikroelektronische Hybridschaltung als ein Bauelement behandelt, und die Beschädigung als Ausfall der Werte elektrischer Parameter der Schaltung aus den Zuständen der Tauglichkeit definiert. Die auf der umformten Arrhenius-Gleichung begründete Transformation gestattet, die Realisierung der Positionsstatistiken der Lebensdauer in reellen Bedingungen abzuschätzen. Gleichzeitig wurde auch eine günstige rechnergestützte Methode für eine geeignete statistische Analyse vorgeschlagen.

З. ВОНСОВИЧ, П. ВЛОКА

ПРИМЕНЕНИЕ ТЕРМИЧЕСКИХ УДАРОВ ДЛЯ УСКОРЕННОГО ОПРЕДЕЛЕНИЯ ХАРАКТЕРИСТИК НАДЕЖНОСТИ ТОЛСТОПЛЕНОЧНЫХ ГИБРИДНЫХ МИКРОСХЕМ

Р е з ю м е

Считая гибридную микросхему в целом элементом и определяя выход его из строя выходом электрических параметров за пределы множества состояний годности, показана реализация программы ускоренных испытаний до первого отказа с применением термоударов. Вводится преобразование, основанное на модифицированном уравнении Арреньюса, которое позваляет на оценку позиционных статистик базотказности в реальных условиях. Предлагается тоже соответствующий метод статистического анализа, существенно облегчающего расчёты на ЭВМ.

ROZPRAWY ELEKTROTECHNICZNE 1982, 28, z. 1-2, ss. 219-231

621.319.43:621.315.592

Własności nierównowagowe kondensatora typu MIS w warunkach oświetlenia

STANISŁAW KRAWCZYK (WARSZAWA)

Instytut Technologii Elektronowej Politechniki Warszawskiej

Otrzymano 22.4.1980

W pracy przedstawiono jakościową i ilościową analizę procesów nierównowagowych w kondensatorze MIS w warunkach oświetlenia.

1. WPROWADZENIE

Do chwili obecnej powstało wiele prac teoretycznych i eksperymentalnych poświęconych procesom nierównowagowym w nieoświetlonym kondensatorze typu MIS. Brak jest natomiast opisu procesów nierównowagowych zachodzących w warunkach stałego oświetlenia. Jedynie w pracach [11, 14] przedstawiono próby uwzględnienia wpływu oświetlenia na przebieg procesów nierównowagowych, ograniczając się jednak raczej do dyskusji niż analizy ilościowej tego zagadnienia. Również pokazane w pracach [5, 7] analizy numeryczne polegające na rozwiązaniu równań transportu mimo swojej ogólności nie zostały zaadaptowane do tego celu.

Tymczasem wydaje się, że opis procesów nierównowagowych zachodzących w kondensatorze MIS w warunkach oświetlenia jest zagadnieniem ważnym, na przykład w związku z realnymi perspektywami wykorzystania przyrządów typu CCD jako pól obrazowych w kamerach telewizyjnych i filmowych [1, 2, 3]. Przypuszcza się także, że znajdą zastosowanie kontrolowane za pomocą oświetlenia struktury MIS pracujące w warunkach impulsowych [11]. Procesy nierównowagowe przebiegające w strukturze MIS w warunkach oświetlenia mogą stwarzać także możliwości opracowania nowych metod badania układu metal-dielektryk-półprzewodnik.

Duża różnorodność zjawisk przebiegających w obszarze przypowierzchniowym półprzewodnika i duża ilość sposobów w jaki można wywołać stan nieustalony sprawia, że trudne jest stworzenie prostego opisu analitycznego, jednolitego dla wszystkich procesów nierównowagowych w kondensatorze MIS. W tej pracy omówiono procesy nierównowagowe towarzyszące sterowaniu oświetlonego kondensatora MIS impulsem prostokątnym, który wytwarza głębokie zubożenia, oraz procesy towarzyszące przełączaniu w kierunku stanu płaskich pasm bądź akumulacji.

Przedstawiona tu analiza bazuje na modelu matematyczno-fizycznym i metodzie opisu obszaru przypowierzchniowego półprzewodnika wprowadzonych w pracach [9, 10]. W pracach tych zdefiniowano efektywny poziom generacji świetlnej ξ

$$\xi = \frac{\Delta n}{n_i}, \tag{1}$$

gdzie:

Δn — koncentracja nośników nadmiarowych,

n_i — koncentracja samoistna.

Wprowadzono tam także, jako istotne parametry półprzewodnika, efektywny potencjał Fermiego u_F^* oraz efektywną koncentrację samoistną n_i^*:

$$u_F^* = \frac{1}{2} \ln \left[\frac{\xi + e^{u_F}}{\xi + e^{-u_F}} \right], \tag{2}$$

$$n_i^* = n_i [(\xi + e^{u_F})(\xi + e^{-u_F})]^{1/2}, \tag{3}$$

gdzie u_F — potencjał Fermiego $\left(\text{w jednostkach } \frac{kT}{q}\right)$ półprzewodnika nieoświetlonego.

Pokazano [9], że opis właściwości kondensatora MIS w warunkach stałego oświetlenia można na ogół przeprowadzić analogicznie jak dla stanu równowagi termodynamicznej, zastępując w dobrze znanych równaniach u_F i n_i przez u_F^* i n_i^*. Wielkości opisujące właściwości oświetlonego kondensatora MIS przyjęto tu oznaczać symbolami z gwiazdkami.

2. PRZEŁĄCZENIE OŚWIETLENIONEGO KONDENSATORA MIS Z WYTWORZENIEM GŁĘBOKIEGO ZUBOŻENIA

W rozważanym tu przypadku kondensator MIS znajdujący się początkowo w stanie akumulacji, płaskich pasm, bądź zubożenia zostaje spolaryzowany impulsem prostokątnym w kierunku inwersji. Prawie natychmiast po przyłożeniu impulsu przy powierzchni półprzewodnika powstaje obszar głębokiego zubożenia. Stała czasowa tego procesu jest równa czasowi relaksacji nośników większościowych, a w praktyce jest ona ograniczona szybkością narastania przyłożonego na bramkę sygnału. Można przyjąć, że w trakcie formowania obszaru głębokiego zubożenia obsadzanie szybkich stanów powierzchniowych nie ulega zmianie. Czas trwania tej pierwszej fazy przełączania nie zależy od oświetlenia półprzewodnika. Na początku drugiej fazy przełączania koncentracja nośników większościowych przy powierzchni jest bardzo mała, natomiast powierzchniowa koncentracja nośników mniejszościowych zaczyna wzrastać ponad swoją wartość z okresu przed przełączaniem. Czas trwania drugiej fazy procesu przełączania trwa aż do momentu, gdy koncentracja nośników mniejszościowych przy powierzchni dojdzie do wartości bliskiej koncentracji nośników większościowych w objętości. Potencjał powierzchniowy ustalił się na poziomie odpowiadającym końcowi pierwszej fazy procesu przełączania (stanowi głębokiego zubożenia) i pozostanie na tej wartości aż do końca drugiej fazy przełączania, ponieważ ładunek nośników mniejszościowych przy powierzchni jest przez cały ten czas znacznie mniejszy od ładunku zjonizowania domieszek. Dostarczanie nośników mniej-

szościowych do powierzchni następuje na skutek dyfuzji z obszaru quasi-neutralnego półprzewodnika, relaksacji centrów pułapkowych w objętości i na powierzchni, generacji termicznej i świetlnej w obszarze ładunku przestrzennego a także generacji powierzchniowej. Przy końcu drugiej fazy procesu przełączania koncentracja nośników mniejszościowych przy powierzchni półprzewodnika zaczyna przewyższać koncentracje nośników większościowych w objętości półprzewodnika.

W trakcie trzeciej fazy przełączania następuje zmniejszenie szerokości obszaru ładunku przestrzennego i wartości potencjału powierzchniowego do ich wartości quasi-równowagowych. Następuje także wzrost koncentracji nośników mniejszościowych przy powierzchni półprzewodnika, tak aby w warunkach quasi-równowagi spełniony był warunek

$$\bigwedge_x n^*(x)p^*(x) = n_i^{*2}. \tag{4}$$

W przypadku typowego, nieoświetlonego kondensatora MIS czas trwania pierwszej fazy wynosi około $10^{-11} \div 10^{-9}$ s, drugiej fazy około $10^{-5} \div 10^{-2}$ s, czas trwania trzeciej fazy może zaś być rzędu kilku lub nawet kilkudziesięciu sekund. Czas trwania drugiej i trzeciej fazy przełączania zależą, jak to pokazano w dalszej części tego artykułu, od oświetlonego półprzewodnika.

2.1. Zjawiska zachodzące w oświetlonej strukturze MIS w warunkach głębokiego zubożenia

Procesy relaksacji cetrów pułapkowych na powierzchni i w objętości półprzewodnika

Bezpośrednio po ukształtowaniu się głębokiego zubożenia po przyłożeniu impulsu napięcia na bramkę następuje proces relaksacji centrów pułapkowych na powierzchni i w obszarze przypowierzchniowym półprzewodnika. Trwa on do chwili osiągnięcia przez nie stanu quasi-równowagi z pasmem nośników mniejszościowych. W przypadku kondensatora MIS z półprzewodnikiem typu *p* przełączonego ze stanu płaskich pasm w kierunku silnej inwersji duża ilość stanów powierzchniowych, początkowo nieobsadzonych, zostaje w trakcie procesu przejściowego zapełniona przez elektrony.

Dla oszacowania wartości czasu relaksacji szybkich stanów powierzchniowych przyjęto, że przyrost koncentracji elektronów w stanach powierzchniowych $\frac{dn_T^*}{dt}$ lub dziur $\frac{dp_T^*}{dt}$ zachodzi przede wszystkim w wyniku emisji dziur do pasma walencyjnego i przechwytu elektronów z pasma przewodnictwa. Tak więc mamy:

$$\frac{dn_T^*}{dt} = -\frac{dp_T^*}{dt} \approx P_T \sigma V_{th}(n_t e^{n_t} + n_s^*), \tag{5}$$

gdzie:

V_{th} — prędkość cieplna nośników,

n_s^* — koncentracja elektronów na powierzchni,

u_t — potencjał poziomu pułapkowego;

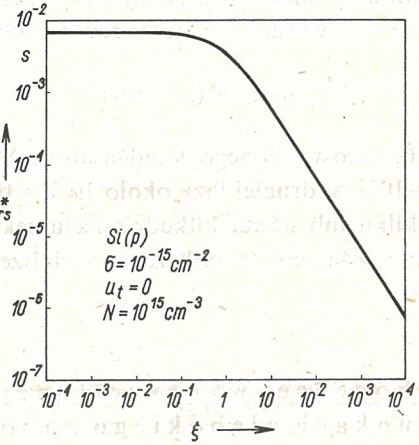
(założono tu dla uproszczenia, że przekroje czynne pułapki na przechwyt elektronów i dziur są sobie równe: $\sigma_p = \sigma_n = \sigma$).

Rozwiązanie równania (5) można zapisać w następującej postaci

$$p_T^*(u_t, t) = P_T^*(u_t, 0)e^{-t/\tau_{rs}^*}, \tag{6}$$

gdzie $\tau_{rs}^* = [\sigma V_{th}(n_i e^{u_t} + n_s^*)]^{-1}$. (7)

Zależność tak oszacowanego czasu relaksacji szybkich stanów powierzchniowych od efektywnego poziomu generacji świetlnej pokazano na rys. 1. Maleje on silnie przy wzroście



Rys. 1. Zależność czasu relaksacji szybkich stanów powierzchniowych (w trakcie przełączania oświetleniowego kondensatora MIS) od efektywnego poziomu generacji świetlnej ξ.

natężenia oświetlenia półprzewodnika. Podobną analizę, można przeprowadzić dla centrów pułapkowych zlokalizowanych w pobliżu powierzchni półprzewodnika.

Dokładne obliczenia numeryczne i eksperyment [5, 13, 14] wskazują, że procesy relaksacji szybkich stanów powierzchniowych i pułapek w półprzewodniku w nieznacznym tylko stopniu wpływają na przebieg dalszych etapów procesu przełączania. Prace te traktowały oczywiście o nieoświetlonym kondensatorze MIS, ponieważ jednak w przypadku oświetlonego kondensatora MIS całkowita zmiana ładunku w stanach powierzchniowych w trakcie podobnego procesu przełączenia nie może być większa, a jednocześnie czas relaksacji tych stanów jest krótszy, wydaje się, że wnioski z prac [5, 13, 14] można zaadoptować również do przypadku oświetlonego kondensatora MIS.

Procesy generacji w oświetlonej strukturze MIS w warunkach głębokiego zubożenia

W drugiej fazie procesu przełączania następuje gromadzenie nośników mniejszościowych przy powierzchni półprzewodnika. Bilans równowagi ładunku nośników mniejszościowych dopływających do powierzchni półprzewodnika w trakcie omawianego tu procesu może być zapisany następująco

$$\frac{dQ_{scn}^*}{dt} = -I_g^* = -I_B^* - I_d^* - I_s^*, \tag{8}$$

gdzie:

I_B^* — gęstość prądu nośników dochodzących z obszaru quasi-neutralnego półprzewodnika,

I_d^* — gęstość prądu nośników generowanych w obszarze ładunku przestrzennego,

I_s^* — gęstość prądu nośników generowanych z udziałem szybkich stanów powierzchniowych.

Wprowadzenie pojęcia generacji znacznie ułatwia analityczny opis procesu przełączania. Wstawiając odpowiednie wyrażenia na prądy generacji trzeba jednak mieć na uwadze, że w omawianych tu warunkach gwałtownie zmienia się koncentracja nośników zapełniających stany pułapkowe.

Omawianie poszczególnych składników zależności (8) zacznijmy od prądu I_B^*. W warunkach głębokiego zubożenia gęstość prądu nośników dochodzących z obszaru quasi-neutralnego może być uważana za stałą. Rozwiązanie równania ciągłości w obszarze quasi-neutralnym półprzewodnika (przy założeniu, że koncentracje nośników w obszarze zubożonym są równe zeru) prowadzi do następującego wyrażenia na gęstość prądu nośników mniejszościowych dyfundujących do krawędzi obszaru ładunku przestrzennego

$$I_B^* = q \frac{n_o^* D}{L}, \tag{9}$$

gdzie:

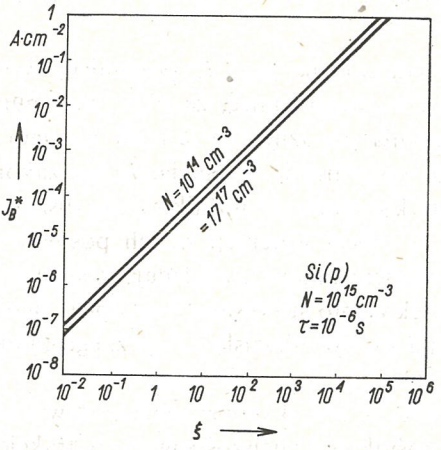
L — długość drogi dyfuzji nośników mniejszościowych

$$L = \sqrt{D\tau}, \tag{10}$$

D — współczynnik dyfuzji nośników mniejszościowych,

τ — czas życia nośników mniejszościowych,

n_o^* — quasi-równowagowa koncentracja nośników mniejszościowych w objętości półprzewodnika.



Rys. 2. Zależność prądu nośników generowanych w obszarze quasi-neutralnym półprzewodnika od efektywnego poziomu generacji świetlnej ξ.

Parametrem tych przebiegów jest koncentracja domieszek w półprzewodniku

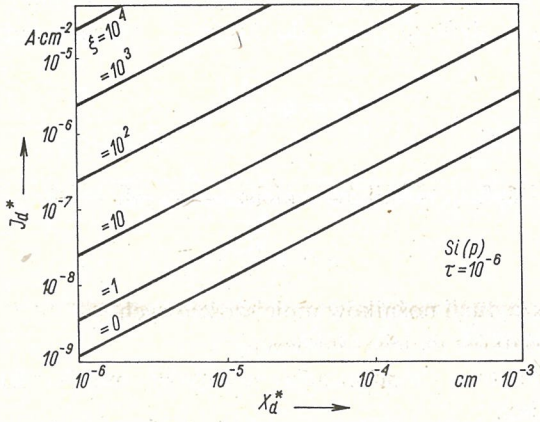
Prąd nośników generowanych w obszarze quasi-neutralnym nie zależy w tej fazie procesu przełączania od szerokości obszaru ładunku przestrzennego (a więc od polaryzacji bramki), a ponadto wzrasta przy wzroście efektywnego poziomu generacji świetlnej ξ. Zależność prądu I_B^* od efektywnego poziomu generacji świetlnej wykreśloną według zależności (9) pokazano na rys. 2.

Przy opracowaniu prostego modelu drugiej fazy procesu przełączania założono także stałą wartość prędkości generacji w obszarze głębokiego zubożenia (szerokość obszaru głębokiego zubożenia pozostaje stała w trakcie tej fazy procesu) [9, 10]:

$$I_d^* = \frac{n_i}{\tau}\left(\frac{1}{2}+\xi\right) x_d^* q, \tag{11}$$

gdzie x_d^* — szerokość obszaru ładunku przestrzennego w warunkach oświetlenia.

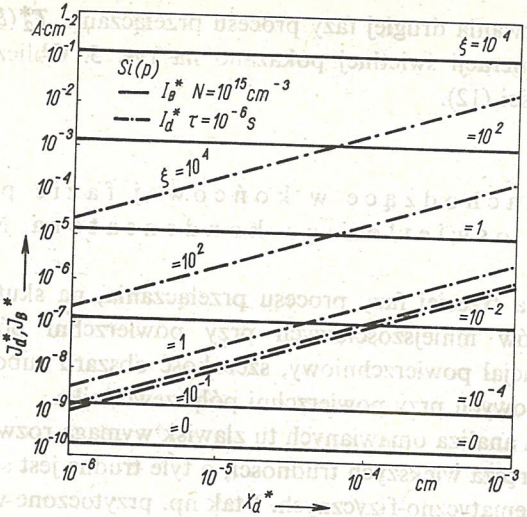
Zależność prądu I_d^* od szerokości obszaru zubożonego x_d^* przedstawiono na rys. 3.



Rys. 3. Zależność prądu generacji I_d^* od szerokości obszaru głębokiego zubożenia x_d^*. Parametrem tych przebiegów jest efektywny poziom generacji świetlnej ξ

Koncentrując się na wyrażeniu na prąd I_s^* powstający w wyniku generacji-rekombinacji powierzchniowej trzeba przypomnieć, że obliczenia numeryczne przeprowadzone dla warunków bez oświetlenia wskazują, że jest on na ogół znacznie mniejszy niż I_B^* i I_d^* [5, 7]. Przyczynę tego upatruje się w fakcie, że po zakończeniu procesu relaksacji stanów powierzchniowych koncentracje nośników mniejszościowych przy powierzchni są już stosunkowo duże, a proces generacji za ich pośrednictwem jest mało wydajny. Proste modele generacji powierzchniowej, zakładające stałą wartość prędkości generacji powierzchniowej (aczkolwiek często stosowane przy opisie przełączanego nieoświetlonego kondensatora MIS), wydają się być nieścisłe z fizycznego punktu widzenia i nie będą tu rozwijane.

Zależności przedstawione na rys. 4 umożliwiają porównanie udziału prądów I_B^* i I_d^* w zjawisku gromadzenia nośników mniejszościowych w trakcie trwania drugiej fazy procesu przełączania. Przyjmując, że czas drugiej fazy procesu przełączania trwa do chwili, gdy ładunek nośników mniejszościowych przy powierzchni półprzewodnika przewyższy



Rys. 4. Porównanie wielkości prądów I_B^* i I_d^* w warunkach głębokiego zubożenia

ładunek nośników większościowych w objętości, może on zostać oszacowany z następującego równania

$$T_2^* = \frac{Q_{scp}^*(0^+)}{I_d^* + I_B^*}, \tag{12}$$

gdzie $Q_{scp}^*(0^+)$ jest ładunkiem obszaru głębokiego zubożenia bezpośrednio po przełączeniu

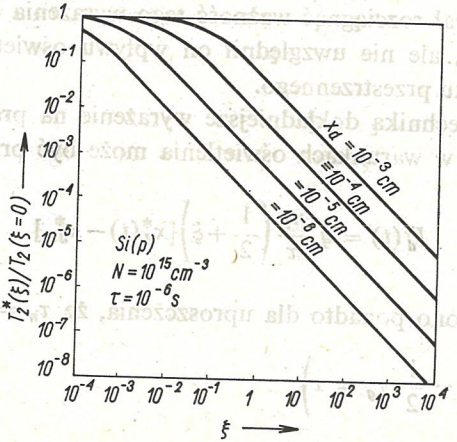
$$Q_{scp}^*(0^+) = \left[\frac{2\varepsilon_s|\varphi_s^*(0^+)|}{qN}\right]^{1/2} Nq, \tag{13}$$

przy czym:

$\varphi_s(0^+)$ — potencjał powierzchniowy bezpośrednio po przełączeniu,

N — koncentracja domieszek w półprzewodniku,

ε_s — przenikalność elektryczna półprzewodnika.



Rys. 5. Zależność czasu trwania drugiej fazy procesu przełączania od efektywnego poziomu generacji świetlnej ξ

Parametrem jest szerokość obszaru głębokiego zubożenia x_d^*

Zależność czasu trwania drugiej fazy procesu przełączania $T_2^*(\xi)/T_2$ ($\xi = 0$) od efektywnego poziomu generacji świetlnej pokazano na rys. 5. Obliczenia przeprowadzono na podstawie zależności (12).

2.2. Zjawiska zachodzące w końcowej fazie procesu przełączania oświetlonego kondensatora MIS

W trakcie trwania trzeciej fazy procesu przełączania, na skutek znacznego wzrostu koncentracji nośników mniejszościowych przy powierzchni półprzewodnika, ulegają zmianie m.in.: potencjał powierzchniowy, szerokość obszaru zubożonego i koncentracja nośników większościowych przy powierzchni półprzewodnika.

O ile numeryczna analiza omawianych tu zjawisk wymaga rozwiązania układu równań transportu i nie nastręcza większych trudności, o tyle trudne jest stworzenie dla ich opisu prostych modeli matematyczno-fizycznych. I tak np. przytoczone w poprzednim rozdziale wyrażenia na prąd I_d^* i I_B^* przestają być słuszne m.in. ze względu na to, że przestaje być prawdziwa aproksymacja głębokiego zubożenia.

Gęstość prądu nośników generowanych w obszarze głębokiego zubożenia w nieoświetlonej strukturze MIS jest w większości prac [8, 16] zapisywana w następującej postaci[1)]

$$I_d(t) = \frac{n_i}{\tau_g}\,[x_d(t) - x_{d\infty}], \tag{14}$$

gdzie:

x_d — quasi-równowagowa szerokość obszaru ładunku przestrzennego,

τ_g — stała czasowa generacji nośników w obszarze całkowicie zubożonym

$$\tau_g = \tau_p e^{-u_t} + \tau_n e^{u_t}, \tag{15}$$

τ_p, τ_n — czasy życia elektronów i dziur.

Schroder [15] próbował rozciągnąć ważność tego wyrażenia dla przypadku oświetlonego kondensatora MIS, ale nie uwzględnił on wpływu oświetlenia na równowagową szerokość obszaru ładunku przestrzennego.

Postępując tą samą techniką dokładniejsze wyrażenie na prąd generacji w obszarze ładunku przestrzennego, w warunkach oświetlenia może być przedstawione następująco

$$I_d^*(t) = q\frac{n_i}{\tau}\left(\frac{1}{2} + \xi\right)[x_d^*(t) - x_{d\infty}^*] \tag{16}$$

(w wyrażeniu tym założono ponadto dla uproszczenia, że $\tau_n = \tau_p$ oraz $u_t = 0$, co daje w konsekwencji $\tau_n = \tau_p = \frac{1}{2}\tau_g = \tau$).

[1)] Mimo, że wyrażenie to nie ma właściwej interpretacji fizycznej daje ono dobrą zgodność teorii z eksperymentem w przypadku przełączania nieoświetlonego kondensatora MIS.

Biorąc pod uwagę analogię ze złączem p-n zależność prądu dyfuzji z obszaru quasi-neutralnego w końcowej fazie procesu przełączania przedstawiono tu w następującej postaci[2)]

$$I^*_{B(t)} = q\frac{n^*_o D_n}{L_n}\{1-\exp[u^*_{s\infty}-u^*_s(t)]\}, \tag{17}$$

gdzie $u_{s\infty}$ oznacza quasi-równowagową wartość potencjału powierzchniowego.

Oszacowanie czasu trwania trzeciej fazy procesu przełączania przeprowadzono w rozdziale następnym na podstawie analizy zależności $I^*_{MIS}(t)$ i $C^*_{MIS}(t)$.

2.3. Zależność $I^*_{MIS}(t)$ i $C^*_{MIS}(t)$ po przełączeniu oświetlonego kondensatora MIS w kierunku silnej inwersji

Zależności $I^*_{MIS}(t)$ i $C(t)^*_{MIS}$ wyprowadzono tu podobnie jak dla przypadku procesów przełączania nieoświetlonego kondensatora MIS [4, 5, 15, 16]. Równanie wyjściowe do obliczania tych zależności przedstawiono poniżej

$$U_G = \varphi^*_s - \frac{1}{C_o}[Q^*_{scn}+Q^*_{scp}+Q^*_{st}], \tag{18}$$

gdzie:

U_G — napięcie bramka-podłoże,
Q^*_{scn} — ładunek nośników mniejszościowych,
Q^*_{scp} — ładunek nośników większościowych,
Q^*_{st} — ładunek szybkich stanów powierzchniowych,
C_o — pojemność dielektryka na jednostkę powierzchni.

Zróżniczkowanie tego równania daje następujące wyrażenie

$$\frac{dU_G}{dt} = 0 = \frac{d\varphi^*_s}{dt} - \frac{1}{C_o}\left[\frac{dQ^*_{scn}}{dt}+\frac{dQ^*_{scp}}{dt}+\frac{dQ^*_{st}}{dt}\right]. \tag{19}$$

Abstrahując od mechanizmów determinujących prąd generacji I^*_g można zapisać następującą równość

$$I^*_g = -\left[\frac{dQ^*_{scn}}{dt}+\frac{dQ^*_{st}}{dt}\right]. \tag{20}$$

Korzystając ponadto z następujących zależności:

$$I^*_x \overset{\text{df}}{=} \frac{dQ^*_{sep}}{dt}, \tag{21}$$

$$\frac{d\varphi^*_s}{dt} = \frac{1}{C^*_{scp}} I^*_x, \tag{22}$$

$$C^*_{scp} = \frac{\varepsilon_s}{x^*_d} \overset{\text{df}}{=} C^*_d, \tag{23}$$

[2)] Położenie quasi-potencjału Fermiego nośników mniejszościowych na powierzchni półprzewodnika można w trakcie tego procesu z dobrym przybliżeniem uważać za stałe; ulega zmianie głównie wartość potencjału powierzchniowego.

można równanie (19) zapisać w następującej postaci

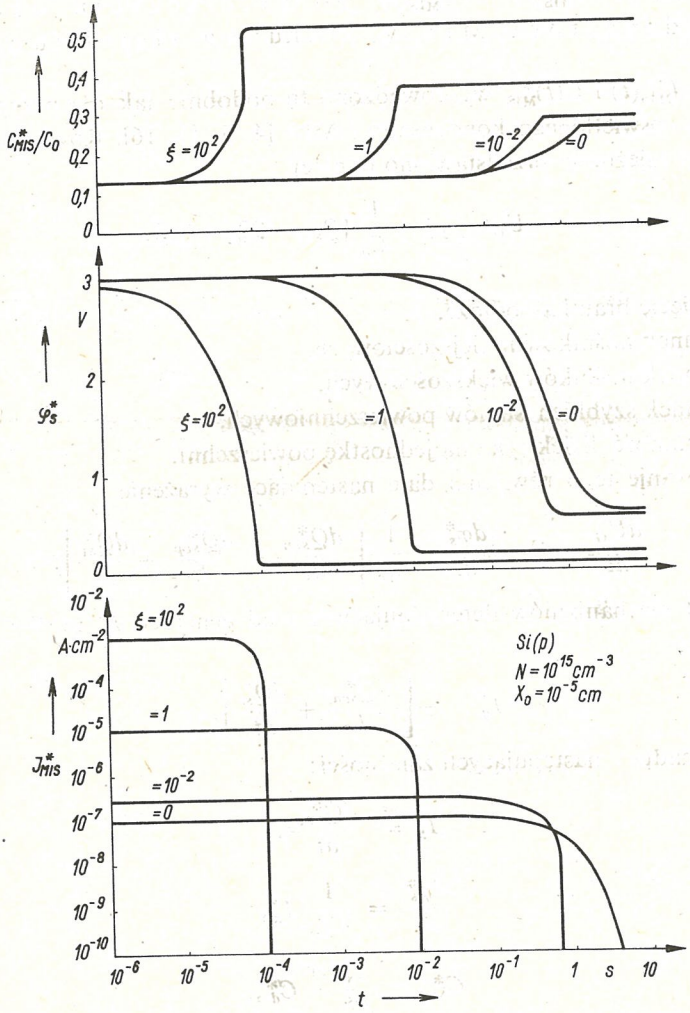
$$\frac{1}{C_d} I_x^* + \frac{1}{C_o} I_g^* + \frac{1}{C_o} I_x^* = 0. \tag{24}$$

Całkowity prąd I_{MIS}^* płynący w obwodzie bramka-podłoże w czasie procesu przełączania określony jest następująco

$$I_{MIS}^* = I_g^* + I_x^*, \tag{25}$$

lub też po elementarnych przekształceniach w postaci:

$$I_{MIS}^*(t) = \frac{C_o}{C_o + C_d^*(t)} I_g(t), \tag{26}$$



Rys. 6. Przykładowe zależności $I_{MIS}^*(t)$ i $C_{MIS}^*(t)$ oświetlonego kondensatora MIS przełączonego ze stanu płaskich pasm w kierunku silnej inwersji

Parametrem jest efektywny poziom generacji świetlnej ξ. Przyjęto następujące parametry struktury MIS: $N = 10^{15}$ cm^{-3}, $x_o = 10^{-5}$ cm, $\tau = 10^{-6}$ s, $U_G = 3{,}9$ V

$$I^*_{\mathrm{MIS}}(t) = 1 - \frac{C^{*HF}_{\mathrm{MIS}}(t)}{C_o} I^*_g(t), \tag{27}$$

gdzie C^{*HF}_{MIS} — wysokoczęstotliwościowa pojemność kondensatora MIS.

Prąd I^*_{MIS} nie jest równy prądowi generacji, ponieważ część generowanych nośników neutralizuje ładunek nośników większościowych z wewnętrznej krawędzi obszaru ładunku przestrzennego i nie przepływa w obwodzie bramka-podłoże.

Przykładowe przebiegi $I^*_{\mathrm{MIS}}(t)$ i $C^*_{\mathrm{MIS}}(t)$ przy przełączeniu oświetlonego kondensatora MIS ze stanu płaskich pasm w kierunku silnej inwersji obliczone na gruncie przedstawionej tu teorii pokazano na rys. 6. Przy przyjętej tu skali przebiegi pierwszej i drugiej fazy przełączania są niewidoczne i zostały pominięte przy wykonywaniu wykresu tych zależności.

Przeprowadzona analiza pokazuje m.in., że oświetlenie półprzewodnika powoduje wyraźne skrócenie trzeciej, najdłużej trwającej fazy procesu przełączania kondensatora MIS.

3. PRZEŁĄCZENIE OŚWIETLONEGO KONDENSATORA MIS W KIERUNKU STANU PŁASKICH PASM I STANU AKUMULACJI

Impulsowe wprowadzenie kondensatora MIS znajdującego się początkowo np. w stanie silnej inwersji do stanu płaskich pasm bądź akumulacji powoduje prawie natychmiastowe (określone procesem relaksacji dielektrycznej) zniknięcie obszaru zubożonego przy powierzchni półprzewodnika. Natychmiast po zniknięciu obszaru zubożonego nadmiarowe nośniki mniejszościowe znajdujące się w obszarze przypowierzchniowym półprzewodnika rozpoczynają dyfuzję ambipolarną w kierunku objętości półprzewodnika i jednocześnie rekombinują z nośnikami większościowymi. Opis matematyczny tego zjawiska otrzymać można poprzez rozwiązanie równań ciągłości, w obszarze przypowierzchniowym półprzewodnika.

Przy pominięciu obecności szybkich stanów powierzchniowych uproszczone rozwiązanie równania ciągłości w tych warunkach można zapisać w następującej postaci

$$Q^*_{scn}(t) = Q^*_{scn}(0)\exp\left(-\frac{t}{\tau}\right), \tag{28}$$

gdzie $Q^*_{scn}(0)$ oznacza ładunek nadmiarowych nośników mniejszościowych w chwili $t = 0$ po przełączeniu.

Przebieg procesu rekombinacji, przy przyjętych tu założeniach, nie zależy od efektywnego poziomu generacji świetlnej [10].

W najprostszym przybliżeniu można przyjąć, że równocześnie z procesami dyfuzji i rekombinacji w obszarze przypowierzchniowym półprzewodnika występuje po przełączeniu proces relaksacji szybkich stanów powierzchniowych. W procesie tym dominuje przechwyt nośników większościowych oraz w mniejszym stopniu emisja nośników mniejszościowych. Emisja nośników większościowych i przechwyt nośników mniejszościowych mogą być na ogół całkowicie pominięte.

Stała czasowa tego procesu, np. dla półprzewodnika typu p może być oszacowana jako $(\sigma_p V_{th} N_a)^{-1}$. Przy koncentracji domieszek $N_a = 10^{15}\mathrm{cm}^{-1}$ oraz $\sigma_p = z$ zakresu

od $10^{-14}cm^{-2}$ do $10^{15}cm^{-2}$ stała czasowa relaksacji szybkich stanów powierzchniowych mieści się w granicach od 10^{-8} do $10^{-6}s$.

Proces relaksacji szybkich stanów powierzchniowych towarzyszący przełączeniu kondensatora MIS w kierunku stanu płaskich pasm lub akumulacji przy przyjętych tu założeniach również nie zależy od oświetlenia. Przeprowadzona analiza nie wykazała więc wpływu oświetlenia na przebieg procesu przełączania kondensatora MIS w kierunku stanu płaskich pasm bądź akumulacji.

4. PODSUMOWANIE

W pracy przedstawiono jakościową i ilościową analizę procesów nierównowagowych w kondensatorze MIS w warunkach oświetlenia. Pokazano, że oświetlenie może znacznie skrócić czas powrotu do stanu quasi-równowagi kondensatora MIS sterowanego impulsem prostokątnym w kierunku głębokiego zubożenia. Oświetlenie nie wywiera natomiast istotnego wpływu na procesy nierównowagowe towarzyszące przełączaniu kondensatora MIS w kierunku stanu płaskich pasm bądź akumulacji.

LITERATURA

1. L. Altman, *CCD devices*, Electronics, 43, 112, 1970.
2. D. F. Barbe, *Imaging devices using the charge-conpled concept*, Proc. IEEE, 63, 38, 1975.
3. H. K. Burke, G. J. Michen, *Charge-injection imaging: operating techniques and performances characteristics*, IEEE Trans. Electron. Dev., ED-23, 189, 1976.
4. P. U. Calzolari, S. Graffi, C. Morandi, *Field-enhanced carrier generation in MOS capacitors*, Solid-State Electron., 17, 1001, 1974.
5. P. U. Calzolari, S. Graffi, G. Pierini, *Effects of surface states relaxation phenomena in MOS capacitor*, Solid-State Electron., 19, 325, 1976.
6. P. U. Calzolari, S. Graffi, C. Morandi, *Minority carrier recombination in MOS capacitors switched from inversion to accumulation*, Solid-State Electron., 20, 205, 1977.
7. T. H. Collins, J. N. Churchill, *Exact modeling of the transient response of MOS capacitor*, IEEE Trans. Electron. Devices, 22, 90, 1975.
8. F. P. Heiman, *On the determination of minority carrier lifetime from the transient response of an MOS capacitor*, IEEE Trans. Electron. Dev., ED-14, 781, 1967.
9. A. Jakubowski, S. Krawczyk, *Electrical properties of MIS capacitor under illumination*, Electron Technology, 11, 3, 1978.
10. A. Jakubowski, S. Krawczyk, *Simple model of generation and recombination processes in semiconductor under illumination*, Bull. Ac. Pol. Sc. Sér. Tech., 26, 185, 1978.
11. A. G. Nassiban, R. B. Galligaro, J. G. Simons, *Digital optical metal-insulator-silicon thyristor* (o.m.i.s.t.), IEE Solid-State Electron. Dev., 2, 149, 1978.
12. H. M. Preier, *Different mechanism affecting the inversion layer transient response*, IEEE Trans. Electron. Dev., ED-15, 990, 1970.
13. C. T. Sah, H. S. Fu, *Current and capacitance transient responses of MOS capacitor. General theory and applications to initially depleted surface without surface states*, Phys. Stat. Sol. A., 11, 297, 1972.
14. C. T. Sah, H. S. Fu, *Current and capacitance transient responses of MOS capacitors. Recombination centres in the surface space charge layer*, Phys. Stat. Sol., 14, 59, 1972.

15. D. K. Schroder, *Bulk and optical generation parameters measured with the pulsed MOS capacitor*, IEEE Trans. Electron. Dev., ED-19, 1018, 1972.
16. M. Zerbst, *Relaxationseffekte an Halbleiter—Isolator—Grenz—Flüchen*, Z. Angew. Phys., 22, 30, 1966.

S. KRAWCZYK

NON-EQUILIBRIUM PROPERTIES OF MIS CAPACITOR UNDER ILLUMINATION CONDITIONS

Summary

In this work a qualitative and quantitative analysis of non-equilibrium properties of MIS capacitor under illumination conditions is presented.

S. KRAWCZYK

PROPRIÉTÉS DU DÉSÉQUILIBRE DE LA CAPACITÉ DU CONDENSATEUR TYPE MIS DANS LES CONDITIONS DE L'ÉCLCAIRAGE

Résumé

Dans le travail on a présenté l'analyse qualitative et quantitative des caractéristiques du déséquilibre de la capacité du condensateur type MIS dans les conditions de l'éclairage.

S. KRAWCZYK

DIE UNGLEICHWERTIGEN EIGENSCHAFTEN DES MIS-KONDENSATORS IN BELEUCHTUNGSVERHÄLTNISSEN

Zusammenfassung

In der Bearbeitung wurde die qualitative und quantitative Analyse ungleichwertiger Eigenschaften des MIS-Kondensators in Beleuchtungsverhältnissen präsentiert.

С. КРАВЧИК

НЕРАВНОВЕСНЫЕ СВОЙСТВА КОНДЕНСАТОРА ТИПА MIS В УСЛОВИЯХ ОСВЕЩЕНИЯ

Резюме

Представлен качественный и количественный анализ неравновесных процессов в конденсаторе MIS в условиях освещения.

ROZPRAWY ELEKTROTECHNICZNE 1982, 28, z. 1-2, ss. 233-245

621.372.8:621.375.826:535.3

Sprawność sprzężenia lasera złączowego ze światłowodem włóknistym za pomocą mikrosoczewki

WŁODZIMIERZ NAKWASKI (ŁÓDŹ)

Instytut Fizyki Politechniki Łódzkiej

Otrzymano 3.5.1980

Przedmiotem niniejszej pracy jest analiza strat promieniowania przy sprzęganiu lasera złączowego ze szklanym włóknem światłowodowym za pomocą mikrosoczewek. Umożliwia ona obliczenie sprawności sprzężenia, jak również daje wskazówki, w jaki sposób ją zwiększyć przez odpowiedni dobór parametrów układu.

1. WPROWADZENIE

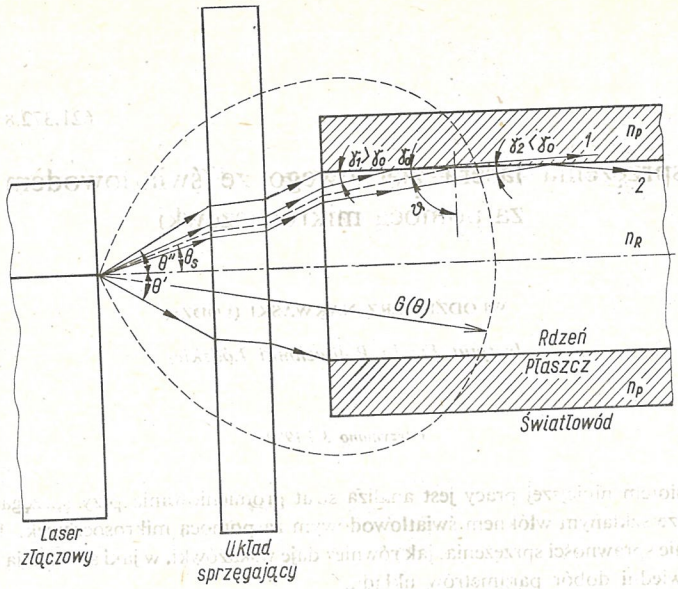
Podstawowym zastosowaniem w przyszłości laserów złączowych wydaje się być telekomunikacja optyczna [1 ÷ 3]. Będzie to, przynajmniej na Ziemi, telekomunikacja światłowodowa, wykorzystująca do przesyłania sygnału świetlnego szklane włókna światłowodowe [4, 5], ponieważ tłumienność promieniowania świetlnego w atmosferze ziemskiej jest silną funkcją warunków atmosferycznych [6].

Poważnym problemem związanym z zastosowaniem w telekomunikacji optycznej światłowodów włóknistych jest sprawność ich sprzężenia z podstawowym źródłem świetlnej fali nośnej, to jest laserem złączowym. W celu zwiększenia powyższej sprawności stosuje się często specjalne układy sprzęgające, opisane w poprzedniej pracy autora [7].

Najczęściej takimi układami sprzęgającymi są specjalne mikrosoczewki ogniskujące promieniowanie lasera złączowego na frontowej płaszczyźnie wejściowej końcówki światłowodu. Celem niniejszego artykułu jest przedstawienie metod, pozwalających obliczyć sprawność sprzężenia laser złączowy — światłowód za pomocą mikrosoczewek oraz dających wskazówki, co do sposobów jej zwiększenia.

2. DEFINICJA SPRAWNOŚCI SPRZĘŻENIA

Na wartość sprawności sprzężenia laser złączowy — szklane włókno światłowodowe niezwykle istotny wpływ ma zarówno niekorzystna charakterystyka kierunkowa emitowanego przez laser promieniowania, cechująca się jak wiadomo znaczną rozbieżnością, jak również trudności związane z wprowadzeniem promieniowania do światłowodu,



Rys. 1. Sprzężenie lasera złączowego ze szklanym włóknem światłowodowym. Układ sprzęgający zaznaczono symbolicznie. Poszczególne symbole zostały wyjaśnione w treści pracy pod wzorem (1). Promień *1* emitowany przez laser pod kątem Θ_1 większym od kąta sprzężenia Θ_s, pada na granicę rdzeń-płaszcz światłowodu pod kątem γ_1 większym od kąta odbicia γ_0 i przenika przez tę granicę. Natomiast promień 2, emitowany w kierunku leżącym wewnątrz kąta sprzężenia, jest odbijany od granicy rdzeń-płaszcz i może być przesyłany przez światłowód

a wynikające z jego małej średnicy rdzenia i niskiej wartości apertury numerycznej. Pokazano to na rys. 1, ułatwiającym zdefiniowanie sprawności w sposób następujący [8]

$$\eta = \frac{\int_{\Theta'}^{\Theta''} G(\Theta)\,T(\Theta)\,\xi(\Theta)\,d\Theta}{\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} G(\Theta)\,d\Theta}, \tag{1}$$

gdzie:

$G(\Theta)$ — rozkład kątowy promieniowania lasera w strefie dalekiej,

Θ — kąt między kierunkiem rozpatrywanego promieniowania emitowanego przez laser, a normalną do płaszczyzny zwierciadła laserowego, którą przy prawidłowym ustawieniu jest oś układu,

$T(\Theta)$ — współczynnik transmisji promieniowania przez granicę powietrze — szkło (w przybliżeniu $T(\Theta) \approx T = \frac{4n_R}{(n_R+1)^2}$, gdzie n_R — współczynnik załamania materiału rdzenia światłowodu; $n_R = 1{,}5$: $T(\Theta) = 0{,}96$),

Θ' i Θ'' — kąty między kierunkiem promieni padających na granicę rdzeń-płaszcz światłowodu na jego powierzchni frontowej, a normalną do zwierciadła lasera,

$\xi(\Theta)$ — współczynnik sprzężenia równy 1 dla $\Theta < \Theta_s$ oraz równy 0 dla $\Theta > \Theta_s$.

Kąt Θ_s jest kątem sprzężenia, będącym maksymalnym kątem Θ, dla którego promieniowanie wniknie do światłowodu i będzie odbijane od granicy rdzeń-płaszcz, to jest będzie się rozchodzić względem osi światłowodu pod kątem γ mniejszym od kąta odbicia[1] γ_o równym

$$\gamma_o = \arccos(n_P/n_R), \tag{2}$$

gdzie przez n_P i n_R oznaczono współczynniki załamania odpowiednio materiału płaszcza i rdzenia światłowodu. Dla kątów Θ zawartych w granicach między Θ_s i Θ' (bądź Θ'') promieniowanie dostanie się do światłowodu lecz, padając na granicę rdzeń-płaszcz pod kątem γ większym od kąta odbicia γ_o, nie jest od niej odbijane i nie może być dlatego propagowane w światłowodzie. Promieniowanie to wnika do płaszcza światłowodu, co pokazano na rys. 1, i zostaje, z punktu widzenia jego użyteczności w telekomunikacji optycznej, stracone.

Zależność (1) jest słuszna w przypadku promieniowania rozbieżnego (z rozkładem natężenia $G(\Theta)$) w jednej płaszczyźnie i ściśle skolimowanego w płaszczyźnie do niej prostopadłej. Jest to założenie bliskie rzeczywistości w przypadku laserów złączowych o konstrukcji paskowej stosowanych w telekomunikacji optycznej, których rozkład kątowy promieniowania w strefie dalekiej wykazuje podobną asymetrię. Mianowicie rozbieżność promieniowania tych laserów w płaszczyźnie prostopadłej do złącza *p-n* jest, w wyniku ugięcia na szczelinie o szerokości równej około 0,1 μm, znacznie większa niż w płaszczyźnie równoległej do złącza, dla której analogiczna szczelina ma szerokość około 10 μm.

Najprostszym rodzajem sprzężenia jest sprzężenie bezpośrednie polegające na bezpośrednim przystawieniu końcówki światłowodu do emitującego promieniowanie lasera złączowego. Wówczas zakładając, że laser usytuowany jest dokładnie na osi światłowodu oraz że rozmiary powierzchni emitującej promieniowanie są dużo mniejsze od średnicy rdzenia światłowodu, można sprawność sprzężenia bezpośredniego wyrazić wzorem [9]

$$\eta_B = T \cdot \operatorname{erf}(2\Theta_s/\beta), \tag{3}$$

gdzie przez erf (x) oznaczono całkę prawdopodobieństwa, której tablica podana jest na przykład w poradniku [10]. Dodatkowym stosowanym tu założeniem jest przyjęcie gaussowskiego kształtu funkcji $G(\Theta)$ o szerokości kątowej β między kierunkami o natężeniu e-krotnie mniejszym od natężenia promieniowania osiowego, gdzie e — podstawa logarytmów naturalnych, oraz założenie $T(\Theta) = T$.

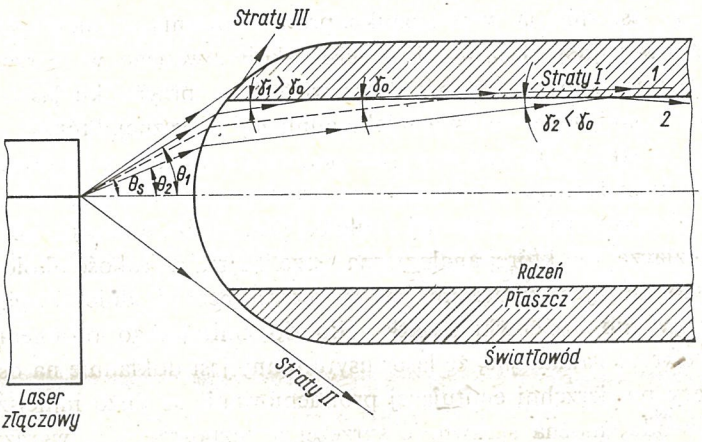
Dla standardowych światłowodów, w przypadku których kąty sprzężenia zawarte są w granicach od 5° do 10° oraz dla laserów złączowych o szerokości kątowej charakterystyki kierunkowej promieniowania $\beta = 50° \div 70°$ otrzymane za pomocą wzoru (3) wartości sprawności sprzężenia bezpośredniego zawarte są w przedziale od 0,15 do 0,40. Realne wartości sprawności sprzężenia η_B są mniejsze, ponieważ przedstawione powyżej założenia nie są ściśle spełnione, przede wszystkim rozmiary powierzchni emitującej promieniowanie laserowe są zwykle porównywalne ze średnicami rdzeni standardowych światłowodów stosowanych w telekomunikacji optycznej.

[1] Zwykle w optyce używa się pojęcia kąta całkowitego wewnątrznego odbicia ϑ, tj. wielkości, będącej dopełnieniem kąta odbicia γ_o do kąta prostego. Oba kąty γ_o i ϑ zostały pokazane na rys. 1.

Wzór (3) może być również stosowany do szacunkowej oceny wartości sprawności sprzężenia w przypadku stosowania bardziej złożonych układów sprzęgających, z tym, że kąt sprzężenia Θ_s będzie w tym przypadku znacznie większy w wyniku ogniskowania promieniowania przez powyższe układy.

3. ANALIZA STRAT PROMIENIOWANIA PRZY SPRZĘŻENIU ZA POMOCĄ MIKROSOCZEWEK PÓŁSFERYCZNYCH LUB PÓŁCYLINDRYCZNYCH

Najczęściej stosowana metoda zwiększania sprawności sprzężenia lasera złączowego ze światłowodem polega na wytwarzaniu na końcu włókna szklanego skupiającej promieniowanie mikrosoczewki. Układ taki pokazano na rys. 2.



Rys. 2. Sprzężenie lasera złączowego ze światłowodem za pomocą mikrosoczewki półcylindrycznej bądź półsferycznej

Na rysunku zaznaczono trzy rodzaje strat promieniowania opisane w treści pracy

Straty promieniowania emitowanego przez laser złączowy przy sprzężeniu ze światłowodem tą metodą (oprócz odbicia części promieniowania na granicy powietrze-szkło) występują w trzech przypadkach [11]:

1) gdy promieniowanie po załamaniu w soczewce pada na granicę rdzeń-płaszcz pod kątem γ większym od kąta odbicia γ_o (straty I),
2) gdy promieniowanie ominie soczewkę (straty II),
3) gdy promieniowanie pada na soczewkę poza rdzeniem światłowodu (straty III).

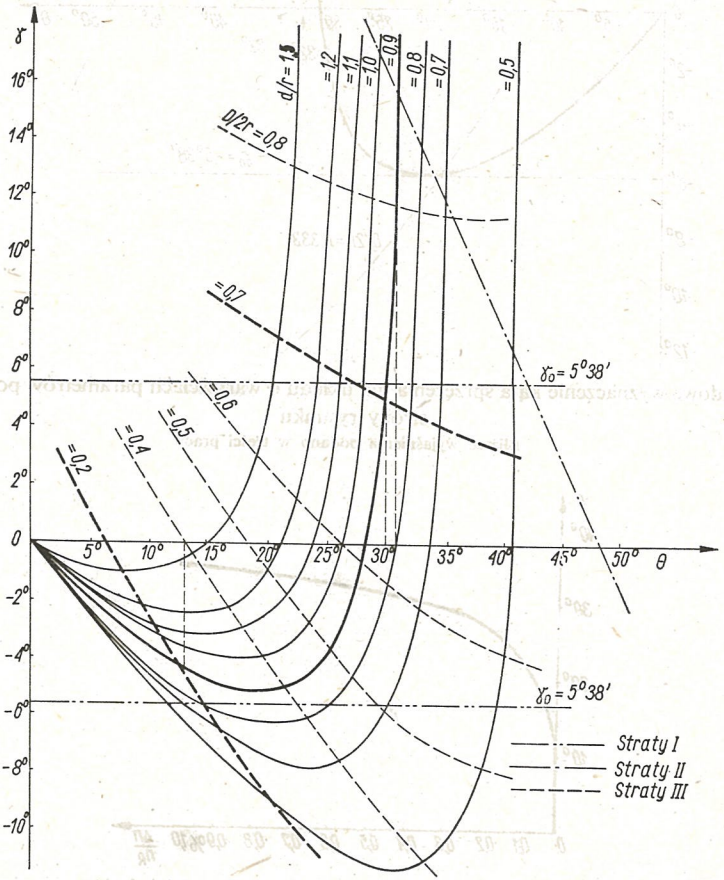
Zależność kąta γ od kąta padania Θ została w postaci wykresu przedstawiona na rys. 3. Parametrem na tym wykresie jest stosunek d/r, gdzie przez d oznaczono odległość soczewki od zwierciadła lasera złączowego, a przez r — promień krzywizny soczewki.

Metoda wyznaczania kąta sprzężenia Θ_s została wyjaśniona na rys. 4, przykładowo dla układu o następujących wartościach parametrów [11]: $D = 80$ μm, $d = 100$ μm, $r = 120$ μm, $n_R = 1{,}5000$ i $n_P = 1{,}4925$, gdzie przez D oznaczono średnicę rdzenia światłowodu. Obliczony stąd kąt odbicia jest równy $\gamma_o \approx 5°38'$. Odpowiadają mu na rys. 4 dwie poziome linie ($\pm\gamma_o$) wyznaczające zakres zmian kąta padania Θ, dla którego kąt γ jest

mniejszy od kąta γ_o (straty I). Jak widać z tego rysunku, z punktu widzenia strat I kąt sprzężenia byłby równy 32°.

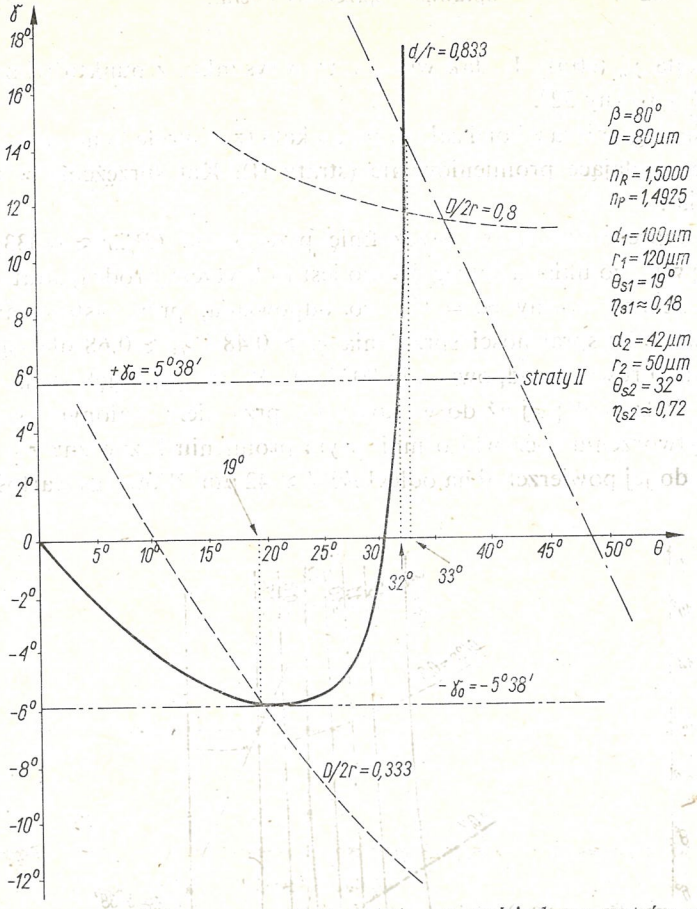
Linia ukośna ogranicza z kolei zakres zmian kąta Θ ze względu na ewentualne ominięcie soczewki przez padające promieniowanie (straty II). Kąt sprzężenia w tym przypadku byłby równy 33°.

Straty III reprezentowane są przez linię przerywaną ($D/2r \approx 0{,}333$). Okazuje się, że dla rozpatrywanego układu sprzęgającego jest to krytyczny rodzaj strat. Stąd kąt sprzężenia jest ostatecznie równy $\Theta_s = 19°$, co odpowiada, przy zastosowaniu wzoru (3), szacunkowej wartości sprawności sprzężenia $\eta_s \approx 0{,}48$ i $\eta_s \approx 0{,}68$ dla kąta rozbieżności promieniowania β równego odpowiednio 80° i 50°. W celu jej zwiększenia należy podwyższyć stosunek $D/2r$, najlepiej aż do wartości 0,8, przy niezmienionym stosunku d/r. Odpowiada to wytworzeniu soczewki o mniejszym promieniu krzywizny $r = 50$ μm i przybliżeniu lasera do jej powierzchni na odległość $d \approx 42$ μm. Wówczas wartość kąta sprzęże-



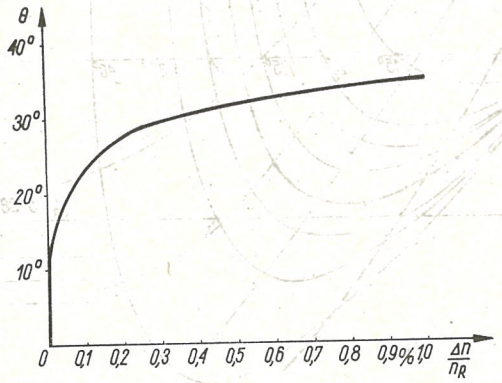
Rys. 3. Zależność kąta γ, pod jakim w stosunku do osi układu promieniowanie rozchodzi się wewnątrz światłowodu, od kąta padania Θ w przypadku sprzężenia za pomocą mikrosoczewki półcylindrycznej bądź półsferycznej [11].

Parametrem na wykresie jest stosunek d/r. Linia ukośna symbolizuje ograniczenie wartości kąta sprzężenia związane ze stratami II, zaś linie przerywane — analogiczne ograniczenie związane ze stratami III dla różnych wartości stosunku $D/2r$

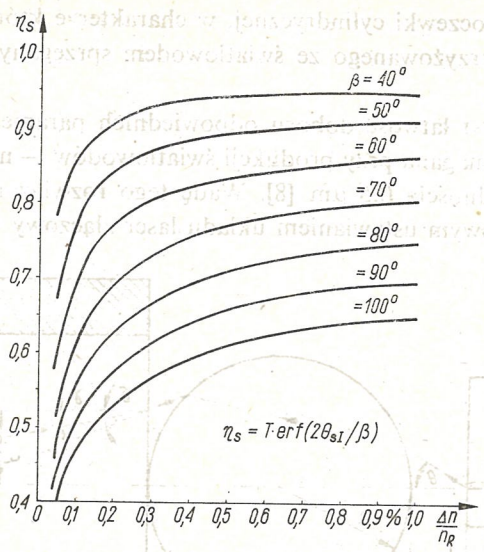


Rys. 4. Przykładowe wyznaczenie kąta sprzężenia dla układu o wartościach parametrów podanych z prawej strony rysunku

Bliższe wyjaśnienia podano w treści pracy



Rys. 5. Zależność kąta sprzężenia Θ_{s1} związanego ze stratami I od wartości skokowej zmiany współczynnika załamania $\Delta n = n_R - n_P$ na granicy rdzeń-płaszcz światłowodu przy sprzężeniu za pomocą mikrosoczewki półsferycznej bądź półcylindrycznej [11]



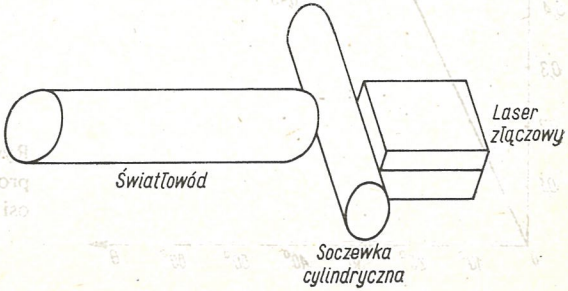
Rys. 6. Zależność sprawności sprzężenia η_s za pomocą mikrosoczewki półsferycznej bądź półcylindrycznej od skokowej zmiany współczynnika załamania Δn na granicy rdzeń-płaszcz światłowodu dla różnych wartości szerokości kątowej β rozkładu kątowego promieniowania lasera złączowego w strefie dalekiej W obliczeniach założono, że kąt sprzężenia Θ_{sI} związany ze stratami I jest nie większy od analogicznych kątów związanych ze stratami II i III. Można to uzyskać przez odpowiednie zaprojektowanie układu sprzęgającego, co pokazano na rys. 4

nia zostanie zwiększona do 32°, a szacunkowe wartości sprawności sprzężenia do 0,71 i do 0,89, odpowiednio dla $\beta = 80°$ i $\beta = 50°$. Inny przykład doboru parametrów układu sprzęgającego przedstawiono w dodatku.

Zależność kąta Θ_s od skokowej zmiany współczynnika załamania na granicy rdzeń-płaszcz $\Delta n = n_R - n_P$ została pokazana na rys. 5. Rys. 6 przedstawia z kolei wyznaczoną teoretycznie analogiczną zależność sprawności sprzężenia η_s przy założeniu gaussowskiego kształtu funkcji $G(\Theta)$, opisującej kątowy rozkład natężenia promieniowania w strefie dalekiej z szerokościami połówkowymi $\beta = 40°$, 50°, 60°, 70°, 80°, 90° i 100°.

4. ANALIZA STRAT PROMIENIOWANIA PRZY SPRZĘŻENIU ZA POMOCĄ SOCZEWKI CYLINDRYCZNEJ

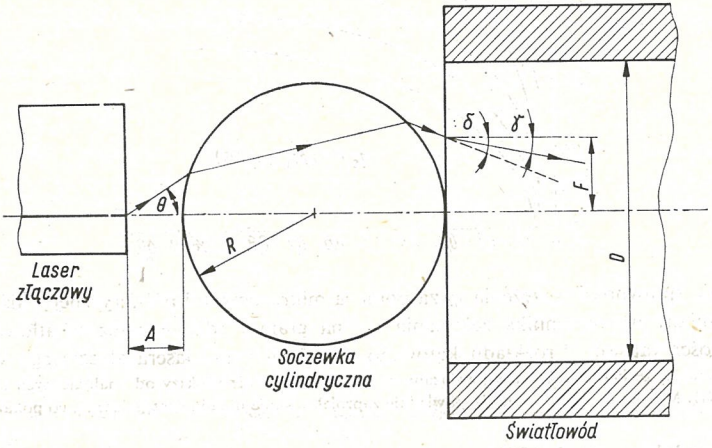
Kłopoty technologiczne związane z koniecznością wykonywania na końcach światłowodów mikrosoczewek o zadanych z góry parametrach spowodowały zainteresowanie się projektantów układów telekomunikacyjnych inną metodą sprzęgania. Polega ona na



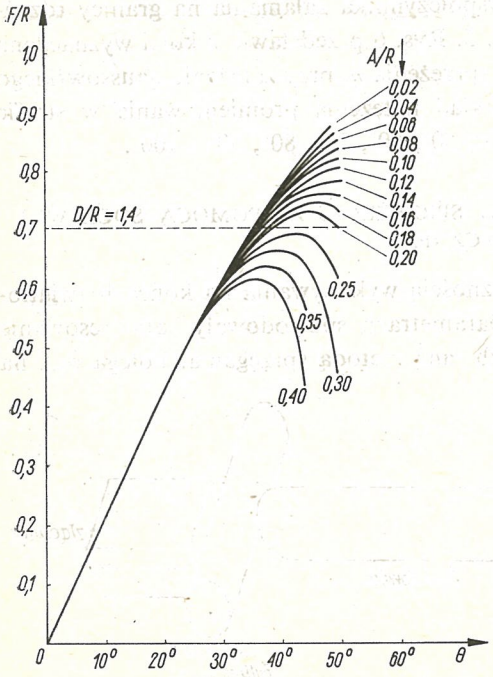
Rys. 7. Sprzężenie lasera złączowego ze szklanym włóknem światłowodowym za pomocą soczewki cylindrycznej

sprzężeniu za pomocą soczewki cylindrycznej, w charakterze której używa się krótkiego odcinka światłowodu skrzyżowanego ze światłowodem sprzęganym w sposób pokazany na rys. 7.

Zaletą tej metody jest łatwość doboru odpowiednich parametrów soczewki z uwagi na dużą dokładność wymaganą przy produkcji światłowodów — na przykład ich średnica wykonana jest z dokładnością 0,5 μm [8]. Wadę tego rozwiązania stanowią natomiast kłopoty związane z osiowym ustawianiem układu laser złączowy — soczewka cylindrycz-



Rys. 8. Parametry układu sprzęgającego laser złączowy ze światłowodem za pomocą soczewki cylindrycznej [9].
Poszczególne symbole wyjaśniono w treści pracy



Rys. 9. Zależność odległości F punktu wnikania promienia świetlnego do światłowodu od jego osi w układzie sprzęgającym pokazanym na rys. 8 od kąta padania Θ [9].
Parametrem na wykresie jest stosunek A/R

na — światłowód włóknisty. Sprawność sprzężenia przy użyciu soczewki cylindrycznej zależy od odpowiedniego doboru parametrów układu (rys. 8): odległości laser-soczewka A, promienia soczewki R i średnicy rdzenia światłowodu D.

Na rys. 9 przedstawiono wykres odległości F punktu wnikania promienia światła do światłowodu od jego osi w funkcji kąta padania Θ. Zmiennym parametrem na tym wykresie jest stosunek A/R.

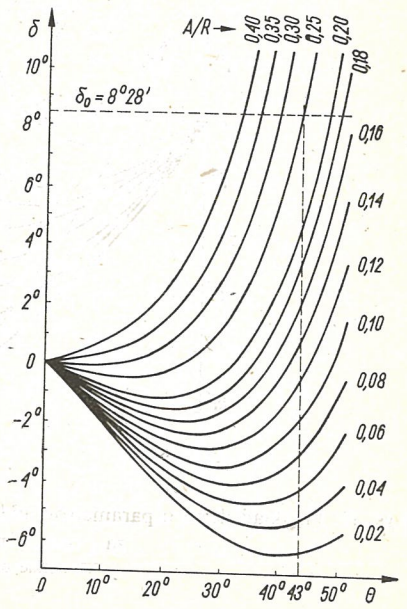
Przyjmując przykładowo średnicę rdzenia światłowodu $D = 1{,}4\ R$ oraz szerokość kątową $\beta = 80°$ charakterystyki kierunkowej promieniowania lasera w strefie dalekiej, otrzymuje się stąd (rys. 9) minimalną wartość odległości soczewki od lasera $A_{min} = 0{,}25\ R$, niezbędną aby całe (poza odbiciem na granicach szkło-powietrze) promieniowanie padające na soczewkę mogło się dostać do światłowodu. Wykres ten umożliwia więc ograniczenie z dołu wartości A.

Ograniczenie z góry wynika z kolei z faktu ewentualnego ominięcia przez część promieniowania laserowego soczewki. Maksymalny kąt padania, dla którego promieniowanie dostanie się do soczewki, określony jest zależnością

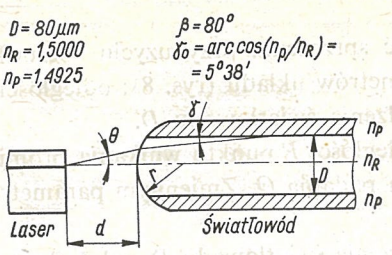
$$\sin\Theta_{max} = \left(\frac{A}{R} + 1\right)^{-1}. \tag{4}$$

Dla rozpatrywanego tutaj przykładu, przyjmując $\Theta_{max} = \frac{\beta}{2}$, mamy $\sin 40° \approx 0{,}6428$ i stąd $A_{max} \approx 0{,}556\ R$. W celu otrzymania wydajnego sprzężenia lasera złączowego ze światłowodem należy go więc w tym przypadku umieścić w odległości $A \in \langle 0{,}25\ R, 0{,}556\ R\rangle$ od soczewki.

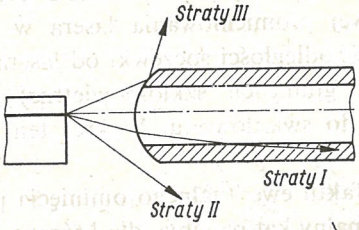
Na rys. 10 pokazano zależność kąta δ, pod jakim promieniowanie po wyjściu z soczewki pada na frontową powierzchnię światłowodu, od kąta Θ, pod jakim promieniowanie

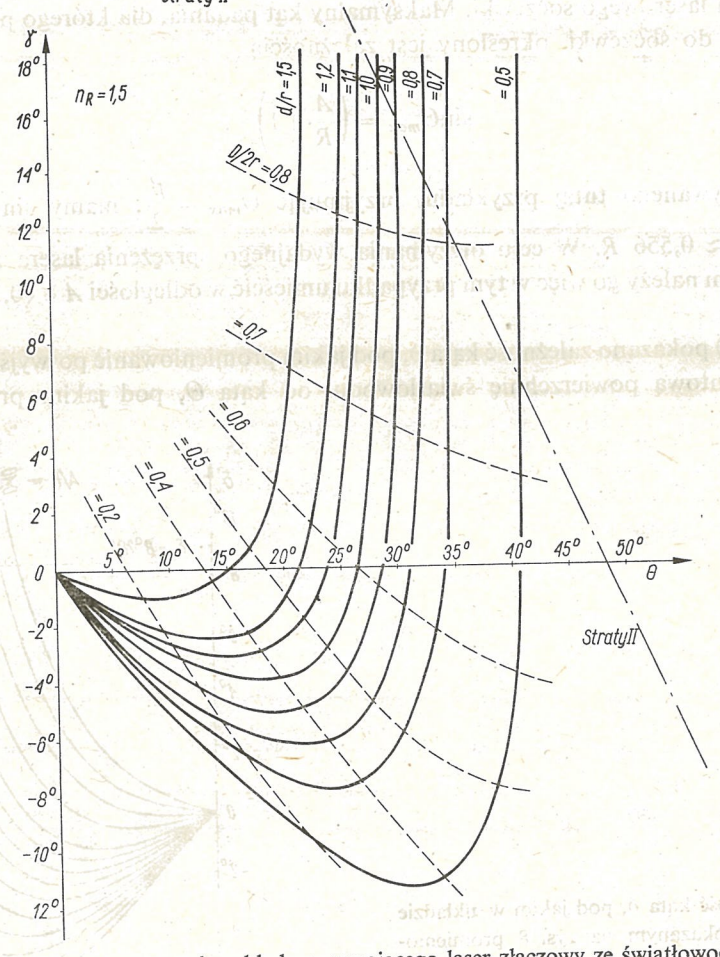


Rys. 10. Zależność kąta δ, pod jakim w układzie sprzęgającym pokazanym na rys. 8 promieniowanie po wyjściu z soczewki pada na frontową powierzchnię światłowodu od kąta padania Θ [9]. Na rysunku wyjaśniono opisaną w treści pracy metodę wyznaczania wartości kąta sprzężenia Θ_s



	1	2
d (μm)	180	51
r (μm)	200	57
d/r	0,9	0,9
D/2r	0,2	0,7
θ_I (...°)	30	30
θ_{II} (...°)	31	31
θ_{III} (...°)	13	30
θ_S (...°)	13	30
η_s	0,18	0,75





Rys. 11. Przykład doboru parametrów układu sprzęgającego laser złączowy ze światłowodem włóknistym za pomocą mikrosoczewki półsferycznej [16].

Wszystkie użyte symbole wyjaśniono na rysunku

wychodzi z lasera złączowego. Światłowód będzie przesyłał tylko promieniowanie padające na granicę rdzeń-płaszcz pod kątem γ mniejszym od kąta odbicia γ_o. Dla światłowodu, którego współczynniki załamania materiału rdzenia i płaszcza są odpowiednio równe $n_R = 1{,}5000$ i $n_P = 1{,}4925$, kąt odbicia $\gamma_o \approx 5°38'$. Stosując prawo Snelliusa

$$\frac{\sin\delta}{\sin\gamma} = \frac{n_R}{n_o}, \tag{5}$$

gdzie: n_o — współczynnik załamania środowiska, najczęściej powietrza ($n_o = 1$), otrzymuje się stąd graniczny kąt padania $\delta_o \approx 8°28'$. Rozpatrywany światłowód przesyła tylko to promieniowanie, którego kąt padania na frontową powierzchnię światłowodu po wyjściu z soczewki cylindrycznej zawiera się w granicach $\delta \in \langle -\delta_o, \delta_o \rangle$. Kątowi δ_o odpowiada z kolei kąt sprzężenia Θ_s. Jak widać na rys. 10 wzrost stosunku A/R powoduje malenie kąta sprzężenia. Wybieramy więc najmniejszą dopuszczalną wartość A: dla rozpatrywanego przykładu $A_{opt} = 0{,}25\,R$. Odpowiada jej (jak widać na rys. 10) kąt sprzężenia $\Theta_s \approx 43°$.

Dużą wartość sprawności sprzężenia uzyskuje się w przypadku, gdy

$$\Theta_s > \beta/2. \tag{6}$$

W rozpatrywanym przykładzie dla kątów rozbieżności promieniowania laserowego $\beta = 80°$ i $\beta = 50°$ stosując wzór (3) otrzymujemy szacunkowe wartości sprawności sprzężenia odpowiednio $\eta_c \approx 0{,}84$ i $\eta_c \approx 0{,}95$.

5. PODSUMOWANIE

W niniejszej pracy przedstawiono analizę strat promieniowania przy sprzęganiu lasera złączowego ze szklanym włóknem światłowodowym za pomocą mikrosoczewek półsferycznych i półcylindrycznych (rozdział 3), bądź soczewek cylindrycznych (rozdział 4). Wnioski wynikające z powyższej analizy umożliwiają obliczenie sprawności sprzężenia, jak również dają wskaźniki dotyczące odpowiedniego doboru parametrów układu, powodującego zwiększenie wyżej wymienionej sprawności.

Ostatnio obserwuje się powrót projektantów układów telekomunikacji optycznej do sprzężenia bezpośredniego w zmodyfikowanej postaci [12÷15]. Zagadnieniu temu będzie poświęcona w przyszłości osobna praca.

DODATEK

Na rys. 11. przedstawiono przykład doboru parametrów układu sprzęgającego laser złączowy ze szklanym włóknem światłowodowym za pomocą mikrosoczewki półsferycznej.

LITERATURA

1. S. J. Buchsbaum, *Lightwave communications. An overview*, Physics Today, Vol. 29, No 5, 1976, pp. 23—25.
5. A. Smoliński, *Światłowody i ich zastosowania*. I Krajowe Sympozjum nt. Światłowody i ich zastosowania, Jabłonna 1976, PWN, Warszawa 1977, ss. 19—42.

3. W. Nakwaski, *Lasery złączowe w telekomunikacji optycznej*. Rozprawy Elektrotechniczne, z. 4, 1981.
4. D. Gloge, *Optical fibers for communication*, Appl. Optics, Vol. 13, No 2, 1974, pp. 249—254.
5. H. Klejman, *Telekomunikacja torami światłowodowymi*, Przegląd Telekomunikacyjny, Vol. 51, No 4, 1978, ss. 111—117.
6. H. Klejman, *Tory telekomunikacyjne optyczne*, Przegląd Telekomunikacyjny, Vol. 49, No 6, 1976, ss. 173—176.
7. W. Nakwaski, *Sprzężenie lasera złączowego ze szklanym włóknem światłowodowym*, Przegląd Telekomunikacyjny, Vol. 53, No 10, 1980, ss. 341—345.
8. M. Maeda, I. Ikushima, K. Nagano, M. Tanaka, H. Nakashima, R. Itoh, *Hybrid laser-to-fibre coupler with a cylindrical lens*, Appl. Optics, Vol. 16, No 7, 1977, pp. 1966—1970.
9. E. Weidel, *Light coupling problems for GaAs laser-multimode fibre coupling*, Optical and Quantum Electronics, Vol. 8, 1976, pp. 301—307.
10. I. N. Bronsztejn, K. A. Siemiendiajew, *Matematyka. Poradnik encyklopedyczny*, PWN, Warszawa 1970, ss. 94—95.
11. C. A. Brackett, *On the efficiency of coupling light from stripe-geometry GaAs lasers into multimode optical fibers*, J. Appl. Phys., Vol. 45, No 6, 1974, pp. 2636—2637.
12. M. Maeda, K. Nagano, M. Tanaka, K. Chiba, *Buried-heterostructure laser packaging for wideband optical transmission systems*, IEEE Trans. on Comunications, Vol. COM-26, No 7, 1978, pp. 1076—1081.
13. Ozeki T., B. S. Kawasaki, *Efficient power coupling using taper-ended multimode optical fibres*, Electronics Letters, Vol. 12, No 23, 1976, pp. 607—608.
14. A. J. Spring Thorpe, M. J. Rider, *Novel double-heterostructure lasers*, International Electron Devices Meeting, Washington 1977, pp. 571—574.
15. A. J. Spring Thorpe, *A novel double-heterostructure p-n-junction laser*, Appl. Phys. Lett., Vol. 31, No 8, 1977, pp. 524—525.
16. W. Nakwaski, *Wyznaczanie sprawności sprzężenia lasera złączowego ze światłowodem włóknistym*, Konferencja „Fizyka dla przemysłu", Kraków 8—10.IX.1980, ss. 299—300.

W. NAKWASKI

THE EFFICIENCY OF COUPLING LIGHT FROM A LASER DIODE INTO AN OPTICAL FIBER WITH THE AID OF A LENS

Summary

The analysis of radiative losses during laser diode-optical fiber coupling is a subject of this paper. It enables one to calculate an efficiency of coupling as well as shows how this efficiency can be increased with the aid of a suitable change of system parameters.

W. NAKWASKI

RENDEMENT DU COUPLAGE D'UN LASER À JONCTION AU CONDUIT DE LUMIÈRE FIBREUX À L'AIDE D'UNE MICROLENTILLE

Résumé

L'étude se pose pour but d'analyser les pertes de rayonnement lors du couplage d'un laser à jonction au conduit de lumière en fibre de verre à l'aide des microlentilles. L'étude contribue à faciliter le calcu du rendement du couplage et donne des indications sur la possibilité d'augmenter ce rendement par le choix convenable des paramètres du système.

W. NAKWASKI

LEISTUNGSFÄHIGKEIT DER KOPPLUNG EINES VERBINDUNGSGLASES MIT EINEM FASRIGEN LICHTLEITER MITTELS EINER MIKROLINSE

Zusammenfassung

Der Stoff der vorliegenden Bearbeitung ist die Verlustanalyse der Strahlung bei der Kopplung eines Verbindungsglases mit lichtleitender Glasfaser mittels Mikrolinsen. Diese Analyse ermöglicht die Berechnung der Leistungsfähigkeit der Kopplung, erteilt auch Hinweise, wie die Leistungsfähigkeit erhöht werden kann durch eine entsprechende Wahl der Systemparameter.

В. НАКВАСКИ

К.П.Д. СЦЕПЛЕНИЯ С ПОМОЩЬЮ ЛИНЗЫ ИНЖЕКЦИОННОГО ЛАЗЕРА СО СТЕКЛЯННЫМ ВОЛОКОННЫМ СВЕТОВОДОМ

Резюме

Проведен анализ потерь излучения при сцеплении инжекционного лазера со стеклянным волоконным световодом с помощью линз. Этот анализ дает возможность вычисления к.п.д. сцепления и показывает возможности повышения к.п.д. при помощи соответственного изменения параметров системы.

ROZPRAWY ELEKTROTECHNICZNE 1982, 28, z. 1-2, ss 247-259

621.395.3

Analiza wpływu prądu obciążenia na sterujący strumień magnetyczny w kontaktronie

JERZY KACZMAREK (GDAŃSK)

Instytut Technologii Elektronicznej Politechniki Gdańskiej

Otrzymano 3.6.1980

Przedstawiono metodę obliczania zmian wielkości strumienia magnetycznego jakie wywołuje przepływ prądu obciążenia przez kontaktron. Przeanalizowano oddziaływanie magnetyczne pomiędzy obwodem obciążenia a obwodem sterującym oraz zjawiska termiczne. Określono zakres maksymalnych obciążeń prądowych kontaktronów.

1. WSTĘP

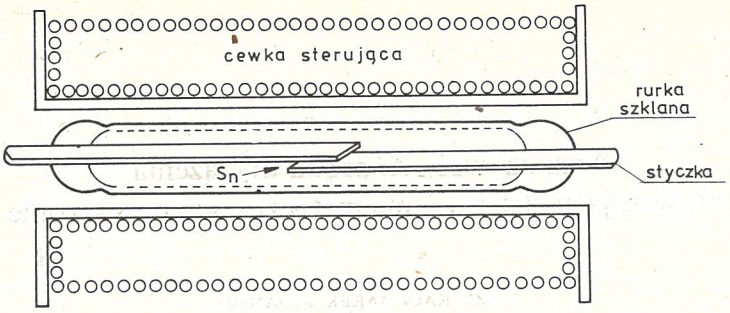
Zakres zastosowań elementów stykowych nie zmniejsza się pomimo konkurencji elementów półprzewodnikowych. Wynika to z licznych zalet elementów stykowych a zwłaszcza: możliwości galwanicznego rozwarcia obwodu elektrycznego, prostoty sterowania i dużego zakresu obciążalności.

Wśród elementów stykowych szczególne miejsce zajmują kontaktrony, które charakteryzują się dużą niezawodnością i odpornością na narażenia technoklimatyczne. Kontaktrony zostały opracowane z przeznaczeniem dla central telefonicznych, jest to nadal ich największe zastosowanie również w Polsce, zwłaszcza przy wzrastającej produkcji telefonicznych central elektronicznych E-10. Zakres zastosowań kontaktronów jest współcześnie rozszerzany zarówno w kierunku komutacji sygnałów mikromocy, jak również w kierunku dużych obciążeń rzędu 10 kV czy 500 VA.

Powstaje problem, jak określić zakres ekstremalnych obciążeń dla kontaktronów i jakie zjawiska ten zakres ograniczają. W pracy przeanalizowano zjawiska fizyczne, jakie występują w kontaktronach przy dużych obciążeniach prądowych. Zjawiska te powodują zmniejszenie wielkości roboczego strumienia magnetycznego, prowadzą do uszkodzenia elementu i ograniczają zakres maksymalnych obciążeń prądowych dla kontaktronów.

2. WPŁYW PRĄDU OBCIĄŻENIA NA PRACĘ KONTAKTRONU

Kontaktron charakteryzuje się tym, że jego styczki wykonane z materiału ferromagnetycznego spełniają dwie funkcje: są częścią obwodu elektrycznego a jednocześnie stanowią drogę dla przepływu strumienia magnetycznego Φ, jaki wytwarza cewka sterująca (rys. 1).



Rys. 1. Budowa przekaźnika kontaktronowego

Strumień magnetyczny Φ_r przepływając przez powierzchnię nakładki S_n wytwarza siłę przyciągania między styczkami F_p równą

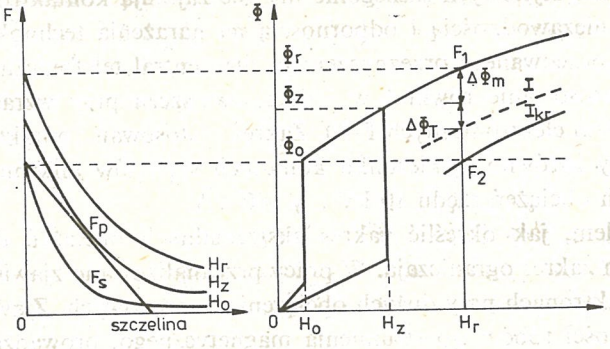
$$F_p = \frac{\Phi_r^2}{2\mu_o S_n}, \tag{1}$$

gdzie:

μ_o — przenikalność magnetyczna próżni,

S_n — powierzchnia przepływu strumienia magnetycznego w nakładce.

Jeżeli siła przyciągania F_p jest większa od siły sprężystości styczek F_s, następuje zadziałanie kontaktronu. Styczki zostają zwarte, gdy w cewce sterującej występuje pole magnetyczne o wartości H_z. W konkretnych warunkach pracy kontaktronu mamy określone i znane natężenie pola w cewce sterującej H_r, strumień roboczy Φ_r i siłę przyciągającą F_p (punkt pracy F_1, rys. 2). Styczki kontaktronu rozewrą się, gdy strumień magnetyczny



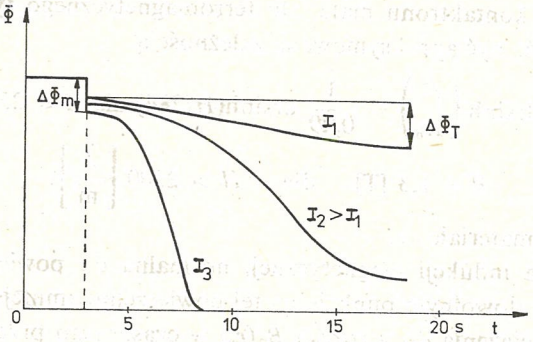
Rys. 2. Wpływ prądu obciążenia na pracę przekaźnika kontaktronowego

zmaleje poniżej wartości strumienia otwarcia Φ_o. Zwykle następuje to w wyniku zmniejszenia w cewce sterującej natężenia pola do wartości H_o. Prąd obciążenia, przepływając przez zamknięte styczki kontaktronu, powoduje jednak zmniejszenie wielkości roboczego strumienia magnetycznego, a tym samym siły przyciągania między styczkami.

Problem ten po raz pierwszy opisany przez T. Harę i współautorów w roku 1974 [1] jest przedmiotem licznych prac [2÷6]. Zmniejszenie strumienia magnetycznego przy do-

statecznie dużym prądzie obciążenia może być tak znaczne, że nastąpi otwarcie styczek kontaktronu. Roboczy strumień magnetyczny zmniejszy się poniżej strumienia otwarcia Φ_o. Punkt pracy kontaktronu przesunie się od punktu F_1 do punktu F_2 (rys. 2). Przerwa w przepływie prądu obciążenia, przy jednoczesnym ciągłym działaniu stałego pola wytwarzanego przez cewkę sterującą, prowadzi po pewnym czasie do ponownego zadziałania kontaktronu. Dochodzi do licznych odskoków styczek, zwanych odskokami prądowymi kontaktronu.

Jak wykazały przeprowadzone badania zmniejszenie strumienia magnetycznego w kontaktronie w czasie przepływu prądu obciążenia następuje w dwóch etapach [3]. W pierwszym etapie strumień zmniejsza się o wartość $\Delta\Phi_m$ bezpośrednio w momencie włączenia prądu (rys. 3). Zmianę tę wywołuje tzw. zjawisko elektrodynamiczne, wynikające z oddziały-

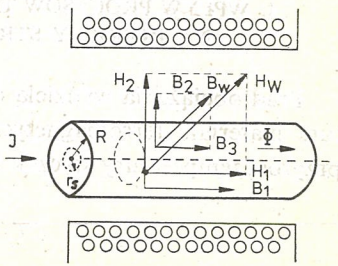


Rys. 3. Zmiany strumienia magnetycznego w kontaktronie spowodowane przepływem prądu obciążenia

wania pomiędzy polem magnetycznym jakie towarzyszy prądowi obciążenia a polem wytworzonym przez cewkę sterującą [1]. W drugim etapie, który trwa przez czas rzędu sekund następuje zmniejszenie strumienia roboczego o wartość $\Delta\Phi_T$. Zmiany te spowodowane są zjawiskami termicznymi, jakie zachodzą w styczkach i w obszarze styku przy przepływie prądu [3, 4].

3. WPŁYW POLA MAGNETYCZNEGO WYTWARZANEGO PRZEZ PRĄD OBCIĄŻENIA NA ROBOCZY STRUMIEŃ MAGNETYCZNY W KONTAKTRONIE

Prąd obciążenia I, przepływając przez kontaktron, wytwarza wewnątrz styczek kołowe pole magnetyczne $\vec{H}_2$ (rys. 4). Przyjmuje się, że styczka ma kształt cylindra o promieniu



Rys. 4. Wykres wektorów pola magnetycznego i indukcji magnetycznej w materiale ferromagnetycznym styczek

R [1]. Moduł wektora $\vec{\mathbf{H}}_2$ zmienia się w zależności od odległości od środka styczki r_s zgodnie z zależnością

$$H_2(r_s) = \frac{Ir_s}{2\pi R^2} \quad \text{dla} \quad r_s \leqslant R. \tag{2}$$

W każdym punkcie przekroju poprzecznego styczki wypadkowy wektor natężenia pola magnetycznego $\overline{\mathbf{H}}_w$ jest sumą wektora $\overline{\mathbf{H}}_2$ i wektora natężenia pola $\overline{\mathbf{H}}_1$, jaki wytwarza cewka sterująca.

W materiale ferromagnetycznym kontaktronu wielkość wypadkowego wektora indukcji magnetycznej $\overline{\mathbf{B}}_w$ zależy od wielkości wypadkowego wektora pola $\overline{\mathbf{H}}_w$, przy czym kierunki obydwu wektorów pokrywają się. Moduł wektora $\mathbf{B}_w$ można obliczyć korzystając z krzywej namagnesowania materiału ferromagnetycznego $B = f(H)$. Dla stosowanego do produkcji styczek kontaktronu materiału ferromagnetycznego typu FeNiCo (kowar) funkcja $B = f(H)$ może być aproksymowana zależnością

$$B = \frac{1}{B_o}\,\text{arsinh}\left(\frac{H}{A_o}\right) = \frac{1}{0{,}49}\,\text{arsinh}(H/284) \text{ dla } H \leqslant 2500\left[\frac{\text{A}}{\text{m}}\right],$$

$$B = 1{,}3\ [\text{T}] \quad \text{dla} \quad H > 2500\left[\frac{\text{A}}{\text{m}}\right], \tag{3}$$

gdzie: A_o, B_o — stałe materiałowe.

Składowa wektora indukcji magnetycznej, normalna do powierzchni przekroju poprzecznego styczki, w dowolnym punkcie na tej powierzchni zmniejsza się od wartości B_1 przy braku prądu obciążenia do wartości $B_3(r_s)$ w czasie jego przepływu (rys. 4). Suma punktowych zmian indukcji w całym przekroju poprzecznym styczki jest równa zmniejszeniu wartości strumienia i wynosi

$$\Delta\Phi_m = \iint_{S_1} [B_1 - B_3(r_s)]\,dS_1 = \iint_{S_1} \left\{ \frac{1}{B_0}\,\text{arsinh}\left(\frac{H_1}{A_o}\right) - \right.$$

$$\left. - \frac{1}{B_o}\,\text{arsinh}\left(\frac{\sqrt{H_1^2 + \dfrac{I^2 r_s^2}{4\pi^2 R^4}}}{A_o}\right) \cos\left[\text{arctg}\left(\frac{Ir_s}{2\pi R^2 H_1}\right)\right] \right\} dS_1, \tag{4}$$

gdzie S_1 — powierzchnia przekroju poprzecznego styczki kontaktronu.

Zależność ta określa zmiany strumienia magnetycznego spowodowane przepływem prądu obciążenia przez kontaktron, które wynikają ze zjawisk magnetycznych [4].

4. WPŁYW PROCESÓW TERMICZNYCH WYWOŁANYCH PRĄDEM OBCIĄŻENIA NA ROBOCZY STRUMIEŃ MAGNETYCZNY W KONTAKTRONIE

Prąd obciążenia wydziela na rezystancji kontaktronu ciepło Joule'a. Wzrasta temperatura materiału ferromagnetycznego styczek T_f. Z badań eksperymentalnych wynika, że przyrost temperatury styczek można określić zależnością

$$\Delta T_f = T_f - T_o = AI\left[1 - \exp\left(\frac{-t}{\tau}\right)\right], \tag{5}$$

gdzie:

T_o — temperatura otoczenia 293 K,

A — stała zależna od wymiaru kontaktronu,

τ — termiczna stała czasowa styczek,

I — prąd obciążenia,

t — czas przepływu prądu.

Stosując miniaturową termoparę do pomiaru temperatury styczki kontaktronu, określono wartości stałych A i τ dla produkowanych w Polsce typów kontaktronów. Dane te zamieszczono w tablicy 1.

Tablica 1

Termiczne stałe kontaktronów

Typ kontaktronu	Współczynnik A [K/A]	Termiczna stała czasowa τ [sek]
ZW-103	24	21
ZM-107	58	12,5
ZM-108	83	7,1

Wzrost temperatury materiału ferromagnetycznego prowadzi zawsze do zmniejszania wartości jego indukcji, która w temperaturze Curie maleje do zera. Zmianę indukcji magnetycznej opisuje równanie [7]

$$\frac{B(T)}{B(T_o)} = \operatorname{tgh}\left[\frac{B(T)}{B(T_o)}\frac{T_c}{T}\right], \tag{6}$$

gdzie:

$B(T_o)$ — indukcja magnetyczna w temperaturze $T_o = 293$ K,

$B(T)$ — indukcja w temperaturze T,

T_c — temperatura Curie równa dla FeNiCo 698 K.

Wzrost temperatury materiału ferromagnetycznego powoduje zmniejszenie strumienia magnetycznego w styczkach kontaktronu o wartość

$$\Delta\Phi_{fo} = B[(T_o) - B(T_f)]S_n. \tag{7}$$

Zmniejszenie strumienia magnetycznego, a tym samym siły przyciągającej F_p, powoduje otwieranie się styczek. Proces ten przebiega wolno w czasie rzędu sekund, co wynika z termicznej stałej czasowej materiału styczek. Pomiędzy wolno otwierającymi się styczkami powstaje statyczny mostek metaliczny z płynnego metalu stykowego [8]. Jego obecność wpływa na lokalny rozkład temperatury w obszarze stykowym w końcowym etapie otwierania się styczek.

Promień statycznego mostka metalicznego o kształcie walca określa wzór [8]

$$a = \frac{\sqrt{L}}{0{,}44}\frac{I}{k}, \tag{8}$$

gdzie:

L — stała Lorenza równa $2{,}4 \cdot 10^{-8} \left[\frac{V}{K}\right]^2$,

k — przewodność cieplna metalu stykowego $\left[\frac{W}{mK}\right]$,

I — prąd płynący przez kontaktron.

Powstanie mostka o temperaturze topnienia materiału stykowego wpływa na wzrost temperatury materiału ferromagnetycznego $T(r)$ w pobliżu styku. Przyrost temperatury materiału w dowolnym punkcie na powierzchni nakładki oddalonym od powierzchni stykowej o odległość r można obliczyć z wyrażenia [9]

$$\arccos\left[\frac{1+\alpha\Delta T(r)}{1+\alpha T_s}\right] = \frac{2}{\pi}\operatorname{arctg}\left(\frac{r}{a}\right)\arccos\left(\frac{1}{1+\alpha T_s}\right), \tag{9}$$

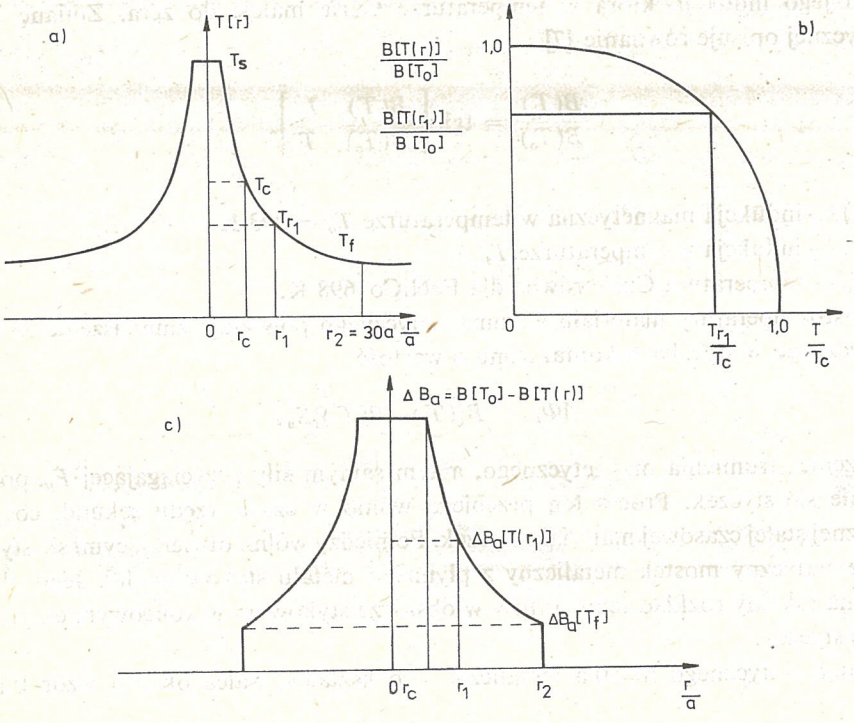
gdzie:

$\Delta T(r) = T(r) - T_o$ dla $r \geqslant a$,

T_s — temperatura topnienia materiału stykowego,

a — promień mostka metalicznego,

α — temperaturowy współczynnik rezystancji materiału ferromagnetycznego.



Rys. 5. Etapy obliczania zmian strumienia magnetycznego spowodowane zjawiskami termicznymi w obszarze stykowym

Na rysunku 5 przedstawiono kolejne etapy obliczenia wielkości zmian strumienia spowodowane zjawiskami termicznymi w obszarze stykowym [10]. Rysunek 5a ilustruje rozkład temperatury w pobliżu styku opisany wyrażeniem (9).

Mostek o promieniu a i temperaturze topnienia T_s znajduje się na powierzchni nakładki kontaktronu. W obszarze r_c przekraczana jest temperatura Curie materiału ferromagnetycznego T_c. Zakłada się, że w odległości $r_2 = 30a$ nie występuje już lokalny wzrost temperatury spowodowany zjawiskami stykowymi. Dlatego rozpatrujemy zmianę indukcji na części powierzchni nakładki $S_2 = \pi r_2^2$. W dowolnym punkcie r_1 na powierzchni S_2 panuje temperatura $T(r_1)$. Wzrost temperatury materiału ferromagnetycznego w tym punkcie prowadzi do zmiany indukcji (wzór 6, rys. 5b) o wartość

$$\Delta B_a[T(r_1)] = B(T_o) - B[T(r_1)]. \tag{10}$$

Zmiany indukcji na powierzchni nakładki w pobliżu styku ilustruje rys. 5c. Płaski odcinek oznacza całkowitą utratę indukcji i wynika z przekroczenia temperatury Curie przez materiał ferromagnetyczny. Zmiana $\Delta B_a(T_f)$ związana jest ze wzrostem temperatury materiału całych styczek T_f. Na rysunku 5c przedstawiono przekrój figury, która w rzeczywistości jest bryłą o symetrii obrotowej.

Zmiany strumienia magnetycznego, spowodowane zjawiskami stykowymi, to suma lokalnych zmian indukcji na całej powierzchni S_2. Określone są one wzorem

$$\Delta\Phi_a = \iint_{S_2} \Delta B_a[T(r)]dS_2. \tag{11}$$

W interpretacji geometrycznej jest to objętość omawianej figury z rys. 5c.

Całkowite zmniejszenie strumienia magnetycznego spowodowane przepływem prądu obciążenia o wartości I oznaczone przez $\Delta\Phi_c(I)$ jest równe sumie zmian strumienia jakie wywołują wszystkie wymienione uprzednio procesy i wynosi

$$\Delta\Phi_c(I) = \iint_{S_1} \Delta B_3(r_s)dS_1 + \iint_{S_2} \Delta B_a[T(r)]dS_2 + [B(T_o) - B(T_f)](S_n - S_2) =$$

$$= \Delta\Phi_m(I) + \Delta\Phi_a(I) + \Delta\Phi_f(I) = \Delta\Phi_m(I) + \Delta\Phi_T(I). \tag{12}$$

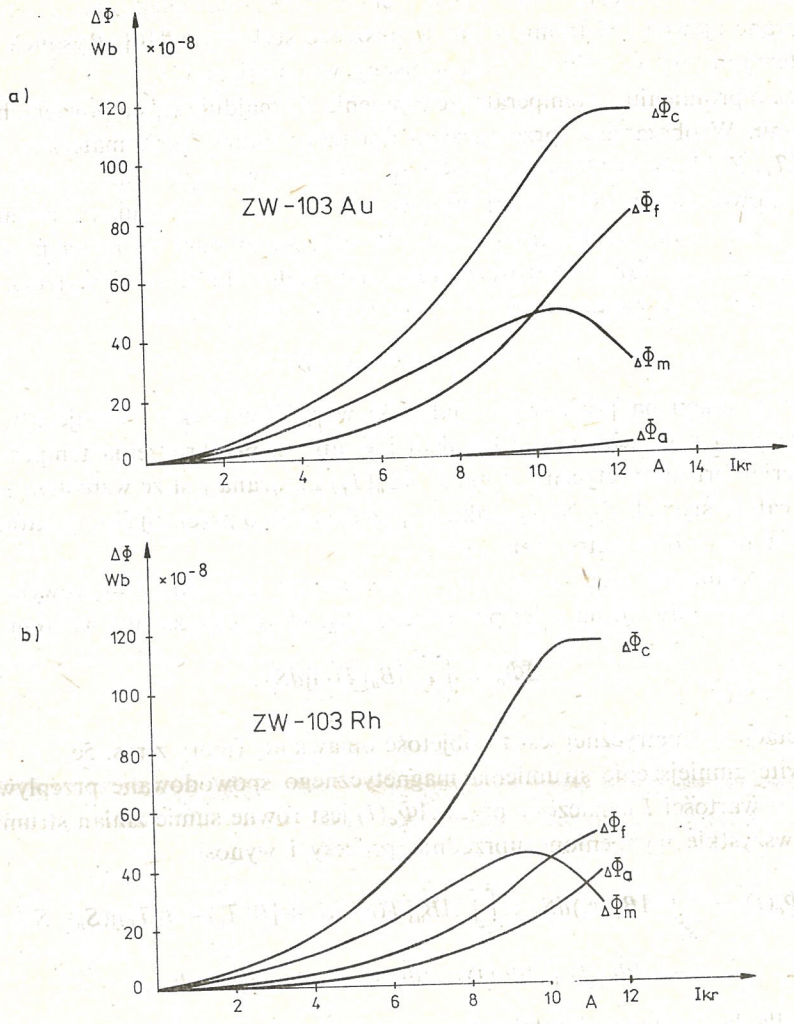
Znając dla konkretnego kontaktronu wartość strumienia roboczego Φ_r i strumienia otwarcia Φ_o można określić wartość prądu krytycznego, przy którym dochodzi do rozwarcia styczek (rys. 2). Dla prądu krytycznego spełniona będzie zależność

$$\Phi_r - \Phi_o = \Delta\Phi_c(I = I_{kr}). \tag{13}$$

Podana metodyka obliczania wartości prądu krytycznego została wykorzystana do określenia zakresu maksymalnych obciążeń prądowych dla kontaktronów.

5. REZULTATY OBLICZEŃ ZMIAN STRUMIENIA MAGNETYCZNEGO

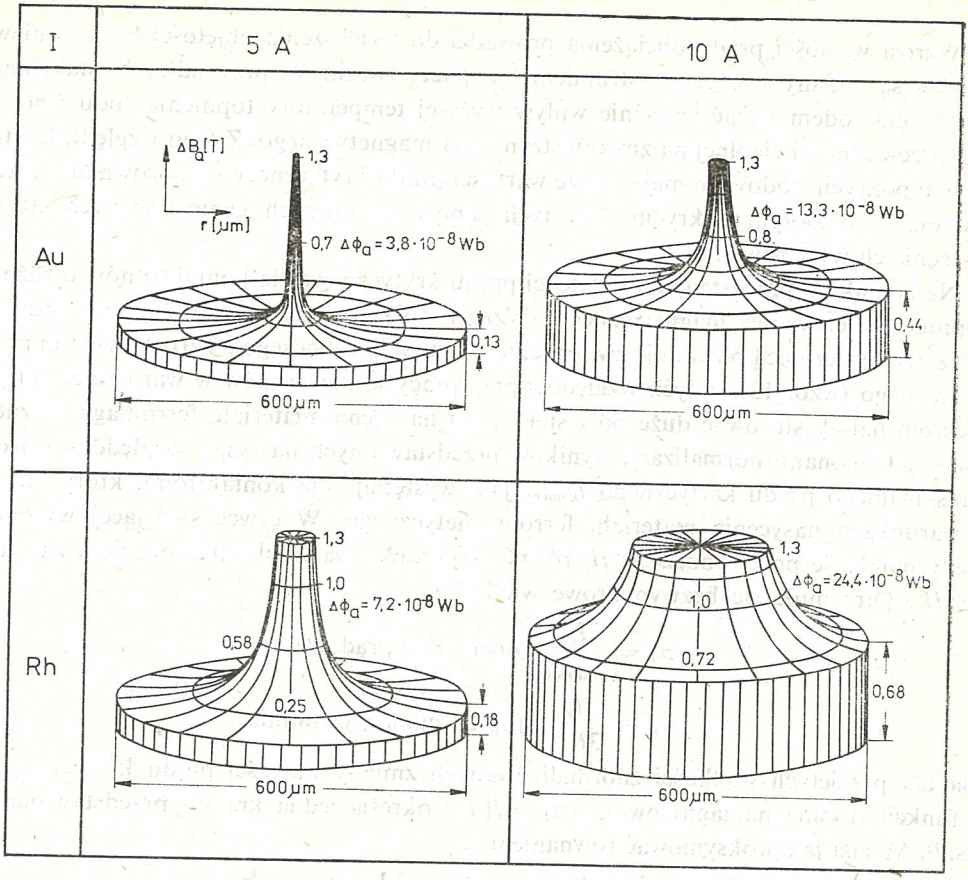
Wykonano obliczenia zmian strumienia roboczego oraz określono wartość prądu krytycznego dla kontaktronów o złotym i rodowym pokryciu stykowym. Styczki kontaktronu wykonane były z materiału ferromagnetycznego FeNiCo (kowar). Na rysunku 6



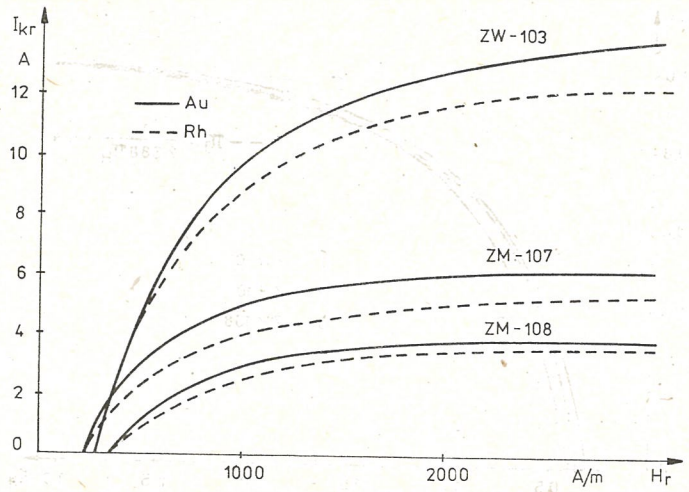
Rys. 6. Zmiany strumienia magnetycznego w kontaktronie przy obciążeniu prądem krytycznym

przedstawiono wyniki obliczeń zmian strumienia magnetycznego, jakie występują przy prądzie krytycznym w kontaktronach typu ZW-103 o różnym stopniu namagnesowania styczek. Przedstawiono całkowite zmiany strumienia $\Delta\Phi_c$ oraz udział poszczególnych składników w tych zmianach. Decydujące znaczenie mają czynniki $\Delta\Phi_f$ i $\Delta\Phi_m$. Czynnik $\Delta\Phi_a$, który reprezentuje zjawiska stykowe ma wpływ na wartość prądu krytycznego i zmiany strumienia roboczego w kontaktronach rodowych (rys. 6b).

Na rysunku 7 przedstawiono przykładowe zmiany wektora indukcji magnetycznej w obszarze stykowym dla kontaktronu o złotym i rodowym pokryciu stykowym. Złoto ma dobrą przewodność cieplną, tworzy mostki metaliczne o małych średnicach. Na niewielkim obszarze przekraczana jest temperatura Curie i następuje całkowity spadek wartości indukcji magnetycznej o wartość 1,3 T.



Rys. 7. Wpływ termicznych zjawisk stykowych na zmiany indukcji magnetycznej w kontaktronach



Rys. 8. Zależność wartości prądu krytycznego od wielkości sterującego pola magnetycznego

Wzrost wartości prądu obciążenia prowadzi do zwiększenia objętości brył, ponieważ większe są zmiany roboczego strumienia magnetycznego. W przypadku kontaktronów o pokryciu rodem widać wyraźnie wpływ wyższej temperatury topnienia rodu i gorszej jego przewodności cieplnej na zmiany strumienia magnetycznego. Z tego względu kontaktrony o pokryciu rodowym mają niższe wartości prądu krytycznego, w porównaniu z kontaktronami o złotym pokryciu przy tych samych wymiarach geometrycznych styczek i warunkach wysterowania.

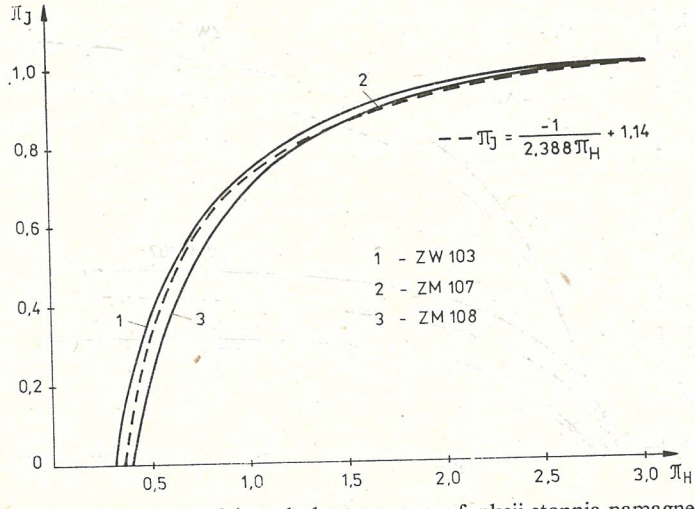
Na rysunku 8 przedstawiono wartości prądu krytycznego dla kontaktronów o różnym stopniu wysterowania magnetycznego. Wzrost pola magnetycznego H_r wytwarzanego przez cewkę sterującą powoduje zwiększenie strumienia roboczego i wzrost wartości prądu krytycznego (wzór 13). Z tych względów przy pracy kontaktronów w warunkach dużych obciążeń należy stosować duże pola sterujące i nasycenie materiału ferromagnetycznego styczek. Dokonano normalizacji wyników przedstawionych na rys. 8 względem wartości maksymalnego prądu krytycznego $I_{k\,max}$, jaki występuje dla kontaktronu, który pracuje w warunkach nasycenia materiału ferromagnetycznego. W cewce sterującej występuje wtedy natężenie pola roboczego H_r równe trzykrotnej wartości natężenia pola zadziałania H_z. Otrzymuje się bezwymiarowe wielkości

$$\pi_I = \frac{I_{kr}}{I_{kr\,max}} \quad \text{obciążenie prądowe,}$$
$$\pi_H = \frac{H_r}{3H_z} \quad \text{stopień namagnesowania.} \tag{14}$$

Dla tak przyjętych wielkości znormalizowanych zmiany wartości prądu krytycznego w w funkcji stopnia namagnesowania $\pi_I = f(\pi_H)$ określa jedna krzywa przedstawiona na rys. 9. Można ją aproksymować równaniem

$$\pi_I = \frac{1}{A_1 \pi_H} + B_1 = \frac{-1}{2{,}388\pi_H} + 1{,}14, \tag{15}$$

gdzie: A_1, B_1 — stałe.



Rys. 9. Znormalizowane zmiany wartości prądu krytycznego w funkcji stopnia namagnesowania materiału ferromagnetycznego styczek

Parametr π_H zmienia się dla wszystkich kontaktronów podobnie i wynika z krzywej namagnesowania materiału ferromagnetycznego. Najbardziej istotny jest parametr $I_{k\,max}$, który określa maksymalne stałoprądowe obciążenie kontaktronu w warunkach nasycenia materiału ferromagnetycznego. W tablicy 2 przedstawiono wartość $I_{k\,max}$ obliczoną przy pomocy wzorów (12) i (13) oraz pomierzoną. Podano wartość średnią otrzymaną z pomiarów dla partii 10 szt. kontaktronów. Widać dobrą zgodność wyników, różnice między wartościami zmierzonymi i obliczonymi nie przekraczają dziesięciu procent w stosunku do wartości obliczonej.

Tablica 2

Wartości maksymalne prądów krytycznych dla kontaktronów

Typ kontaktronu	Materiał pokrycia	$I_{k\,max}$ obliczone [A]	$I_{k\,max}$ pomierzone [A]	%
ZW-103	Au	12,6	13,8	8,7
	Rh	11,6	10,8	7,4
ZM-107	Au	5,95	5,8	2,5
	Rh	5,43	5,26	3,2
ZM-108	Au	4,0	3,84	4,1
	Rh	3,8	3,63	4,6

Wielkość maksymalnego prądu krytycznego $I_{k\,max}$ oraz wzór (15) określają jednoznacznie zakres stałoprądowego obciążenia kontaktronu. Dlatego $I_{k\,max}$ jako parametr charakterystyczny dla danego typu kontaktronu powinien być traktowany jako parametr katalogowy.

6. PODSUMOWANIE

W pracy przeanalizowano przyczyny zmniejszania się wartości roboczego strumienia magnetycznego w kontaktronie przy obciążeniu prądowym. Uwzględniono oddziaływania magnetyczne między obwodem obciążenia a obwodem sterującym kontaktron oraz zjawiska termiczne wywołane przepływem prądu. Zjawisko zmniejszania wartości strumienia magnetycznego ogranicza zakres obciążeń stałoprądowych dla kontaktronów. Stwierdzono, że zakres obciążalności zależy od wymiarów geometrycznych styczek, rodzaju materiału ferromagnetycznego i stykowego a także od stopnia namagnesowania styczek.

LITERATURA

1. T. Hara, H. Tanaka, F. Kume, *Strang Behaviors in High Current Dry Reed Switch*, Proc. 20th Holm Seminar on Electric Contact Phenomena, s. 179—184, Chicago 1974.
2. R. H. Rovnyak, R. Ewell, *Determining High Current Carring Capability of Magnetic Self-Latching*, Proc. of 23th Relay Conf. Stillwater, Oklahoma USA, 1975.
3. B. B. Kosmowski, J. Kaczmarek, H. J. Wierzba, *Influence of High Current Load on Dry Reed Operation*, Proc. of 21th Ann. Holm Seminar on Electric Contact Phenomena, s. 111—114, Chicago 1975.
4. J. Kaczmarek, H. J. Wierzba, *Phenomena Occuring in Dry-Reed at High Current Loads*, Proc. of 8th Int'l Conf. on Electric Contact Phenomena, s. 611—613, Chicago 1978.
5. B. Miedziński, Z. Okraszewski, *Wpływ obciążeń prądowych na pracę przekaźnika kontaktronowego*, PAK, nr 6, s. 214—215, 1977.
6. B. Miedziński, Z. Pytel, *Własności komutacyjne kontaktronów próżniowych*, PAK, nr 10, s. 400—403, 1975.
7. R. H. Bozorth, *Ferromagnetism*, D. van Nostrand, New York 1951.
8. T. Utsumi, *Theoretical and Experimental Investigations of the Dynamic Molten Bridge*, IEEE Trans. on Parts, Material and Packaging PMP-5, No 1, s. 62—68, 1969.
9. J. A. Greenwood, J. B. P. Williamson, *Electrical Conduction in Solids, Theory of Temperature, Dependent Conductors*, Proc. Roy. Soc. A, vol. 246, plate 2, 1958.
10. J. Kaczmarek, *Analiza zjawisk stykowych w kontaktronach przy obciążeniach prądowych*, Praca doktorska, Politechnika Gdańska, 1979, s. 1—111.

J. KACZMAREK

ANALYSIS OF LOAD CURRENT INFLUENCE ON DRY-REED CONTROL MAGNETIC FLUX

Summary

A method of calculating the range of control flux changes caused by load current of dry-reed is presented. The magnetic interaction between load and control circuits and thermic proceses are analysed. The range of dry-reeds maximal current loads is determined.

J. KACZMAREK

ANALYSE DE L'INFLUENCE DU COURANT DE CHARGE SUR LE FLUX MAGNÉTIQUE DE COMMANDE DANS LE RELAIS À LAMES SOUPLES

Résumé

Dans l'article on décrit la méthode de calcul des variations du flux magnétique produit par le courant de charge. On analyse la réaction magnétique entre le circuit de charge et celui de commande ainsi que les effets thermiques. On définit le domaine des charges maximales du courant des relais à lames souples.

J. KACZMAREK

ANALYSE DES LASTSTROMEINFLUSSES AUF DEN STEUERMAGNETFLUß IM SCHUTZROHRKONTAKT

Zusammenfassung

Es wird eine Methode zur Berechnung der durch das Fließen des Laststromes durch den Schutzrohrkontakt hervorgerufenen Änderungen des Magnetflusses beschrieben. Die gegenseitige magnetische Einwirkung des Last- und Steuerkreises sowie die Wärmeerscheinungen werden analysiert. Ferner wird die maximale Strombelastung der Schutzrohrkontakte bestimmt.

Е. КАЧМАРЕК

АНАЛИЗ ВЛИЯНИЯ ТОКА НАГРУЗКИ НА УПРАВЛЯЮЩИЙ МАГНИТНЫЙ ПОТОК ПРОХОДЯЩИЙ ЧЕРЕЗ ГЕРКОН

Р е з ю м е

Изложен метод вычисления изменения магнитного потока в герконе вызванного током нагрузки. Рассмотрены тепловые явления и взаимное влияние цепи нагрузки и упрвляющей цепи. Определена максимальная нагрузка герконов по току.

ROZPRAWY ELEKTROTECHNICZNE 1982, 28, z. 1-2, ss. 261-268

621.38

Zastosowanie metody korekcji macierzy odwrotnej w programach projektowania układów elektronicznych

BOGUSŁAW WOLSZCZAK (WROCŁAW)

Instytut Telekomunikacji i Akustyki Politechniki Wrocławskiej

Otrzymano 26.5.1980

Opisano metodę korekcji macierzy odwrotnej po zmianie wartości elementu w macierzy oryginalnej. Podano odpowiednie zależności na korekcję odwrotnych macierzy sieciowych w analizie układów elektronicznych za pomocą EMC. Przedstawiono przykłady zastosowań tej metody do modyfikcji parametrów elementów sieciowych podczas procesu projektowania oraz zastosowań w procedurach iteracyjnych i w wielkoprzyrostowej analizie wrażliwości funkcji układowej.

1. WSTĘP

W procesie projektowania układów elektronicznych za pomocą EMC można wyodrębnić fazy polegające na powtórzeniu analizy (tzw. projektowanie iteracyjne). Również metody optymalizacyjne w większości przypadków polegają na tworzeniu ciągu kolejnych przybliżeń. Z punktu widzenia programowania w obu przypadkach zachodzą zjawiska iteracyjnego (przyrostowego) zmieniania parametrów elementów gałęziowych analizowanych sieci, co w konsekwencji powoduje zmianę wartości elementów odpowiednich macierzy.

Po każdej takiej iteracji pewne etapy procesu obliczeniowego trzeba powtarzać wielokrotnie od początku. Jeżeli na tych etapach występują macierze odwrotne, to w warunkach podanych poniżej można znacznie zmniejszyć nakłady obliczeniowe stosując metodę korekcji macierzy odwrotnej. Typowe zastosowania tej metody i uwagi dotyczące pisania odpowiedniej procedury programowej przedstawiono w podsumowaniu.

Temat zastosowań formuły korekcji macierzy odwrotnej poruszono w kilku publikacjach [2, 3, 6, 7, 12, 13]. W niniejszej pracy podano szczegółowy opis wykorzystania tej formuły z punktu widzenia wymagań programów projektowania układów elektronicznych, a więc w postaci jednolitej procedury dla wszystkich typów gałęzi sieciowych (tj. sterujących i sterowanych). Wskazano również nową możliwość zastosowania formuły do korekcji macierzy odwrotnej jakobianu przy rozwiązywaniu układów równań nieliniowych w analizie stałoprądowej układów elektronicznych.

2. FORMUŁA KOREKCJI MACIERZY ODWROTNEJ

Podstawą do zastosowań w procedurach korekcji macierzy odwrotnych jest poniższa zależność

$$(M+VSH)^{-1} = M^{-1} - M^{-1}V(S^{-1}+HM^{-1}V)^{-1}HM^{-1}, \tag{1}$$

gdzie:

M, S — nieosobliwe macierze kwadratowe $(n \times n)$,

V, H — macierze prostokątne spełniające warunki mnożenia w zapisie (1).

Poprawność zależności (1) można sprawdzić mnożąc równanie lewostronnie przez $(M+VSH)$.

Zależność (1) jest szczególnie użyteczna gdy macierz $S = s$ jest wielkością skalarną, a macierze V i H są wektorami $v(n \times 1)$ i $h(1 \times n)$. Wówczas przybiera ona postać

$$(M+svh)^{-1} = M^{-1} - \frac{(M^{-1}v)(hM^{-1})}{\frac{1}{s} + hM^{-1}v} \tag{2}$$

spotykaną w literaturze pod nazwą formuły Householdera [3, 6].

Przechodząc do macierzy sieciowych podstawiono w (2) pod macierz M macierz admitancji węzłowych Y, a w miejsce s podstawiono Δy (zmiana admitancji określonej gałęzi). Przyjęto również, że wszystkie elementy macierzy $v(n \times 1)$ i $h(1 \times n)$ są równe zeru, z wyjątkiem niezerowych elementów v_i i h_j, gdzie na współrzędnych i, j są wartości $+1$ lub -1. Po tych podstawieniach zależność (2) ma postać

$$(Y+\Delta yvh)^{-1} = Y^{-1} - \frac{(Y^{-1}v)(hY^{-1})}{\frac{1}{\Delta y} + hY^{-1}v}, \tag{3}$$

lub po podstawieniu $Y^{-1} = Z$, $(Y+\Delta Y)^{-1} = Z'$

$$Z' = Z - \frac{(Zv)(hZ)}{\frac{1}{\Delta y} + hZv}, \tag{4}$$

gdzie Z' — nowa macierz Z po uwzględnieniu zmiany admitancji w elemencie macierzy Y.

Poniżej rozpatrzono 3 warianty wykorzystania zależności (4) w programach analizy układów elektronicznych.

Wariant I. Modyfikacja wartości elementu pomiędzy węzłami $i-j$.

Modyfikacji podlega k-ty element pomiędzy węzłami $i \neq 0$ i $j \neq 0$. Zgodnie z zasadami tworzenia macierzy Y [4, 5] zaburzenie w macierzy Y będzie na współrzędnych:

$$\begin{aligned} Y_{ii} &= Y_{ii} + \Delta y_k, & Y_{ij} &= Y_{ij} - \Delta y_k, \\ Y_{ji} &= Y_{ji} - \Delta y_k, & Y_{jj} &= Y_{jj} + \Delta y_k. \end{aligned} \tag{5}$$

Macierze v i h mają postać:

$$v = \begin{array}{c} \left[\begin{array}{c} 0 \\ 1 \\ 0 \\ \vdots \\ -1 \\ 0 \end{array}\right] \begin{array}{c} \\ i \\ \\ \\ j \\ \\ \end{array} \end{array}, \quad h = [0\overset{i}{1}0\ldots0-\overset{j}{1}0]. \tag{6}$$

Na podstawie (4) otrzymano

$$Z' = Z - \frac{\Delta y_k (Z_{.i} - Z_{.j})(Z_{i.} - Z_{j.})}{1 + \Delta y_k (z_{ii} - z_{ij} - z_{ji} + z_{jj})}, \tag{7}$$

gdzie:

z_{mn} — elementy macierzy $Z = Y^{-1}$,

$Z_{.i}$, $Z_{.j}$ — i-ta i j-ta kolumna macierzy Z,

$Z_{i.}$, $Z_{j.}$ — i-ty i j-ty wiersz macierzy Z.

Zależność (7) obejmuje również przypadek, gdy modyfikacja dotyczy elementu, którego jeden z zacisków połączony jest z masą (podstawą) — wówczas jeden z indeksów jest równy zeru, tzn. $i = 0$ lub $j = 0$. Tak więc w (7) znikają wszystkie elementy opisane indeksem i lub j.

Wariant II. Modyfikacja wartości współczynnika przy źródle sterowanym

Wariant I dotyczy modyfikacji wartości elementu biernego. Wówczas macierze v, h (6) są wzajemnie transponowane, a ich iloczyn jest macierzą symetryczną. W przypadku źródła sterowanego (np. prądowego sterowanego napięciem o transkonduktancji gm) współczynnik gm położony jest w macierzy Y niesymetrycznie względem głównej przekątnej.

Poprawkę Δgm_p, dla współczynnika gm przy p-tym źródle sterowanym, podstawia się w następujący sposób [4]:

$$\begin{aligned} Y_{ki} &= Y_{ki} + \Delta gm_p, \quad & Y_{kj} &= Y_{kj} - \Delta gm_p, \\ Y_{li} &= Y_{li} - \Delta gm_p, \quad & Y_{lj} &= Y_{lj} + \Delta gm_p, \end{aligned} \tag{8}$$

gdzie:

$$\begin{aligned} i &= P[OD(p)], \quad & k &= P[DO(p)], \\ j &= K[OD(p)], \quad & l &= K[DO(p)], \end{aligned} \tag{9}$$

$OD(r)$ — tablica zawierająca numery gałęzi sterujących (sterowanie od gałęzi r),

$DO(s)$ — tablica zawierająca numery gałęzi sterowanych (sterowanie jest skierowane do gałęzi s),

$P(\cdot)$, $K(\cdot)$ — tablica zawierająca numery węzłów wszystkich gałęzi sieci; uporządkowana para $P(n)$, $K(n)$ zawiera numer węzła początkowego gałęzi n i numer jej węzła końcowego (informacja o zastrzałkowanym kierunku prądu w gałęzi).

W tym przypadku macierze v, h mają postać

$$v = \begin{bmatrix} 0 \\ \cdot \\ 1 \\ 0 \\ \cdot \\ -1 \\ 0 \end{bmatrix} \begin{matrix} \\ \\ k \\ \\ \\ l \\ \\ \end{matrix}, \qquad h = [0..\overset{i}{1}..0\overset{j}{-1}0..0]. \tag{10}$$

Na podstawie (4) otrzymano

$$Z' = Z - \frac{\Delta gm_p (Z_{.k} - Z_{.l})(Z_{i.} - Z_{j.})}{1 + \Delta gm_p (z_{ik} - z_{ij} - z_{lk} + z_{jl})}, \tag{11}$$

gdzie:

$Z_{.k}$, $Z_{.l}$ — k-ta i l-ta kolumna macierzy $Z = Y^{-1}$,

$Z_{i.}$, $Z_{j.}$ — i-ty i j-ty wiersz macierzy Z.

Wariant III. Korekcja macierzy odwrotnej jakobianu

W procesach iteracyjnych (np. rozwiązywanie układów równań nieliniowych) i optymalizacyjnych [1, 6] bardzo często występuje operacja inwersji macierzy Jakobiego. Jeśli zaburzeniu podlega jednocześnie większość elementów macierzy, to trzeba wykonać ponownie inwersję lub obliczyć przybliżenie tej macierzy odpowiednią metodą (np. Broydena [1, 6]).

Praktyka wykazuje, że bardzo często w każdym kroku iteracyjnym zaburzeniu podlega tylko niewielka część elementów macierzy, więc stosowanie formuły korekcyjnej (2) jest bardziej ekonomiczne. Poniżej, przykładowo, przedstawiono etap rozwiązywania układu równań nieliniowych w analizie stałoprądowej układów elektronicznych metodą iteracyjną opartą na koncepcji Newtona. Punktem wyjścia jest utworzenie funkcji błędu $\varphi(U)$ [8, 11]

$$\varphi(U) = I(U) - AU - B, \tag{12}$$

gdzie:

U — wektor napięć na elementach nieliniowych,

$I(U)$ — wektor funkcji opisujących prądy elementów nieliniowych w funkcji napięć U,

A, B — macierze współczynników (opisujące tzw. liniową część sieci).

Poniżej podano jakobian funkcji błędu $\varphi(U)$ w k-tej iteracji

$$J^{(k)}(U) = \frac{\partial \varphi(U)}{\partial U} = -A + \frac{\partial I(U)}{\partial U}\bigg|_{U=U^{(k)}}. \tag{13}$$

Jeżeli funkcje $I(U)$ będą opisywały złącza półprzewodnikowe, to jakobian przyjmie postać

$$\frac{\partial I(U)}{\partial U} = \begin{bmatrix} \dfrac{\partial I_1(u_1)}{\partial u_1} & & & 0 \\ & \dfrac{\partial I_2(u_2)}{\partial u_2} & & \\ & & \ddots & \\ 0 & & & \dfrac{\partial I_n(u_n)}{\partial u_n} \end{bmatrix} = \begin{bmatrix} \Delta_1 & & & 0 \\ & \Delta_2 & & \\ & & \ddots & \\ 0 & & & \Delta_n \end{bmatrix}, \tag{14}$$

a więc w każdej (k-tej) iteracji zmiany będą tylko na głównej przekątnej jakobianu (13)

$$J = \begin{bmatrix} a_{11}+\Delta_1 & a_{12} & \cdots & a_{1i} & \cdots & a_{1n} \\ a_{21} & a_{22}+\Delta_2 & & & & \vdots \\ \vdots & & \ddots & & & \\ & & & a_{ii}+\Delta_i & & \\ \vdots & & & & \ddots & \\ a_{n1} & \cdots & \cdots & a_{ni} & \cdots & a_{nn}+\Delta_n \end{bmatrix} \tag{15}$$

(dla uproszczenia zapisu założono, że znak „$-$" przed A jest przeniesiony na wartości a_{ij}).

Obecna sytuacja przypomina wariant I dla uziemionego jednego zacisku elementu sieciowego, ponieważ modyfikacja tego typu wprowadza zaburzenie na głównej przekątnej macierzy. Wprowadzając do (2) nowe zmienne, otrzymano

$$J' = J - \frac{\Delta_i J_{.i} J_{i.}}{1+\Delta_i a_{ii}}, \tag{16}$$

gdzie:

$J_{.i}$ — i-ta kolumna macierzy J,

$J_{i.}$ — i-ty wiersz macierzy J,

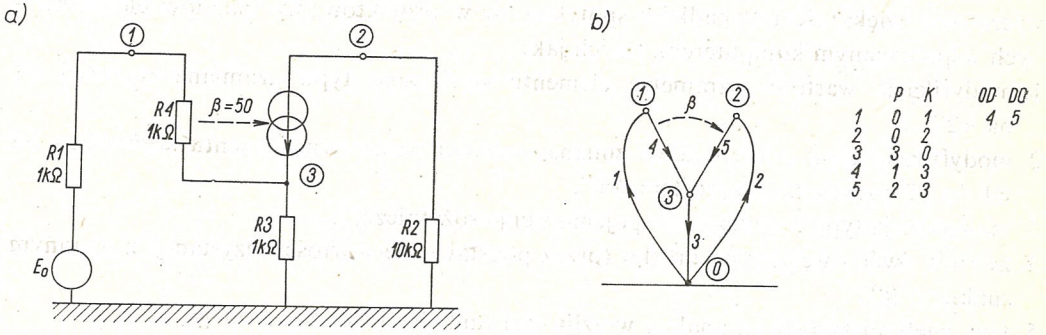
a_{ii} — element jakobianu

(uwaga — w (16) macierz J jest już po pierwotnej inwersji).

Operację (16) trzeba wykonać n razy dla $i = 1, 2, \ldots, n$, biorąc do kolejnej (następnej) korekcji stan macierzy J po poprzedniej korekcji.

3. PRZYKŁAD

Poniżej rozważono przypadek zmiany wartości rezystancji rezystora $R4$ będącego jednocześnie gałęzią sterującą — rys. 1.



Rys. 1. Przykład modyfikacji wartości parametru gałęzi sterującej $R4$.
a) idealizowany schemat stopnia wzmacniającego, b) graf układu i tablice z zapamiętaną topologią

Macierz admitancji węzłowych sieci ma następującą postać

$$Y = \begin{bmatrix} G4+G1 & 0 & -G4 \\ gm & G2 & -gm \\ -gm-G4 & 0 & gm+G3+G4 \end{bmatrix} = \begin{bmatrix} 2 & 0 & -1 \\ 50 & .1 & -50 \\ -51 & 0 & 52 \end{bmatrix} [\text{mS}]. \tag{17}$$

Po inwesji otrzymano

$$Z = Y^{-1} = \begin{bmatrix} .98113 & 0.0 & .01887 \\ -9.43396 & 10.0 & 9.43396 \\ .96226 & 0.0 & .03774 \end{bmatrix} [\text{k}\Omega]. \tag{18}$$

Przy warunku $\beta = 50 = \text{const}$ zwiększono konduktancję $G4$ o 10%, czyli $\Delta g4 = \cdot 1$ mS, a $\Delta gm = \beta \Delta g4 = 50*.1 = 5$ mS. Korekcję wykonano w dwóch etapach:

1) zastosowano formułę (7) do korekcji elementu biernego $G4$ i uzyskano

$$Z^{(1)} = (Y + \Delta g4vh)^{-1} = \begin{bmatrix} .97932 & 0.0 & .02068 \\ -9.39850 & 10.0 & 9.39850 \\ .96053 & 0.0 & .03947 \end{bmatrix} [\text{k}\Omega), \tag{19}$$

2) do uzyskanej w ten sposób macierzy $Z^{(1)}$ zastosowano wzór (11) do korekcji transkonduktancji gm o wartość Δgm i uzyskano wynik końcowy

$$Z^{(2)} = Z' = \begin{bmatrix} .98110 & 0.0 & .01890 \\ -9.45017 & 10.0 & 9.45017 \\ .96392 & 0.0 & .03608 \end{bmatrix} [\text{k}\Omega]. \tag{20}$$

Macierz wynikową Z' (20) uzyskano stosując dwukrotnie korekcję macierzy odwrotnej Z (18) — najpierw dla elementu biernego $G4$, a następnie dla współczynnika gm przy źródle sterowanym.

4. PODSUMOWANIE

Przedstawione 3 warianty wykorzystywania procedury korekcji macierzy odwrotnej wyczerpują większość przypadków spotykanych w projektowaniu układów elektronicznych wspomaganym komputerem, takich jak:

1. modyfikacja wartości parametru elementu sieciowego typu „zamiana wartości $w1$ na $w2$",
2. modyfikacja typu „przyrostowa zmiana wartości parametru elementu sieciowego od $w1$ do $w2$ z zadanym przyrostem Δw",
3. modyfikacja typu „przyłączenie pojemności pasożytniczej",
4. zwarcie dwóch węzłów na krótko (przez podstawienie wartości rezystancji z ujemnym znakiem) [8],
5. tzw. wielkoprzyrostowa analiza wrażliwości funkcji układowej sieci [2, 3, 7],
6. zastosowanie w procesach iteracyjnych [11] i optymalizacyjnych [1, 6].

Z punktu widzenia programowania procedury korekcji macierzy trzeba zwrócić uwagę na dwa aspekty: a) minimalizację zajętości pamięci operacyjnej, b) spełnienie warunku obliczania wszystkich wariantów przez tę samą procedurę.

Ad a). Analiza formuły (2) prowadzi do wniosku, że od macierzy M^{-1} odejmuje się pewną macierz poprawek ΔM. Zamiast tworzenia dodatkowej, pomocniczej macierzy $\Delta M(n \times n)$ trzeba poprawki nanosić na macierz M^{-1} używając do zapamiętania zmiennych pomocniczych (np. autor zastosował dwie tablice $X1(n)$, $X2(n)$ do pamiętania odpowiedniego wiersza i kolumny macierzy M^{-1} przed korekcją).

Ad b). Stosowanie procedury można ujednolicić zauważając, że poprawki mogą być nanoszone cyklicznie na tę samą macierz M^{-1} (wykorzystując do następnej korekcji macierz M^{-1} już poprzednio skorygowaną). Proponuje się, aby procedura korekcji była z następującymi parametrami wejściowymi $(i, j, \Delta W)$, gdzie i, j — współrzędne modyfikowanego elementu w macierzy M^{-1}, ΔW — przyrost wartości parametru. Procedura ta może być wywoływana jednokrotnie z parametrami $i, i, \Delta y$ w przypadku modyfikacji

elementu z uziemionym zaciskiem, lub czterokrotnie z parametrami $(i, i, \Delta y)$, $(i, j, -\Delta y)$, $(j, i, -\Delta y)$, $(j, j, \Delta y)$ według (5) w przypadku modyfikacji przedstawionej w wariancie I. Ostatnie spostrzeżenie czyni wykorzystanie procedury korekcji bardzo elastycznym i pozwala traktować wszystkie zdarzenia programowe w sposób jednolity.

W uzupełnieniu trzeba wspomnieć o dogodności stosowania procedury korekcji macierzy odwrotnej w analizie zmian transmitancji i impedancji zaciskowej sieci na wielkoprzyrostowe zmiany parametrów elementów gałęziowych. W pracach [8, 9] podano odpowiednie zależności wyrażone za pomocą elementów macierzy odwrotnej.

LITERATURA

1. D. A. Calahan, *Projektowanie układów elektronicznych za pomocą maszyny cyfrowej*, WNT, Warszawa 1978.
2. J. K. Fidler, *Calculation of network sensitivity and its application to network analysis*, Proc. IEE, v. 118, nr 11, Nov. 1971.
3. J. K. Fidler, *Some topics in network sensitivity*, Prace Nauk. Inst. Telekom. i Akustyki, nr 32, seria Monografie, nr 14, Politechnika Wrocławska, Wrocław 1977.
4. M. Grobelny, *Zarys liniowej teorii układów elektronicznych*, WKŁ, Warszawa 1976.
5. M. Grobelny, *Projektowanie układów elektronicznych za pomocą komputerów*, WKŁ, Warszawa 1973.
6. F. F. Kuo, W. G. Magnuson, *Computer Oriented Circuit Design*, Prentice-Hall, Inc. Englewood Cliffs, N. J. 1969.
7. K. H. Leung, R. Spence, *Multiparameter Large-Change Sensitivity Analysis and Systematic Exploration*, IEEE Trans., v. CAS-22, nr 10, Oct. 1975, ss. 796—804.
8. B. Wolszczak, *Analiza pewnej klasy układów elektronicznych prądu stałego jako ogniwo procesu projektowania wspomaganego komputerem*, praca doktorska, Inst. Telekom. i Akustyki Politechniki Wrocławskiej, komunikat I-28/K/D/31/74, Wrocław 1974.
9. B. Wolszczak, *Analiza czułościowa układów elektronicznych za pomocą EMC*, Archiwum Elektroniki, t. XXIII, z. 1, 1974, ss. 191—199.
10. B. Wolszczak, *Modelowanie układów elektronicznych na EMC w dziedzinie prądu stałego*, Rozprawy Elektrotechniczne, 1977, v. 23, z. 4, ss. 927—940.
11. B. Wolszczak, *Efektywność zmodyfikowanych metod iteracyjnych Newtona w analizie stanu ustalonego układów elektronicznych*, Rozprawy Elektrotechniczne, 1978, v. 24, z. 2, ss. 253—264.
12. P. J. Goddard, P. A. Villalaz, R. Spence, *Method for the efficient computation of the large-change sensitivity of linear nonreciprocal networks*, Electronic Letters, v. 7, nr 4, 1971, ss. 112—113.
13. G. C. Temes, *Multiple fault simulation in linear active circuits*, Electronic Letters, v. 12, nr 18, 1976, ss. 467—468.

B. WOLSZCZAK

APPLICATION OF INVERSE MATRIX CORRECTION METHOD IN ELECTRONIC CIRCUIT DESIGN PROGRAMS

Summary

In this paper an inverse matrix correction method after a change in elements of the original matrix is described. Suitable formulas to correct the inverse for computer-aided analysis of electronic circuits are given. Some examples of application of this method for network element parameters modifications during design process are presented. Attention is given to application of this method in iterative procedures and large-change sensitivity analysis of the network function.

B. WOLSZCZAK

APPLICATION DE LA MÉTHODE DE CORRECTION DE LA MATRICE INVERSE DANS LES PROGRAMMES DES PROJETS DES CIRCUITS ÉLECTRONIQUES

Résumé

On a décrit la méthode de correction de la matrice inverse après le changement de la valeur d'un élément de la matrice originale. On a donné les relations convernables pour la correction des circuits électriques à l'aide d'un calculateur. On a présenté des exemples de l'application de cette méthode à la modification des paramètres des éléments du circuit au cours du processus d'établissement du projet. On a présenté également son application dans les procédures itéraires et dans l'analyse d'un grand accroissement de sensibilité de la fonction du circuit.

B. WOLSZCZAK

ANWENDUNG DER KORREKTIONSMETHODE EINER UMKEHRMATRIX FÜR PROJEKTIONSPROGRAMME ELEKTRONISCHER SYSTEME

Zusammenfassung

Es wurde die Korrektionsmethode einer Umkehrmatrix nach Wertveränderung der Elemente der Originalmatrix beschrieben. Angegeben wurden entsprechende Abhängigkeiten der Korrektion einer Umkehrmatrix des Netzes bei der Analyse elektronischer Systeme mit Hilfe eines Digitalrechners. Im weiteren werden Anwendungsbeispiele für die Modifizierung der Parameter der Netzelemente während des Projektierungsprozesses sowie Anwendungsbeispiele für die Iterationsprozedur und für die Analyse einer großen Empfindlichkeitsänderung der Systemfunktion dargestellt.

Б. ВОЛЬЩАК

ПРИМЕНЕНИЕ МЕТОДА КОРРЕКЦИИ ОБРАТНОЙ МАТРИЦЫ В ПРОГРАММАХ АНАЛИЗА ЭЛЕКТРОННЫХ СХЕМ

Резюме

Описан метод коррекции обратной матрицы поспе изменения величины элемента в исходной матрице. Предложены соответствующие выражения для коррекции обратных матриц электронных схем в анализе с помощю ЭВМ. Предложены примеры использования этих методов для преобразования параметров электронных схем в процессе проектирования, использования в итерационных процедурах и в анализе чувствительности при больших приращениях.

ROZPRAWY ELEKTROTECHNICZNE 1982, 28, z. 1-2, ss. 269-283

621.396.67

Metody pomiaru gęstości mocy w polu bliskim

HUBERT TRZASKA (WROCŁAW)

Instytut Telekomunikacji i Akustyki Politechniki Wrocławskiej

Otrzymano 6.7.1978

W pracy przedstawiono metody pomiaru gęstości mocy w polu bliskim z punktu widzenia dokładności pomiaru. Przedyskutowano metody oparte na:
1. pomiarze gęstości mocy elektrycznej lub magnetycznej,
2. wykorzystaniu dwustronnie obciążonej anteny ramowej,
3. średniej arytmetycznej gęstości mocy elektrycznej i magnetycznej,
4. średniej geometrycznej gęstości mocy elektrycznej i magnetycznej.

Pokazano, że przy częstotliwościach większych od 300 MHz najkorzystniejsze jest stosowanie metody (1). Natomiast przy częstotliwościach mniejszych od 300 MHz najmniejsze błędy pomiarowe wskazują na potrzebę stosowania metody (4). Wynikiem pomiaru tą metodą jest moduł wektora Poyntinga.

1. WSTĘP

Nasycenie naturalnego środowiska elektromagnetycznego sztucznie wytwarzaną energią elektromagnetyczną systematycznie wzrasta, przy czym w większości przypadków zahamowanie tego wzrostu byłoby równoznaczne z zahamowaniem ogólnie rozumianego postępu cywilizacji. Aby jednak wzrost ten nie doprowadził do niepożądanych oddziaływań na organizmy żywe, a w szczególności do niepożądanych zmian genetycznych, fizjologicznych, czy nawet psychicznych u człowieka, w większości krajów wysoko rozwiniętych wprowadzono ograniczenia dopuszczalnych wartości natężeń pól elektromagnetycznych (*PEM*), dopuszczalnych dla człowieka bez ograniczeń lub w ograniczonym czasie napromieniowania. Wartości dopuszczalnych natężeń *PEM* zostały przyjęte po wieloletnich badaniach prowadzonych głównie przez biologów i lekarzy nad działaniem biologicznym *PEM* z różnych przedziałów częstotliwości. Wskutek przyjęcia różnych kryteriów oceny szkodliwości biologicznej *PEM* w różnych krajach przyjęto różne wartości natężeń *PEM* jako dopuszczalne, a więc i nieszkodliwe dla organizmu człowieka. Poglądy w tej mierze zmieniają się w miarę postępu techniki badań biologicznych, jednak do chwili obecnej, poza przeważającymi poglądami opierającymi się na aktualnym stanie wiedzy, występują także poglądy skrajne i to zarówno negujące szkodliwość *PEM*, jak i przypisujące prawodawcom zbytnią pobłażliwość dla interesów (głównie finansowych) instytucji odpowiedzialnych za wytwarzanie *PEM*.

W większości obowiązujących przepisów ochrony przed promieniowaniem *EM* można by się dopatrzeć wielu nieścisłości i niedopracowań, jednak najważniejszy jest tu fakt ich istnienia i działania w kierunku ochrony najwyższego dobra, jakim jest zdrowie i życie ludzkie.

Wśród badaczy zajmujących się biologicznym działaniem *PEM* nie ma do chwili obecnej ustalonego poglądu na wielkość charakteryzującą *PEM*, a raczej zarówno *PEM* jak i obiekt biologiczny, która byłaby najbardziej charakterystyczna dla ich wzajemnej interakcji [8]. Rezultatem tego są głosy domagające się pomiaru mocy w jednostce objętości [W/m^3], mocy wydzielającej się w jednostce masy [W/kg] lub też energii absorbowanej w jednostce masy [J/kg]. Pomiar tych wielkości jest realizowany przez pomiar przyrostu temperatury, co znacznie utrudnia badania in vivo; natomiast badania in vitro nie są w pełni reprezentatywne dla oddziaływania *PEM* z organizmami żywymi.

Pomiar absorbcji energii *PEM* drogą pomiaru przyrostu temperatury fantomu jest najdokładniejszą miarą oddziaływania *PEM* na organizm żywy, a także najdokładniejszą metodą pomiaru gęstości energii lub gęstości mocy *PEM* w dowolnych warunkach. Przyjmując najkorzystniejszy model absorbcji energii *PEM* można stwierdzić, że przyrosty temperatury odpowiadające gęstościom mocy jakie powinny być mierzone [16] leżą znacznie poniżej możliwości pomiaru. Dlatego też powszechnie stosuje się metody pomiaru gęstości mocy opierające się na wykorzystaniu anten „krótkich".

Wszystkie przepisy ochronne w zakresie częstotliwości powyżej 300 MHz (a niekiedy także powyżej 10 MHz) dopuszczają określone wartości gęstości mocy *PEM*. Przy częstotliwościach niższych określa się zwykle wartości natężenia pola elektrycznego, a czasem także i magnetycznego.

Ze względu na to, że pomiary *PEM*, z punktu widzenia ochrony zdrowia, są prowadzone zwykle w polu bliskim pierwotnych lub wtórnych źródeł promieniowania, konieczne jest stosowanie specyficznej metodyki pomiarów, jak też i przyrządów przeznaczonych do pomiarów w polu bliskim. O ile zagadnienia pomiaru natężeń pól elektrycznych i magnetycznych w.cz. zostały dość szczegółowo opracowane [1, 2, 3, 5, 9, 11] i w chwili obecnej konstrukcja przyrządów pomiarowych stanowi tu raczej problem natury inżynierskiej lub technologicznej, o tyle stosunkowo niewiele prac zostało poświęconych problemowi pomiaru gęstości mocy w polu bliskim [4, 7, 10], a większość stosowanych metod opiera się na pewnych przybliżeniach. Celowi dyskusji tych przybliżeń, a więc i dokładności pomiaru a także przedstawieniu stosowanych metod ma służyć niniejsza praca.

2. GĘSTOŚĆ MOCY W POLU BLISKIM

Znając natężenie pola elektrycznego **E** i natężenie pola magnetycznego **H** gęstość mocy można w sposób jednoznaczny (co do wielkości i kierunku) wyznaczyć w oparciu o wektor Poyntinga **P** [6]:

$$\mathbf{P} = \mathbf{E} \times \mathbf{H}. \tag{1}$$

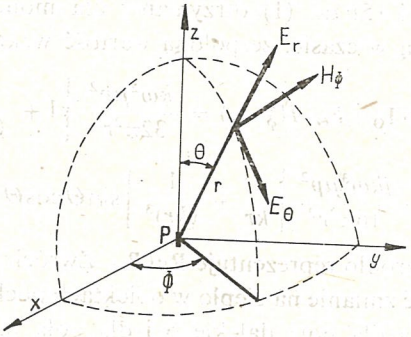
Biorąc pod uwagę zależność

$$\mathbf{E} = Z(\mathbf{H} \times \mathbf{1}_r) \tag{2}$$

gdzie:

$\mathbf{1}_r$ — wersor kulistego układu współrzędnych (rys. 1),

Z — impedancja falowa ośrodka,



Rys. 1. Elementarny dipol elektryczny w układzie współrzędnych

i podstawiając (2) do (1), możemy wyrazić wektor Poyntinga w oparciu o **E** lub **H** oraz Z. Jednak w polu bliskim wzajemny stosunek **E** do **H** jest zależny od struktury źródła promieniowania i od odległości do źródła promieniowania. Tak więc, kiedy błąd pomiaru gęstości mocy w oparciu o pomiar natężenia pola elektrycznego lub magnetycznego jest zbyt duży, konieczny jest niezależny pomiar **E** i **H** i określenie gęstości mocy w oparciu o wzór (1).

Pewnego wyjaśnienia wymaga tu problem rodzaju stosowanej anteny pomiarowej. Nie są znane anteny reagujące sensu stricto na gęstość mocy, znane są natomiast anteny reagujące równocześnie na **E** i **H**. Jednakże anteny takie mają wymiary porównywalne z długością fali, a także pewne własności kierunkowe; czyni je to całkowicie nieprzydatnymi do pomiarów w polu bliskim. Do pomiarów w polu bliskim mogą być stosowane jedynie anteny małe w porównaniu z długością fali, reagujące tylko na **E** lub tylko na **H** (wyjątek od tej reguły opisano w p. 4) [10].

Rozpatrzymy obecnie gęstość mocy w polu bliskim elementarnego dipola elektrycznego, którego amplitudy składowych są wyrażone wzorami [13]:

$$E_r = \frac{p}{2\pi\varepsilon}\left(\frac{1}{r^3} - \frac{jk}{r^2}\right)\cos\Theta\exp(-jkr), \tag{3}$$

$$E_\Theta = \frac{p}{4\pi\varepsilon}\left(\frac{1}{r^3} - \frac{jk}{r^2} - \frac{k^2}{r}\right)\sin\Theta\exp(-jkr), \tag{4}$$

$$H_\phi = \frac{-j\omega p}{4\pi}\left(\frac{1}{r^2} - \frac{jk}{r}\right)\sin\Theta\exp(-jkr) \tag{5}$$

oraz $E_\phi = H_r = H_\Theta = 0$,

gdzie:

p — moment dipolowy

$$p = |\boldsymbol{p}| = \left|\frac{Idz}{j\omega}\mathbf{1}_z\right|,$$

I — prąd w dipolu,

ε, μ — przenikalność elektryczna i magnetyczna ośrodka, $k = \frac{2\pi}{\lambda}$,

inne oznaczenia jak na rys. 1.

Podstawiając (3), (4) i (5) do (1) otrzymamy dla monochromatycznego przebiegu harmonicznego uśrednioną w czasie, zespoloną wartość wektora Poyntinga $\mathbf{P}_s$:

$$\mathbf{P}_s = \frac{1}{2}(E_r \cdot H_\phi^* \cdot \mathbf{1}_\Theta + E_\Theta \cdot H_\phi^* \cdot \mathbf{1}_r) = \frac{k\omega^2\mu p^2}{32\pi^2 r^2}\left[1 + \frac{1}{(kr)^3}\right]\sin^2\Theta \cdot \mathbf{1}_\Theta -$$

$$- \frac{jk\omega^3\mu p^2}{16\pi^2 r^2}\left[\frac{1}{kr} + \frac{1}{(kr)^3}\right]\sin\Theta\cos\Theta \cdot \mathbf{1}_r. \tag{6}$$

Moc promieniową przez źródło reprezentuje $\mathrm{Re}(P_s)$. Zwrócimy tu uwagę na fakt, że tylko ta część mocy może ulegać zmianie na ciepło w obiektach pochłaniających; część ta wyraża się identyczną zależnością dla pola dalekiego i dla pola bliskiego. Z punktu widzenia efektów termicznych oddziaływania *PEM* na organizmy żywe interesuje nas więc pomiar gęstości mocy promieniowanej przez źródło P_r, której moduł wynosi

$$P_r = \mathrm{Re}\,(P_s) = \frac{k\omega^3\mu p^2}{32\pi^2 r^2}. \tag{7}$$

Wielkość określoną wzorem (7) będziemy w dalszej części pracy przyjmować jako poziom odniesienia dla prezentowanych tu metod pomiaru gęstości mocy.

3. POMIAR GĘSTOŚCI MOCY PRZEZ POMIAR E LUB H

Słuszność koncepcji pomiaru gęstości mocy w polu dalekim w oparciu o pomiar **E** lub **H** nie budzi żadnych zastrzeżeń [7]. Natomiast w polu bliskim może występować „dominacja" jednej ze składowych pola i wskutek tego gęstość mocy określona na podstawie pomiaru „dominującej" składowej *PEM* będzie zawyżona. Biorąc pod uwagę składowe elementarnego dipola elektrycznego łatwo zauważyć, że (zwłaszcza dla $\Theta \to 0°$) słuszna jest zależność: $\lim_{r\to 0} |\mathbf{E}/\mathbf{H}| \to \infty$.

Dla wyznaczenia granicy stosowalności metody pomiaru gęstości mocy przez pomiar **E** lub **H** obliczymy na podstawie (3) i (4) wartość gęstości mocy P_E określoną w oparciu o pomiar pola **E** oraz wartość gęstości mocy P_H określoną na podstawie pomiaru pola **H** dla źródła w postaci elementarnego dipola elektrycznego. Ze względu na warunki wzorcowania przyjmiemy $Z = Z_o$ oraz

$$P_r = P_E = P_H. \tag{8}$$

Wtedy:

$$P_E = \frac{|\mathbf{E}|^2}{2Z_o} = \frac{\omega k^3 p^2}{8\pi^2 \varepsilon r^2}\left\{\frac{\sin^2\Theta}{4}\left[1 - \frac{1}{(kr)^2} + \frac{1}{(kr)^4}\right] + \cos^2\Theta\left[\frac{1}{(kr)^2} + \frac{1}{(kr)^4}\right]\right\}, \tag{9}$$

gdzie: $Z_o = \sqrt{\mu_o/\varepsilon_o}$ — impedancja falowa swobodnej przestrzeni, oraz

$$P_H = \frac{|\mathbf{H}|^2 Z_o}{2} = \frac{\omega k^3 p^2}{32\pi^2 \varepsilon r^2}\left[1 + \frac{1}{(kr)^2}\right]\sin^2\Theta. \tag{10}$$

Porównując gęstości mocy określone wzorami (9) i (10) z (7) możemy określić błędy pomiaru gęstości mocy promieniowanej przez elektryczny dipol elementarny przez pomiar pola elektrycznego η_{EE} i przez pomiar pola magnetycznego η_{EH}

$$\eta_{EE} = \frac{|P_r - P_E|}{P_r} \tag{11}$$

oraz

$$\eta_{EH} = \frac{|P_r - P_H|}{P_r}. \tag{12}$$

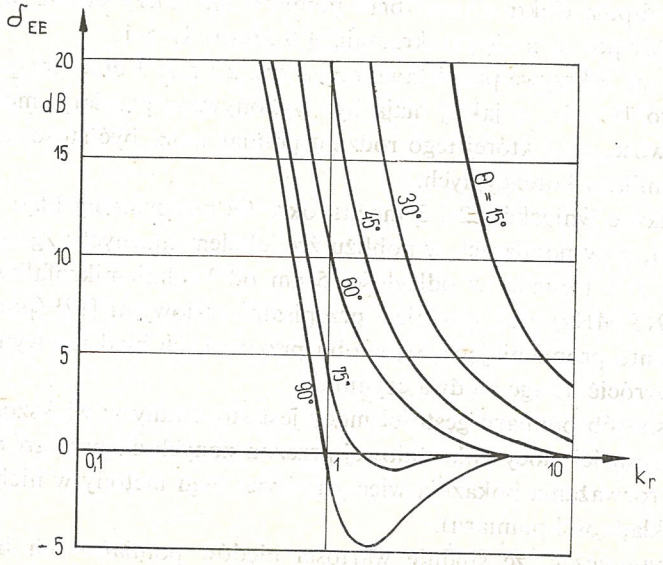
Wzory (11) i (12) są przydatne do liczenia błędów pomiaru, gdy wartość błędu nie przekracza około 15%, dla błędów większych wartości błędów η_{EE} i η_{EH} stają się trudne do porównania. Tymczasem pomiary natężenia pola i gęstości mocy są jednymi z najmniej dokładnych pomiarów wielkości fizycznych i niekiedy przyjmuje się jako zadowalające dokładności ± 3 dB, a nawet ± 6 dB. Dlatego też na rys. 2 i rys. 3 pokazano błędy δ_{EE} i δ_{EH} sformułowane jako:

$$\delta_{EE} = 20\log\frac{|P_E|}{P_r}, \tag{13}$$

$$\delta_{EH} = 20\log\frac{|P_H|}{P_r}. \tag{14}$$

Prowadząc analogiczne rozważania dla źródła w postaci elementarnego dipola magnetycznego otrzymamy błędy pomiaru gęstości mocy w otoczeniu tego źródła δ_{HE} i δ_{HH} wynikające z oparcia pomiaru gęstości mocy o pomiar **E** i **H** odpowiednio. Łatwo zauważyć, że są spełnione następujące zależności:

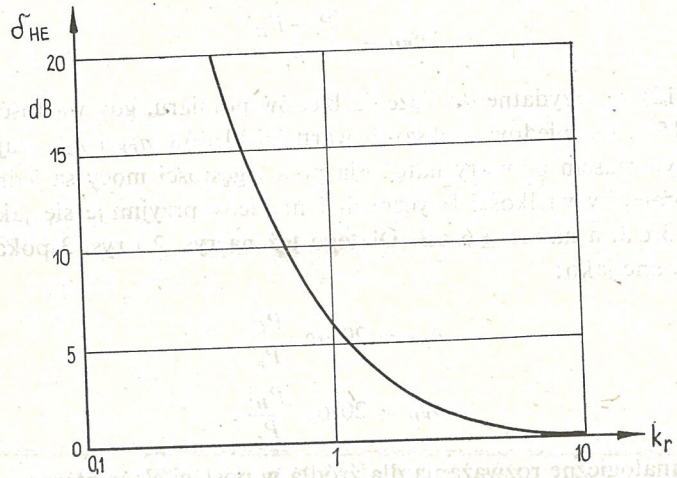
$$\begin{aligned} \delta_{EE} &= \delta_{HH}, \\ \delta_{EH} &= \delta_{HE}. \end{aligned} \tag{15}$$



Rys. 2. Przebieg błędu δ_{EE} w funkcji kr

Przeprowadzone rozważania pozwalają na wyciągnięcie kilku wniosków.

1. Pomiar gęstości mocy w polu bliskim na podstawie pomiaru natężenia pola elektrycznego lub magnetycznego jest obarczony błędem, którego wielkość zależy od rodzaju źródła i od mierzonej składowej *PEM*, a także od odległości od źródła.



Rys. 3. Przebieg błędu δ_{EH} w funkcji kr

2. Dla źródła w postaci elementarnego dipola elektrycznego lub magnetycznego błąd pomiaru gęstości mocy przez pomiar **E** lub **H** odpowiednio jest zawsze duży dla $\Theta = 0°$ i maleje dla $\Theta = 90°$. Przy pomiarze pola **E** w otoczeniu dipola magnetycznego i pola **H** w otoczeniu dipola elektrycznego błąd pomiaru nie zależy od Θ. Można przyjąć, że dla $\Theta \approx 90^0$ błędy pomiaru nie przekraczają ± 6 dB dla $kr > 1$.

3. Opierając się o krzywe przedstawione na rys. 2 i rys. 3 oraz przyjmując minimalną odległość od źródła pola, w jakiej mają być wykonywane pomiary, można określić najniższą częstotliwość, przy której tego rodzaju pomiar może być stosowany bez dodatkowych współczynników korekcyjnych.

4. Na podstawie wniosków 2 i 3 można określić maksymalny błąd pomiaru gęstości mocy tą metodą, przy pomiarach w pobliżu źródeł elementarnych; zgodnie z zaleceniami amerykańskimi [14] (pomiar w odległości 5 cm od kuchni mikrofalowej pracującej na częstotliwości 915 MHz) oraz zgodnie z przepisami krajowymi [17] (pomiar w odległości 15 cm od elementu promieniującego); w obu przypadkach błąd ten wyniesie około 6 dB.

Należy tu zwrócić uwagę na dwa czynniki:

— omawiany sposób pomiaru gęstości mocy jest stosowany w większości wytwarzanych mierników gęstości mocy mikrofalowej przeznaczonych do pomiarów w polu bliskim; przedstawione rozważania pokazują więc graniczny błąd metody w nich stosowanej (niezależnie od dokładności pomiaru),

— można przypuszczać, że średnie wartości błędów pomiarowych będą mniejsze, co wynika ze skończonych wymiarów źródeł fizycznych.

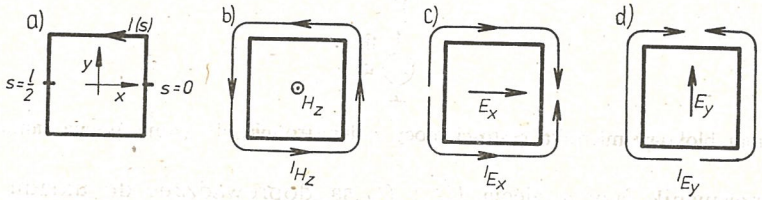
4. POMIAR GĘSTOŚCI MOCY PRZY WYKORZYSTANIU „EFEKTU ANTENOWEGO"

Pojęcie „efekt antenowy" określa wrażliwość anteny ramowej także na działanie pola elektrycznego. Jak już wspomniano, do pomiarów w polu bliskim konieczne jest stosowanie anten o wymiarach znacznie mniejszych od długości fali, które są wrażliwe bądź na **E** bądź na **H**; jednak jeżeli np. średnica kołowej anteny ramowej D nie spełnia zależności [9]

$$D \leqslant 0{,}01\,\lambda_{min}, \tag{16}$$

gdzie λ_{min} — najkrótsza długość fali przy jakiej antena ma pracować, to *SEM* indukowana w antenie przez **E** nie jest pomijalnie mała w stosunku do *SEM* indukowanej przez **H** i antena taka jest nieprzydatna do pomiarów. Wzór (16) dotyczy anteny ramowej nieekrawanej i obciążonej jednostronnie.

Szczególnie ciekawy sposób wykorzystania „efektu antenowego" został przedstawiony w [15]. Sposób ten został zastosowany w przyrządzie przeznaczonym do pomiaru gęstości mocy w polu bliskim [10]. Istotę tego sposobu przedstawimy w oparciu o rys. 4. Kwadra-



Rys. 4. Prądy w dwustronnie obciążonej antenie ramowej

towa antena ramowa o obwodzie l, wykonana z przewodu o średnicy $2a$, leży w płaszczyźnie xy kartezjańskiego układu współrzędnych (rys. 4a). Sumaryczny prąd płynący w antenie jest indukowany przez składową normalną natężenia pola magnetycznego H_z (rys. 4b) oraz przez składowe E_x i E_y natężenia pola elektrycznego (rys. 4c i rys. 4d odpowiednio). Wprowadzając w obwód anteny dwa symetrycznie położone obciążenia, np. dla $s = 0$ i $s = l/2$, można napisać:

$$I(0) = I_{H_z} + I_{E_y}, \tag{17}$$

$$I(l/2) = I_{H_z} - I_{E_y}. \tag{18}$$

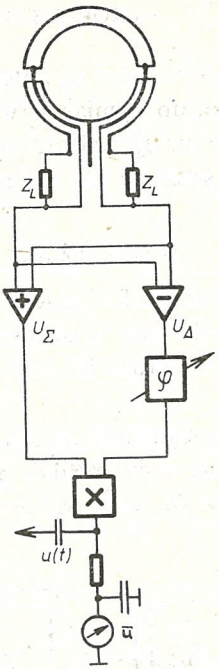
Zależności (17) i (18) są słuszne, gdy prąd I_{E_x}, skojarzony ze składową E_x, dla $s = 0$ i dla $s = l/2$ jest równy zeru, co jest jednym z założeń tej metody i wymaga odpowiedniego umieszczenia ramy względem pola.

Zasadę działania przyrządu opartego o tę metodę pokazano na rys. 5 [10]. Napięcia wyjściowe z obu obciążeń anteny ramowej są doprowadzone do układów sumujących, na wyjściu których otrzymujemy napięcie sumy U_Σ i napięcie różnicy U_Δ:

$$U_\Sigma = 2K_E E_y, \tag{19}$$

$$U_\Delta = 2K_H H_z, \tag{20}$$

gdzie: K_E i K_H — transmitancje układu dla składowej elektrycznej i magnetycznej *PEM* odpowiednio.



Rys. 5. Schemat blokowy miernika gęstości mocy z dwustronnie obciążoną anteną ramową

Przez przesuwnik fazy napięcia U_Σ i U_Δ są doprowadzane do układu mnożącego. Układ ten jest wyposażony w wyjście o napięciu proporcjonalnym do wartości chwilowej gęstości mocy i w wyjście o napięciu proporcjonalnym do wartości średniej gęstości mocy. Rozważania prowadzone przez [15] dla dwustronnie obciążonej anteny ramowej stanowią rozwinięcie rozwiązania stosowanego w radionamiernikach systemu ADCOC. Celem pracy było, obok wykazania możliwości pomiaru gęstości mocy za pomocą dwustronnie obciążonej anteny ramowej, wykazanie zmniejszenia wpływu „efektu antenowego" na dokładność pomiaru natężenia pola magnetycznego anteną ramową obciążoną dwustronnie w porównaniu z anteną ramową obciążoną jednostronnie. Przez analogię do wzoru (16) w [15] została wyprowadzona zależność dla anteny obciążonej dwustronnie

$$D \leqslant 0{,}15\lambda_{min}. \tag{21}$$

Dalsze prace autorów [15] idą w kierunku możliwości dalszego zwiększenia wymiarów anteny ramowej w porównaniu z ograniczeniem określonym wzorem (21). Osiąga się to przez stosowanie anteny ramowej wielokrotnie obciążonej; jednak, jak wykazano w [3], bardziej istotnym czynnikiem ograniczającym gabaryty anteny ramowej przeznaczonej do pomiarów w polu bliskim jest minimalna odległość od źródła *PEM*, w której mają być prowadzone pomiary, oraz wynikający z tego efekt uśredniania *PEM* na powierzchni anteny pomiarowej.

Przedstawiony zarys metody pomiaru gęstości mocy w oparciu o wykorzystanie „efektu antenowego" pozwala na sformułowanie następujących wniosków.

1. Metoda ta pozwala na stosunkowo prosty pomiar gęstości mocy, przy czym możliwe jest tu uzyskanie dużej czułości, co wynika zarówno z możliwości stosowania anten ramo-

wych o odpowiednio większych rozmiarach, jak i możliwości wzmocnienia sygnałów sumy i różnicy.

2. Stosowany w układzie miernika (rys. 5) przesuwnik fazy oraz transmitancje K_E i K_H są funkcjami częstotliwości, co ogranicza znacznie przydatność tej metody do omawianego celu, gdzie szerokopasmowość pomiaru jest jednym z podstawowych wymagań.

3. Przewody łączące antenę z układem elektronicznym w polu bliskim mogą być źródłem błędów pomiarowych wynikających z deformacji *PEM* w otoczeniu anteny pomiarowej oraz wskutek oddziaływania na źródło promieniowania — w innych rozwiązaniach sygnał z anteny, po detekcji, jest przesyłany do układu pomiarowego za pomocą linii o dużej rezystywności, która jest „przeźroczysta" dla *PEM*.

4. Ze względu na sposób połączenia anteny z układem pomiarowym budowa czujnika o sferycznej charakterystyce promieniowania może być mocno utrudniona.

5. Zależności (16) i (21) zostały wyprowadzone dla fali płaskiej, natomiast w polu bliskim nie są one słuszne; jest to istotne zwłaszcza dla źródeł dla których jest spełniona zależność

$$|\mathbf{E}|/|\mathbf{H}| > Z_o. \tag{22}$$

6. Konieczność umieszczenia anteny pomiarowej względem *PEM* tak, aby uzyskać kompensację składowej E_x, jest czynnikiem utrudniającym pomiary w pobliżu wielu źródeł wzajemnie nieskorelowanych oraz w warunkach wielotorowej propagacji.

5. POMIAR ŚREDNIEJ ARYTMETYCZNEJ P_E i P_H

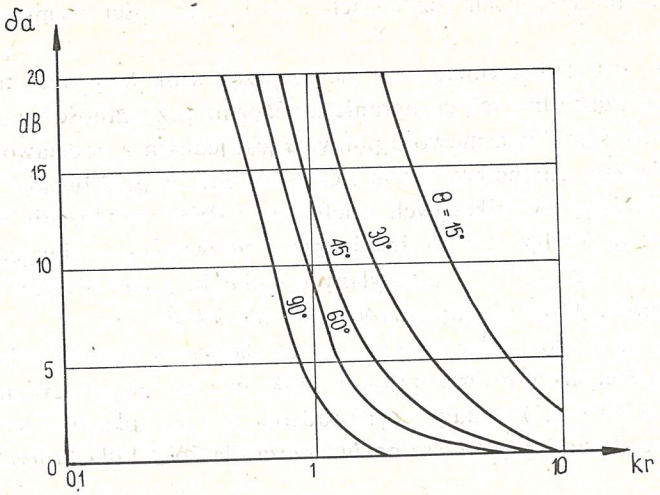
Polskie przepisy ochrony przed promieniowaniem [16] przewidują dopuszczalne wartości natężenia pola elektrycznego i magnetycznego w zakresie częstotliwości $0{,}1 \div 10$ MHz. Przepisy te nie określają jednak sposobu interpretacji wyników pomiarów w przypadku równoczesnego wystąpienia pola elektrycznego i magnetycznego o natężeniach zbliżonych do wartości dopuszczalnej, jednak jej nie przekraczających. Należałoby więc przepis taki interpretować jako alternatywny, tzn. natężenie jednej ze składowych powinno przekroczyć wartość dopuszczalną, aby pole to mogło być uznane za szkodliwe.

Przepisy wprowadzone w niektórych innych krajach, np. w Szwecji [12], poszły dalej i określając dopuszczalne wartości natężeń pola elektrycznego i magnetycznego określają także dopuszczalną sumę gęstości mocy pola elektrycznego i pola magnetycznego. Procedura ta jest równoważna wprowadzeniu — jako mierzonej wartości gęstości mocy — średniej arytmetycznej gęstości mocy pól elektrycznego i magnetycznego.

Błąd popełniony przy takim założeniu obliczymy w oparciu o gęstość mocy elektrycznej P_E i gęstość mocy magnetycznej P_H dane wzorami (9) i (10); przez analogię do (11) i (12) błąd ten (η_a) określimy w postaci

$$\eta_a = \frac{|2P_r - P_E - P_H|}{2P_r}. \tag{23}$$

Wzór (23) jest słuszny zarówno dla źródła w postaci elementarnego dipola elektrycznego, jak i elementarnego dipola magnetycznego. Ze względów przedstawionych w p. 3, na rys. 6



Rys. 6. Przebieg błędu δ_a w funkcji kr

pokazano przebiegi błędów pomiarowych określonych przez analogię do (13) i (14) w postaci

$$\delta_a = 20\log \frac{|P_E + P_H|}{2P_r}. \tag{24}$$

Porównując krzywe pokazane na rys. 6 z krzywymi na rys. 2 widać pewną poprawę dokładności pomiaru gęstości mocy w oparciu o pomiar średniej arytmetycznej gęstości mocy elektrycznej i gęstości mocy magnetycznej w porównaniu z pomiarem opartym o pomiar natężenia pola elektrycznego w pobliżu dipola elektrycznego. W kierunku $\Theta = 90°$ widać także pewną poprawę dokładności pomiaru w stosunku do krzywych pokazanych na rys. 2.

6. POMIAR ŚREDNIEJ GEOMETRYCZNEJ P_E i P_H

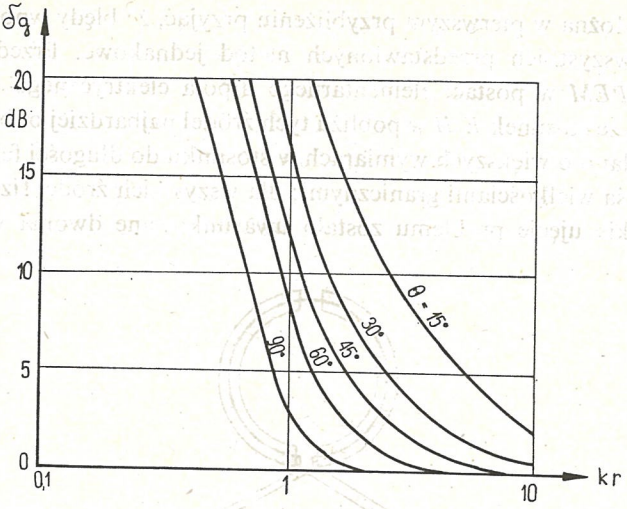
Przez analogię do rozważań przedstawionych powyżej w p. 5 policzymy dla źródła w postaci elementarnego dipola elektrycznego lub magnetycznego dokładność pomiaru gęstości mocy w jego otoczeniu w oparciu o pomiar średniej geometrycznej P_E i P_H. Błąd δ_g określimy w postaci

$$\delta_g = 20\log \frac{|\sqrt{P_E P_H}|}{P_r}. \tag{25}$$

Przebieg błędu δ_g w funkcji kr pokazano na rys. 7. W porównaniu z krzywymi pokazanymi na rys. 6 widać pewne przesunięcie poszczególnych krzywych w lewo, a więc podniesienie dokładności pomiaru.

Dla źródeł elementarnych jest spełniona zależność: $\mathbf{E} \perp \mathbf{H}$. W oparciu o wzory (1) i (2) oraz (6), pomijając zapis wektorowy, możemy napisać

$$P_s = \sqrt{P_E P_H}. \tag{26}$$



Rys. 7. Przebieg błędu δ_g w funkcji kr

Tak więc krzywe pokazane na rys. 7 pokazują równocześnie stosunek modułu gęstości mocy do gęstości mocy promieniowanej przez źródło. Zależność (26) jest słuszna dla dowolnego źródła *PEM*.

7. UWAGI I WNIOSKI

W pracy przedstawiono cztery metody pomiaru gęstości mocy w polu bliskim. Metoda oparta o pomiar natężenia pola elektrycznego lub magnetycznego (p. 3), przy czym skalę przyrządu pomiarowego cechuje się wprost w jednostkach gęstości mocy, jest stosowana m.in. w USA, Polsce i innych krajach. Metoda ta wydaje się być najwygodniejszą w stosowaniu i należałoby przyjąć jej stosowanie wszędzie tam, gdzie to jest możliwe z punktu widzenia dokładności pomiaru.

Metoda opisana w p. 4 stanowi przykład wykorzystania jednej anteny ramowej do pomiaru natężenia pola elektrycznego i magnetycznego. Metoda ta może być stosowana jedynie do pomiarów pól monochromatycznych, pól w których wzajemne położenie wektorów **E** i **H** jest dokładnie znane. Potwierdzeniem tego jest także i to, że Autorzy pracy [10] stosowali tę metodę do pomiaru rozkładu gęstości mocy w otoczeniu anten Yagi.

Metody opisane w p. 5 i p. 6 mogą być rozpatrywane dwojako:

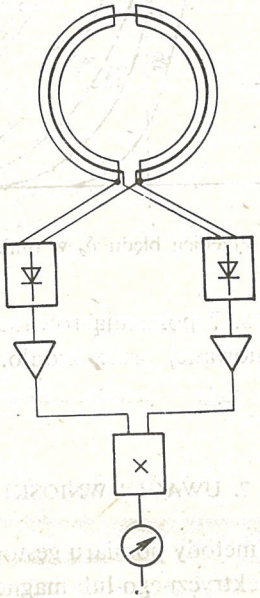
— jako sposób interpretacji wyników pomiarów natężeń pól elektrycznego i magnetycznego (wykonanych oddzielnie),

— jako wytyczne przy konstruowaniu miernika, w którym pomiary natężeń pola elektrycznego i magnetycznego są wykowywane równocześnie.

Metoda opisana w p. 6 została zaproponowana przy opracowywaniu miernika, którego schemat blokowy pokazano na rys. 8.

W pracy podano relacje zachodzące pomiędzy składowymi *PEM* bez uwzględnienia błędów wnoszonych przez przyrządy pomiarowe. Przedstawiona dyskusja dotyczy więc

błędów metody. Można w pierwszym przybliżeniu przyjąć, że błędy wnoszone przez przyrządy będą dla wszystkich przedstawionych metod jednakowe. Przedstawioną analizę oparto o źródła *PEM* w postaci elementarnego dipola elektrycznego i magnetycznego. Ze względu na to, że stosunek E/H w pobliżu tych źródeł najbardziej odbiega od Z_o (w porównaniu ze źródłami o większych wymiarach w stosunku do długości fali), przedstawione przebiegi błędów są wielkościami granicznymi; dla wszystkich źródeł fizycznych będą one majorantami. Takie ujęcie problemu zostało uwarunkowane dwoma względami:



Rys. 8. Schemat blokowy miernika gęstości mocy w polu bliskim opartego o pomiar średniej geometrycznej P_E i P_H

1. pomiary *PEM*, z omawianego punktu widzenia, są zwykle prowadzone bez jakichkolwiek danych „a priori"; wykrycie źródeł promieniowania jest jednym z celów pomiaru,
2. doświadczenie pomiarowe wskazuje na występowanie źródeł o wymiarach niewielkich w stosunku do długości fali, np. wzbudniki pieców indukcyjnych, szpary i otwory w ekranach urządzeń generacyjnych.

Niezależnie od przyjętej metody pomiaru z przedstawionych rozważań widać, że dla $kr < 1$ pomiary zatracają sens, choć dla źródeł zdeterminowanych możliwe jest potraktowanie rozważanych błędów jako systematycznych i uwzględnienie ich w wyniku pomiaru — procedura taka wydaje się być nazbyt złożona dla praktyki metrologicznej. Pewną poprawę dokładności daje stosowanie metody opartej o pomiar średniej geometrycznej gęstości mocy elektrycznej i magnetycznej. Niezależnie od rodzaju źródła *PEM* wynik takiego pomiaru jest równy modułowi wektora Poyntinga i może być miarą maksymalnej gęstości mocy. Jest to szczególnie korzystne z punktu widzenia pomiarów związanych z ochroną przed biologicznym działaniem *PEM*. Stwierdzenie powyższe w niczym nie zmienia faktu, że część urojona wektora Poyntinga nie ma określonego sensu fizycznego.

Ze względu na stosowanie w miernictwie pól bliskich anten pomiarowych o wymiarach znacznie mniejszych od długości fali, oraz największe rozpowszechnienie (i największą sprawność promieniowania) źródeł o wymiarach porównywalnych z długością fali, najbardziej reprezentatywnym źródłem byłby przewód z prądem, w otoczeniu którego błędy pomiaru gęstości mocy są znacznie mniejsze od wyżej przedstawionych. Zagadnienie to wymaga jednak obszerniejszego opracowania i zostanie przedstawione w oddzielnej publikacji. Ponadto, jak wyżej wspomniano, ze względu na nieoznaczoność zagadnienia konieczne jest oparcie się na błędach granicznych, które zostały przedstawione.

LITERATURA

1. E. Aslan, *A low frequency H-field radiation monitor*, Proc. 1975 USNC/URSI Meeting, Boulder, CO.
2. T. M. Babij, H. Trzaska, *Szerokopasmowe własności czujników pola elektrycznego*, Archiwum Elektrotechniki, Nr 2/78, s. 335—352.
3. T. M. Babij, H. Trzaska, *Properties of wideband magnetic field probes*, Proc. 1976 IEEE EMC Intl. Symp. Washington, DC.
4. T. M. Babij, H. Trzaska, *Power density flux measurements*, Proc. 1979 BEMS Meeting, Seattle, WA.
5. H. I. Bassen, *A broadband, miniature, isotropic electric field measurement system*, Proc. 1975 IEEE EMC Intl. Symp., San Antonio, TX.
6. D. J. Bem, *Anteny i rozchodzenie się fal radiowych*, WNT, 1973.
7. R. R. Bowmann, *Power density measurements*, wykład w Politechnice Wrocławskiej, wrzesień 1973.
8. A. W. Guy, *Nonionizing electromagnetic wave effects in biological materials and systems*, wykład w Politechnice Wrocławskiej, wrzesień 1975.
9. F. M. Greene, *Final report on NBS field measurements, probe design and use*, NBS Project 2720376, NBS Boulder, CO.
10. D. Hoff, G. Monich, *Eine Sonde zur direkten Messung von Energiestrommungen im Nahfeld von Sende und Empfangsantennen*, NTZ, Bd. 27/1974, Heft 8, s. 313-318.
11. H. Kucia, A. Koperski, *Zagadnienia miernictwa pól elektromagnetycznych szkodliwych dla zdrowia*, Prace CIOP, Nr 62/1970.
12. K. H. Mild, G. Olsson, *Preliminary studies of electric and magnetic fields at 27 MHz*, Proc. 1977 Intl. Symp. on the Biological Effects of EM Waves, Airlie, VA.
13. B. Minin, *S. W. Cz. i bezopasnost' čeloveka*, Moskwa 1974.
14. S. W. Rosenthal, *Bioelectromagnetics at PINY*, wykład w Politechnice Wrocławskiej, lipiec 1980.
15. H. Whiteside, R. W. P. King, *The loop antenna as a probe*, IEEE Trans., vol. AP-7, Nr 5, May 1964, s. 291—297.
16. Rozporządzenie Ministrów Pracy, Płacy i Spraw Socjalnych oraz Zdrowia i Opieki Społecznej z dnia 19 lutego 1977 r. w sprawie bezpieczeństwa i higieny pracy przy stosowaniu urządzeń wytwarzających pola elektromagnetyczne w zakresie od 0.1 do 300 MHz (Dz.U.PRL Nr 8 z 19 marca 1977 r, poz. 33).
17. Rozporządzenie Rady Ministrów z dnia 25 maja 1972 r. w sprawie bezpieczeństwa i higieny pracy przy stosowaniu urządzeń wytwarzających pola elektromagnetyczne w zakresie mikrofalowym (Dz.U.Nr 21 z 8 czerwca 1972 r., poz. 153).

H. TRZASKA

METHODS OF THE NEAR FIELD POWER DENSITY MEASUREMENTS

Summary

The methods of the near field power density measurements from the point of view of the measurement accuracy are presented. The work discusses methods based upon:
1. electric or magnetic power density measurement,
2. the use of the doubly loaded loop antenna,
3. arithmetical mean of electric and magnetic power densities,
4. geometrical mean of electric and magnetic power densities.

It has been shown, that at frequencies above 300 MHz the most convenient is method (1). At frequencies below 300 MHz, measuring errors lead to the necessity of using the 4-th method. The result of this measurement is modulus of the Poynting vector.

H. TRZASKA

MÉTHODE DE MESURE DE LA DENSITÉ DE PUISSANCE DANS LE CHAMP PROCHE DE LA SOURCE DE RAYONNEMENT ÉLECTROMAGNÉTIQUE

Résumé

Dans l'étude on a présenté les méthodes de mesure de la densité de puissance dans la champ proche, du point de vue de la précision de mesure. On a discuté les méthodes basées sur:
1. la mesure de la densité de puissance,
2. l'utilisation de l'antenne en cadre chargée bilatéralement,
3. la moyenne arithmétique de la densité de puissance électrique et magnétique,
4. la moyenne géométrique de la densité de puissance électrique et magnétique.

On a montré que pour les fréquences dépassant 300 MHz l'application de la méthode (1) est la plus avantageuse, tandis que pour les fréquences inférieures de 300 MHz les moindres erreurs de mesure sont la preuve de la nécessité de l'application de la méthode (4). Le module du vecteur de Poynting est le résultat de mesure avec cette méthode.

H. TRZASKA

MEßMETHODEN FÜR DIE LEISTUNGSFLUSSDICHTE IN EINEM NAHFELD

Zusammenfassung

Im obigen Aufsatz wurden Meßmethoden für die Leistungsflussdichte im Nahfeld vom Gesichtspunkt der Meßgenauigkeit dargestellt. Diskutiert wurden Meßmethoden in Anlehnung an:
1. Messung der elektrischen und magnetischen Leistungsflussdichte,
2. Nutzung einer zweiseitig belasteten Rahmenantenne,
3. den arithmetischen Mittelwert der elektrischen und magnetischen Leistungsflussdichte,
4. den geometrischen Mittelwert der elektrischen und magnetischen Leistungsflussdichte.

Es wurde bewiesen, daß bei Frequenzen über 300 MHz die Anwendung der Methode (1) am günstigsten ist, wogegen bei Frequenzen unter 300 MHz weisen die geringsten Meßfehler darauf hin, daß hier die Methode (4) angebracht ist. Das Meßergebnis, das mittels dieser Methode gewonnen worden ist, ist der Poynting-Vektormodul.

X. ТШАСКА

МЕТОДЫ ИЗМЕРЕНИЯ ПЛОТНОСТИ ПОТОКА ЭНЕРГИИ В БЛИЖНЕИ ЗОНЕ

Р е з ю м е

Представлены методы измерений плотности потока энергии в ближнеи зоне источника электромагнитного излугения с точки зрения погрешности измерений. Рассмотрены методы обоснованные на:

1. измерении плотности потока энергии электрического или магнитного поля,
2. измерении двухсторонно нагруженной рамочной антенной,
3. арифметической средней плотности потоков электрического и магнитного полей,
4. геометрической средней плотности потоков электрического и магнитного полей.

Показано, что при частотах свыше 300 МГц наиболее удобным явльяется метод (1). При частотах ниже 300 МГц самые малые измерительные ошибки диктуют необходимость ползования методом (4). Результатом измерений с применением этого метода является модуль вектора Умова-Поинтинга.

ROZPRAWY ELEKTROTECHNICZNE 1982, 28, z. 1-2, ss. 285-313

621.394.18

Kryptograficzne metody zabezpieczenia informacji

JÓZEF PIEPRZYK (BYDGOSZCZ)

Instytut Telekomunikacji i Elektrotechniki Akademii Technicznej w Bydgoszczy

DOMINIK RUTKOWSKI (SOPOT)

Instytut Telekomunikacji Politechniki Gdańskiej

Otrzymano 12.5.1980

Praca stanowi przegląd, stosowanych zarówno w systemach jak również w sieciach komputerowych, kryptograficznych metod zabezpieczenia informacji. Mimo, że metody kryptograficzne są stosowane od czasów starożytnych, to dopiero w ostatnich latach dało się zaobserwować wyraźne ożywienie w podejmowanych pracach badawczych, a co za tym idzie nastąpił znaczny rozwój teorii kryptografii. Dlatego w pracy położono nacisk na przedstawienie najnowszych tendencji w rozwoju kryptografii. W tym celu, po podaniu klasyfikacji metod kryptograficznego zabezpieczenia i jej krótkim omówieniu, zaprezentowano dwie metody ochrony informacji przesyłanych pomiędzy końcówką a systemem komputerowym. Następnie omówiono algorytmy szyfrowania, a mianowicie amerykański algorytm *DES* oraz algorytm oparty na modułach szyfrujących. Algorytmy szyfrowania można traktować jako podstawowe elementy, z których można budować złożone systemy kryptograficzne, np. rekurencyjne systemy kryptograficzne. Również ochronę informacji w sieci komputerowej można wprowadzić w oparciu o algorytmy szyfrowania. Taka koncepcja została przedstawiona w pracy oraz omówione zostały problemy związane z zarządzaniem kluczami kryptograficznymi. Praca kończy się zaprezentowaniem interesującego problemu projektowania publicznych systemów kryptograficznych, które to systemy są szczególnie użyteczne przy tworzeniu kryptograficznego zabezpieczenia informacji w sieciach komputerowych.

1. WSTĘP

Pojawienie się w latach sześćdziesiątych wielodostępowych systemów komputerowych, a w latach siedemdziesiątych sieci komputerowych spowodowało wzrost zainteresowania problemami zabezpieczenia przetwarzanych, przechowywanych oraz przesyłanych informacji komputerowych przed dostępem do nich nieautoryzowanego użytkownika (*NU*). Dostęp *NU* może prowadzić do tak zwanych przestępstw komputerowych [6, 22, 32, 33], które wiążą się w szczególności z nieautoryzowanym

— odczytaniem informacji,
— modyfikacją informacji,
— zniszczeniem informacji,
— zablokowaniem dostępu do zbiorów informacji dla autoryzowanego użytkownika (*AU*).

Znane metody zabezpieczenia informacji można podzielić na następujące kategorie, a mianowicie

— metody zabezpieczenia fizycznego,
— metody zabezpieczenia operacyjnego,
— metody zabezpieczenia kryptograficznego.

Metody fizycznego zabezpieczenia informacji związane są z fizycznym ograniczeniem dostępu dla *NU* oraz z ochroną systemów i sieci komputerowych przed działaniem sił przyrody [4, 13, 22]. Z kolei metody zabezpieczenia operacyjnego są określone przez organizację i sposób zarządzania zasobami systemów komputerowych przez nadzorujące ich działanie systemy operacyjne [21, 25, 27, 39]. Natomiast metody zabezpieczenia kryptograficznego polegają na wykorzystaniu przekształceń kryptograficznych w celu ochrony informacji przed jej modyfikacją i odczytaniem przez *NU*. Warto tutaj podkreślić, że metody zabezpieczenia kryptograficznego w stosunku do poprzednio wymienionych metod są bardziej uniwersalne i pozwalają dostosować stopień zabezpieczenia informacji do wymagań użytkowników oraz zapewnić ochronę końcówek, kanałów oraz baz danych przed nieautoryzowanym dostępem [7, 34, 40, 41]. Ze względu na wyżej wymienione zalety metod zabezpieczenia kryptograficznego poświęcimy im w dalszym ciągu główną uwagę.

2. KLASYFIKACJA METOD ZABEZPIECZENIA KRYPTOGRAFICZNEGO

Klasyfikację metod zabezpieczenia kryptograficznego można przeprowadzać uwzględniając różne kryteria, a mianowicie:

A — typ obiektów systemu lub sieci komputerowej, do ochrony których są one stosowane,
B — sposób wytwarzania kryptogramów (zaszyfrowanych wiadomości),
C — rodzaj dostarczanej ochrony,
D — sposób przekazywania klucza kryptograficznego do urządzenia szyfrującego.

W dalszym ciągu pracy przedstawimy i omówimy poszczególne klasyfikacje. I tak metody zabezpieczenia kryptograficznego z punktu widzenia typu obiektów, do ochrony których są one stosowane, można podzielić na:

A1. metody kryptograficznego zabezpieczenia końcówek,
A2. metody kryptograficznego zabezpieczenia kanałów,
A3. metody kryptograficznego zabezpieczenia baz danych,
A4. metody kryptograficznego zabezpieczenia sieci komputerowych.

W przypadku końcówek metody kryptograficznego zabezpieczenia są stosowane do stwierdzania, po pierwsze tożsamości zgłaszającego się użytkownika poprzez kontrolę posiadanego przez niego hasła lub klucza kryptograficznego [30, 31, 45], po drugie tożsamości systemu komputerowego, z którym końcówka współpracuje [15, 42]. Natomiast ochrona kanałów polega na zaprojektowaniu odpowiedniego przekształcenia kryptograficznego oraz na doborze procedury jej stosowania w taki sposób, aby prawdopodobieństwa zarówno odczytania wiadomości z przesyłanych kryptogramów, jak też wymiany oryginalnych kryptogramów na kryptogramy dla fałszywych wiadomości były mniejsze od z góry określonych liczb [18, 34, 35]. Z kolei ochrona baz danych wymaga stosowania

przekształceń kryptograficznych do zabezpieczenia informacji przechowywanych w pamięci operacyjnej i pamięci zewnętrznej systemu komputerowego oraz w pamięci zewnętrznej tworzącej tzw. duplikaty bazy danych [1, 20, 34, 44]. Kryptograficzna ochrona sieci komputerowych polega na stosowaniu przekształceń kryptograficznych do ochrony informacji przekazywanych w kanałach łączących poszczególne systemy komputerowe (węzły) ze sobą. Ze względu na specyfikę sieci komputerowych powstają tutaj dodatkowe problemy związane z generowaniem i rozdziałem kluczy kryptograficznych pomiędzy [12, 26, 28]:
— użytkownikami a systemem komputerowym,
— końcówkami a systemem komputerowym,
— systemami komputerowymi.

Jeżeli jako kryterium podziału weźmiemy sposób wytwarzania kryptogramów, to metody zabezpieczenia kryptograficznego można podzielić na:
B1. blokowe,
B2. rekurencyjne.

W blokowych metodach zabezpieczenia kryptograficznego klucz kryptograficzny jest wybierany dla każdej wiadomości niezależnie od poprzednich wiadomości, a kryptogram jest wytwarzany wyłącznie na podstawie aktualnej wiadomości i klucza [2, 34]. Dla rekurencyjnej metody zabezpieczenia kryptograficznego kluczami są podciągi ciągu liczb, wytwarzanych przez generator pseudolosowy na podstawie tzw. „nasienia" [3, 24, 37] lub są to ciągi liczb otrzymywanych z klucza kryptograficznego oraz kryptogramu nadanego w chwili poprzedniej, tzw. systemy kryptograficzne ze sprzężeniem zwrotnym [7, 24, 12, 36].

Przechodzimy teraz do następnej klasyfikacji metod zabezpieczenia kryptograficznego w przypadku, gdy kryterium podziału jest rodzaj ochrony dostarczanej przez te metody. I tak metody zabezpieczenia kryptograficznego można podzielić na metody, które stanowią ochronę przed:
C1. odczytaniem wiadomości z kryptogramu przez *NU*, gdy zna on pewien zbiór zaobserwowanych kryptogramów,
C2. wymianą wiadomości na inną, fałszywą wiadomość, gdy *NU* zna pewien zbiór zaobserwowanych kryptogramów,
C3. wymianą wiadomości na inną, fałszywą wiadomość, gdy *NU* zna pewien zbiór par obserwacji (wiadomość, kryptogram).

Metody zabezpieczenia kryptograficznego stosowane do ochrony wiadomości przed ich odczytaniem z kryptogramów są stosunkowo najlepiej zbadane i doczekały się wielu opracowań [10, 16, 40, 41]. Taki stan rzeczy wynika z faktu, że rozwój kryptografii był przede wszystkim związany z ochroną wiadomości przesyłanych przez kanał, który może być kontrolowany przez *NU*. Jest to tak zwany klasyczny problem kryptografii telekomunikacyjnej [40, 41]. Natomiast metody zabezpieczenia kryptograficznego stosowane do ochrony wiadomości przed ich wymianą muszą spełniać pewne dodatkowe warunki związane z faktem, że *NU* może znać pewien zbiór obserwacji par (wiadomość, kryptogram), jak ma to miejsce, gdy *NU* jest osobą mającą bezpośredni kontakt z systemem komputerowym (programista, operator). W praktyce, gdy zależy nam tylko na zabezpieczeniu informacji przed jej wymianą przez *NU*, stosuje się takie przekształcenia kryptograficzne, dla których kryptogram daje się podzielić na dwie części. Pierwszą część stanowi

wiadomość w jawnej postaci (chronimy wiadomości przed nielegalną wymianą nie zaś przed odczytaniem ich z kryptogramów), natomiast drugą częścią jest tzw. autentyfikator, którego długość jest tak dobrana, aby prawdopodobieństwo udanej, nielegalnej wymiany wiadomości było mniejsze od z góry zadanej liczby [18, 34, 35]. Warto dodać, że metody zabezpieczenia kryptograficznego nie chronią informacji przed działaniem *NU*, a jedynie pozwalają na wykrycie takiej działalności.

Ostatnim kryterium podziału metod zabezpieczenia kryptograficznego jest sposób przekazywania kluczy kryptograficznych do szyfratorów (deszyfratorów). Z tego punktu widzenia metody zabezpieczenia kryptograficznego można podzielić na:

D1. klasyczne systemy kryptograficzne,

D2. publiczne systemy kryptograficzne.

W klasycznych systemach kryptograficznych klucz jest przekazywany do szyfratora (deszyfratora) poprzez kanał, który jest poza zasięgiem działania *NU*. W praktycznych rozwiązaniach klucze przesyłane są za pomocą poczty [5], lub też za pomocą odpowiedniego kanału telekomunikacyjnego z zabezpieczeniem kryptograficznym [10, 12]. W pierwszym rozwiązaniu stosowany „kanał" ma duże opóźnienie, co w poważnym stopniu ogranicza jego stosowanie w sieciach komputerowych. W drugim rozwiązaniu opóźnienie jest o wiele mniejsze, niemniej jednak, gdy *NU* określi klucz na podstawie zaobserwowanych kryptogramów, to może on bez przeszkód odczytywać i zmieniać przesyłane wiadomości. W publicznych systemach kryptograficznych, które będą omówione później, stosowaną są trzy klucze, przy czym klucz, którego używa nadawca kryptogramów, jest przesyłany w jawny sposób przez odbiorcę kryptogramów. Odbiorca kryptogramów w celu zdeszyfrowania kryptogramu używa klucza tajnego, znanego tylko po stronie odbiorczej. Para kluczy (klucz jawny, klucz tajny) jest wytwarzana z klucza podstawowego również nie znanego dla nadawcy [8, 11, 38]. Taka koncepcja tworzenia systemów kryptograficznych w znakomity sposób ułatwia rozdział kluczy pomiędzy użytkownikami i może znaleźć duże pole do zastosowań w sieciach komputerowych.

W dalszym ciągu pracy podamy praktyczne metody zabezpieczenia kryptograficznego informacji przekazywanych pomiędzy końcówką a komputerem. Następnie przedstawimy znane algorytmy szyfrowania oraz możliwości ich zastosowania do ochrony informacji. Z kolei określimy podstawowe problemy występujące podczas projektowania zabezpieczenia kryptograficznego w sieciach komputerowych. W dalszym punkcie zaprezentujemy koncepcję publicznych systemów kryptograficznych. A na zakończenie zastanowimy się nad problemem oceny jakości zabezpieczenia kryptograficznego.

3. KRYPTOGRAFICZNE ZABEZPIECZENIE WIADOMOŚCI PRZESYŁANYCH POMIĘDZY KOŃCÓWKĄ A KOMPUTEREM

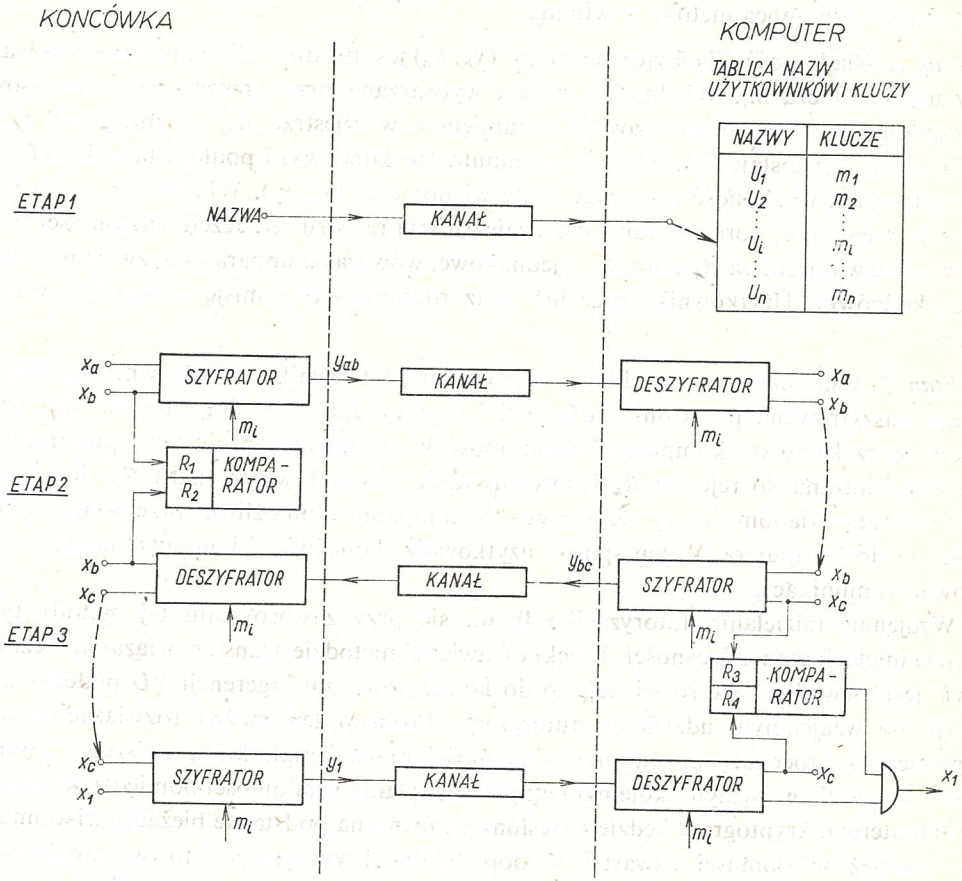
Omówimy teraz stosowane w systemach komputerowych metody inicjowania seansu łączności i kryptograficzne zabezpieczenie transmisji wiadomości pomiędzy końcówką a komputerem. Przedstawimy dwie metody kryptograficznego zabezpieczenia transmisji wiadomości pomiędzy końcówką a komputerem, a mianowicie:

— metodę powitania [15, 16, 42],

— metodę hasła [15, 45, 46].

3.1. Metoda powitania

Przed rozpoczęciem transmisji wiadomości między końcówką a komputerem konieczna jest przede wszystkim wzajemna identyfikacja użytkownika i komputera, gdyż może zdarzyć się, że *NU* dołączy do kanału zewnętrzny komputer i symulując komputer obsługujący końcówkę będzie odbierał wiadomości przesyłane przez użytkownika, albo podając się za *AU* będzie przesyłał własne wiadomości do komputera obsługującego końcówkę. Każdy z *AU* posiada identyfikującą go nazwę oraz charakterystyczny dla niego klucz kryptograficzny. W pamięci komputera znajduje się zbiór nazw wszystkich użytkowników i przyporządkowane im klucze kryptograficzne. Omawiana tutaj kryptograficzna metoda zabezpieczenia transmisji, zwana metodą powitania, odbywa się w trzech etapach.



Rys. 1. Schemat etapów inicjacji seansu łączności przy wykorzystaniu metody powitania

Etap 1. Pierwszą wiadomością przesyłaną z końcówki przez kanał do komputera jest nazwa użytkownika i jest ona przekazywana w postaci nie szyfrowanej. Drugą wiadomością przesyłaną bezpośrednio po pierwszej jest kryptogram $y_{ab} = f[(x_a, x_b), m_i]$ (patrz

rys. 1), gdzie wiadomość (x_a, x_b) oznacza parę wiadomości, przy czym x_a reprezentuje żądanie udzielenia końcówce autoryzacji przez komputer, natomiast x_b służy do udzielenia komputerowi autoryzacji przez końcówkę. Wiadomość x_b jest wybierana w końcówce w sposób losowy i niezależny od tego, który użytkownik korzysta z końcówki. Ciąg x_b zapisywany jest w rejestrze R_1 komparatora końcówki. Następnie para wiadomości (x_a, x_b) jest szyfrowana za pomocą klucza kryptograficznego m_i przyporządkowanego i-temu użytkownikowi. Kryptogram y_{ab} zostaje odebrany przez komputer, który na podstawie nazwy użytkownika odczytuje z tablicy odpowiadający mu klucz m_i. Klucz ten jest używany przez komputer do deszyfracji kryptogramu y_{ab}. Po rozszyfrowaniu kryptogramu po stronie odbiorczej wiadomo, na podstawie wiadomości x_a, że określony użytkownik żąda udzielenia autoryzacji. Wiadomość x_b zostaje zapamiętana w komputerze w celu realizacji dalszego etapu kryptograficznego zabezpieczenia przesyłanych wiadomości za pomocą metody powitania.

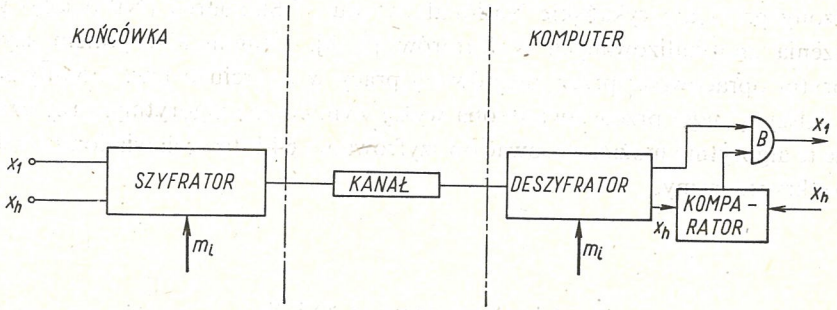
Etap 2. Wiadomość składająca się z pary (x_b, x_c) jest po stronie komputera szyfrowana przy użyciu klucza m_i. Wiadomość x_c jest wytwarzana przez generator pseudolosowy umieszczony w komputerze i zostaje zapamiętana w rejestrze R_3. Kryptogram $y_{bc} = f[(x_b, x_c), m_i]$ zostaje wysłany przez komputer do końcówki i podlega tam deszyfracji. Rozszyfrowana wiadomość ma postać pary wiadomości (x_b, x_c). Wiadomość x_b po wpisaniu do rejestru R_2 porównywana jest z zawartością rejestru R_1. Jeżeli wiadomości przechowywane w rejestrach R_1 oraz R_2 są jednakowe, wówczas komparator zezwala na dalszą pracę końcówki. Użytkownik może już teraz rozpocząć transmisję właściwych wiadomości.

Etap 3. Wiadomość x_1 ze zbioru wiadomości użytkownika wraz z wiadomością x_c zostaje zaszyfrowana po stronie końcówki i kryptogram $y_1 = f[(x_c, x_1), m_i]$ jest przesyłany przez kanał do komputera. Wiadomość x_c po jej wydobyciu z kryptogramu y_1 jest wprowadzana do rejestru R_4 i porównywana z zawartością rejestru R_3. Jeżeli obie porównywane wiadomości są identyczne, to komparator umożliwia przekazanie wiadomości x_1 do komputera. W ten sposób użytkownik, końcówka i komputer mogą kontynuować komunikację.

Wzajemne udzielanie autoryzacji odbywa się przy zastosowaniu tej metody tylko w fazie inicjacji seansu łączności. Dzięki omawianej metodzie szansa nawiązania łączności z *NU* jest niewielka. Nie rozwiązuje to do końca problemu ingerencji *NU* podczas transmisji, po wzajemnym udzieleniu autoryzacji. Problem ten można rozwiązać poprzez odpowiednią procedurę przesyłania wiadomości przez kanał, którą opiszemy poniżej. Jeśli mianowicie w każdym, kolejnym etapie przesyłania wiadomości pomiędzy końcówką a komputerem, kryptogram będzie określony zarówno na podstawie bieżącej wiadomości, jak również wiadomości zawartej w poprzednim kryptogramie, to w kanale będą przesyłane powiązane ze sobą kryptogramy o postaci $f[(x_c, x_1), m_i]$, $f[(x_1, x_2), m_i]$, $f[(x_2, x_3), m_i]$, ... Dzięki takiemu rozwiązaniu istnieje podwójne zabezpieczenie przed ingerencją *NU*. Pierwsze z nich polega na szyfrowaniu wiadomości, drugie zaś na wzajemnym uzależnieniu każdego kryptogramu od wiadomości zawartej w poprzednim kryptogramie. Mankamentem tego rozwiązania jest wzrost średniego opóźnienia transmisji wiadomości między końcówką a komputerem.

3.2. Metoda hasła

Identyfikacja użytkownika i udzielanie autoryzacji końcówce przez komputer przy wykorzystaniu zabezpieczenia kryptograficznego w oparciu o metodę hasła zostały przedstawione schematycznie na rys. 2. Hasłem x_h jest pewien ciąg znaków alfanumerycznych,



Rys. 2. Schemat kryptograficznego zabezpieczenia transmisji przy wykorzystaniu metody hasła

znany tylko użytkownikowi i komputerowi. Ciąg ten można traktować jako nazwę użytkownika. W pamięci komputera znajduje się spis kluczy wraz z nazwami użytkowników. Określony użytkownik przesyła najpierw swoją nazwę w jawnej postaci przez kanał do komputera, następnie końcówka dokonuje szyfracji pary wiadomości, którą tworzy: hasło, pierwsza wiadomość ze zbioru wiadomości użytkownika, za pomocą klucza m_i. Komputer po odbiorze nazwy odszukuje na liście nazw użytkowników i kluczy odpowiadający odebranej nazwie klucz kryptograficzny. Dzięki temu deszyfrator po stronie komputera może rozszyfrować pierwszy kryptogram. Z rozszyfrowanego kryptogramu wydzielone zostaje hasło i wiadomość. Otrzymane w ten sposób hasło jest porównywane w komparatorze (rys. 2) z hasłem pamiętanym w komputerze dla danego użytkownika. Jeśli porównanie wykaże identyczność haseł, wówczas bramka B umożliwia przesyłanie wiadomości do komputera. Takie rozwiązanie jest równoznaczne z udzieleniem końcówce autoryzacji przez komputer.

Opisywana tu metoda hasła jest, jak widać, prostsza od metody powitania. Przy stosowaniu tej metody nie istnieje potrzeba dwukrotnej transmisji tej samej wiadomości w postaci zaszyfrowanej, a zabezpieczenie przed dostępem NU do kanału podczas transmisji polega na odpowiedniej budowie kryptogramów. Jeżeli użytkownik przesyła po kolei wiadomości $x_1, x_2, \ldots, x_u$, to do kanału wprowadzane są kolejno kryptogramy $f[(x_1, x_h), m_i], f[(x_2, x_h), m_i], \ldots, f[(x_u, x_h), m_i]$. Wynika stąd, że każdy, kolejny kryptogram zawiera to samo hasło. Porównując metody powitania i hasła można zauważyć pewne podobieństwa i różnice. Każdy, kolejny kryptogram w obu metodach zawiera powtórzenie jednej z wiadomości, która występowała w poprzednim kryptogramie z tym, że w metodzie powitania jest to kolejna wiadomość użytkownika, a w metodzie hasła zawsze to samo hasło. Ponieważ hasło może być krótsze niż wiadomość, wobec tego metoda hasła umożliwia uzyskanie mniejszego opóźnienia transmisji wiadomości niż metoda powitania. Jednakże w metodzie hasła jedną z wiadomości, na podstawie której wyznaczany jest każdy kryptogram, jest ta sama wiadomość (hasło), co w porównaniu z metodą powitania zwiększa szansę złamania klucza.

4. ALGORYTMY SZYFROWANIA

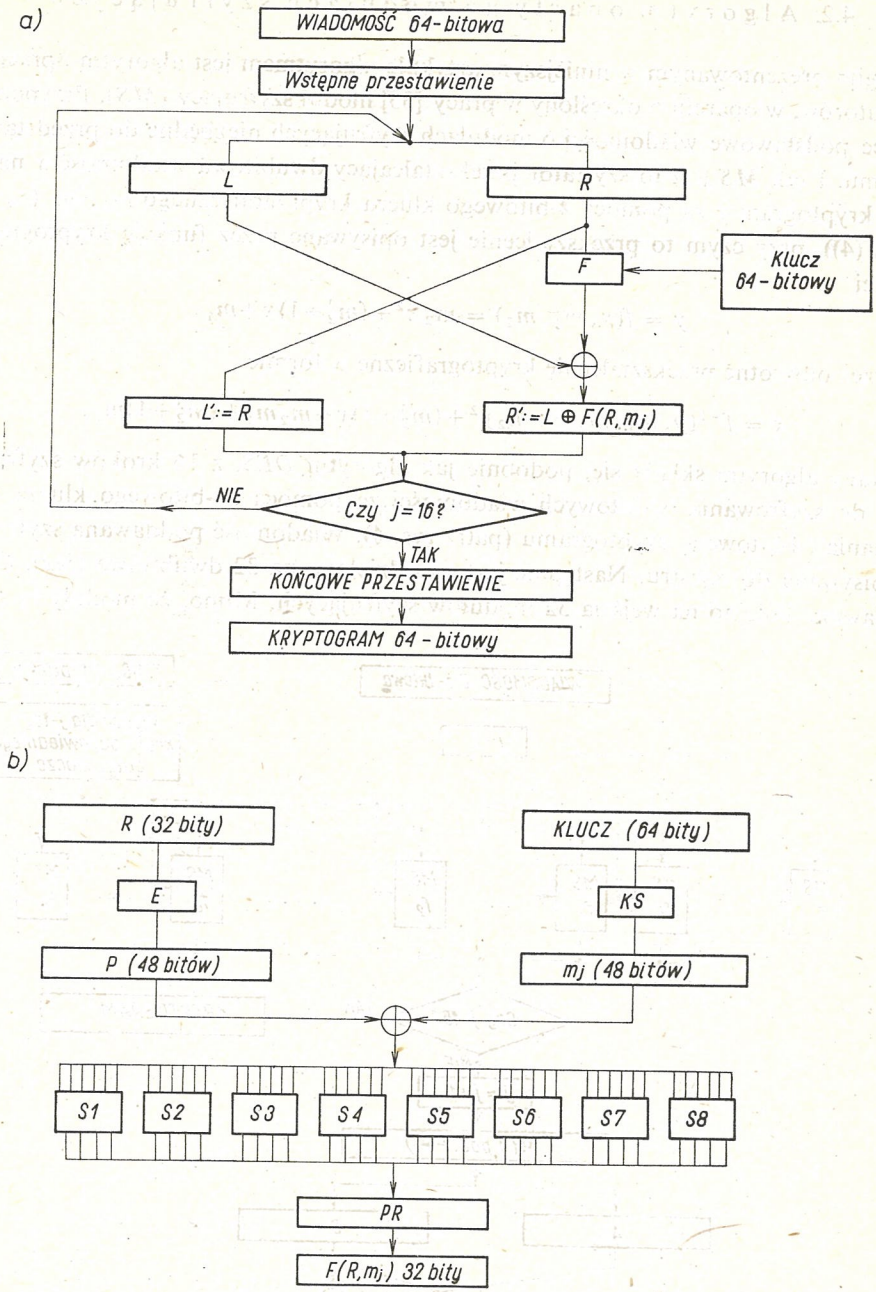
W praktyce daje się zauważyć tendencja do stosowania szyfratorów, które pracują w oparciu o znormalizowane algorytmy wytwarzania kryptogramów. Pierwszym z takich algorytmów jest algorytm *DES* (*Data Encryption Standard*) opracowany przez IBM oraz zatwierdzony przez amerykańskie National Bureau of Standards (NBS) jako podstawa do tworzenia znormalizowanych szyfratorów [2, 9]. Drugim algorytmem szyfrowania jest algorytm opracowany przez autorów tej pracy w oparciu o tzw. moduły szyfrujące [34]. W dalszym ciągu przedstawimy oba wyżej wymienione algorytmy, przy czym warto dodać, że te algorytmy można stosować do szyfrowania 64-bitowych wiadomości uzyskując 64-bitowe kryptogramy.

4.1. Algorytm DES

Jak już wspomniano, algorytm *DES* służy do otrzymywania 64-bitowych kryptogramów na podstawie 64-bitowej wiadomości i 64-bitowego klucza. Przy czym klucz składa się z 56 bitów wybieranych w sposób niezależny oraz z 8 bitów, które są dobierane zgodnie z odpowiednimi równaniami kontroli parzystości jedynek. Algorytm zawiera 16 kroków szyfrowania (patrz rys. 3), z których każdy składa się z tych samych operacji wykonywanych dla klucza pomocniczego m_j, a m_j jest różne dla każdego kroku (gdzie $1 \leqslant j \leqslant 16$ jest numerem wykonywanego kroku). Proces szyfrowania wiadomości jest następujący. Wiadomość 64-bitowa jest wprowadzana do rejestru. Po wstępnym przestawieniu bitów następuje podział wiadomości na dwa 32-bitowe ciągi, które są umieszczane w rejestrach L oraz R (rys. 3a). Zawartości rejestrów L oraz R po ich przekształceniu są wpisywane do rejestrów L' oraz R', przy czym obowiązują następujące zależności

1. $L' := R$; czyli zawartość rejestru R jest przepisywana bezpośrednio do rejestru L',
2. $R' := L \oplus F(R, m_j)$; gdzie $\oplus$ oznacza sumę modulo dwa odpowiednich bitów ciągów L oraz $F(R, m_j)$.

Na uwagę zasługuje sposób uzyskiwania 32-bitowego ciągu $F(R, m_j)$ (patrz rys. 3b). Punktem wyjściowym do otrzymania ciągu $F(R, m_j)$ są zawartość rejestru R oraz klucz kryptograficzny m. Zawartość rejestru R po powtórzeniu niektórych z bitów (operacja E) jest wpisywana do 48-bitowego rejestru P. Z drugiej strony z 64-bitowego klucza m jest wybierany 48-bitowy klucz pomocniczy m_j i wpisywany do odpowiedniego rejestru. Zawartość rejestru P oraz klucz pomocniczy m_j są dodawane modulo dwa bit po bicie, a uzyskany 48-bitowy ciąg jest podawany na wejścia ośmiu różnych bloków realizujących operacje podstawienia (bloki od $S1$ do $S8$), gdzie każdy z tych bloków logicznych przetwarza 6-bitowy ciąg wejściowy na 4-bitowy ciąg wyjściowy. Bity w tak otrzymanym 32-bitowym ciągu są przestawiane w odpowiedni sposób (blok PR) i otrzymujemy końcową postać 32-bitowego ciągu $F(R, m_j)$. Z kolei zawartości rejestrów L' oraz R' są przepisywane do rejestrów L oraz R i jest wykonywany następny krok szyfrowania. Po wykonaniu 16 takich kroków i po zastosowaniu końcowego przestawienia otrzymujemy 64-bitowy kryptogram.



Rys. 3. Standardowy algorytm szyfrowania (*DES*)
a) sieć działań algorytmu, b) sposób wytwarzania funkcji F

Cenną własnością algorytmu *DES* jest fakt, że proces deszyfrowania jest realizowany przez ten sam algorytm z tą różnicą, że klucze pomocnicze m_j $(1 \leqslant j \leqslant 16)$ są generowane w odwrotnej kolejności z klucza kryptograficznego m.

4.2. Algorytm oparty na modułach szyfrujących

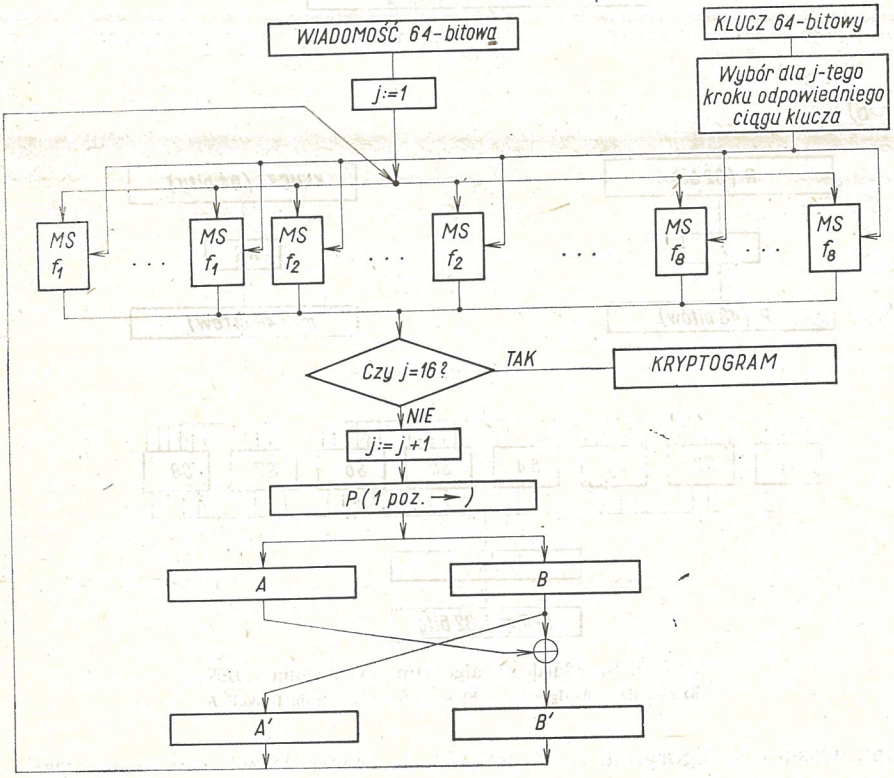
Drugim prezentowanym w niniejszym artykule algorytmem jest algorytm opracowany przez autorów, w oparciu o określony w pracy [35] moduł szyfrujący (*MS*). Przypomnimy pokrótce podstawowe wiadomości o modułach szyfrujących niezbędne do przedstawienia algorytmu. I tak *MS* jest to szyfrator przekształcający dwubitową wiadomość x na dwubitowy kryptogram y za pomocą 4-bitowego klucza kryptograficznego m_1, m_2 ($x, y, m_1, m_2 \in GF(4)$), przy czym to przekształcenie jest opisywane przez funkcję kryptograficzną o postaci

$$y = f(x, m_1, m_2) = m_2 x^2 + (m_2^3 + 1)x + m_1 \tag{4.1}$$

oraz przez odwrotne przekształcenie kryptograficzne w formie

$$x = f^{-1}(y, m_1, m_2) = m_2 y^2 + (m_2^3 + 1)y + m_2 m_1^2 + (m_2^3 + 1)m_1 . \tag{4.2}$$

Opisywany algorytm składa się, podobnie jak algorytm *DES*, z 16 kroków szyfrowania i służy do szyfrowania 64-bitowych wiadomości za pomocą 64-bitowego klucza w celu otrzymania 64-bitowego kryptogramu (patrz rys. 4). Wiadomość poddawana szyfrowaniu jest wpisywana do rejestru. Następnie jest ona dzielona na 32 dwubitowe ciągi. Ciągi te są podawane kolejno na wejścia 32 modułów szyfrujących. Mimo, że moduły te mają tę



Rys. 4. Uproszczony schemat sieci działań algorytmu szyfrowania zbudowany w oparciu o moduły szyfrujące (*MS*)

samą budowę realizują one różne funkcje kryptograficzne (funkcje od f_1 do f_8). Uzyskano to dzięki ustaleniu, na czas procesu szyfrowania, sygnałów na wejściach m_1 modułów szyfrujących, przy czym prawdziwe są następujące zależności:

$$f_1(x, m) = f(x, m_1 = 00, m_2 = m),$$
$$f_2(x, m) = f(x, m_1 = 01, m_2 = m),$$
$$f_3(x, m) = f(x, m_1 = 10, m_2 = m),$$
$$f_4(x, m) = f(x, m_1 = 11, m_2 = m),$$
$$f_5(x, m) = f^{-1}(x, m_1 = 11, m_2 = m),$$
$$f_6(x, m) = f^{-1}(x, m_1 = 10, m_2 = m),$$
$$f_7(x, m) = f^{-1}(x, m_1 = 01, m_2 = m),$$
$$f_8(x, m) = f^{-1}(x, m_1 = 00, m_2 = m),$$

gdzie funkcje f oraz f^{-1} są określone równaniami (4.1) oraz (4.2). Sygnały uzyskane z wyjść modułów szyfrujących, przy określonym dla każdego kroku j kluczu pomocniczym m_j ($1 \leqslant j \leqslant 16$), są wpisywane do rejestru przesuwnego P i przesuwane o jedną pozycję w prawo. Tak otrzymany ciąg dzielony jest na dwa 32-bitowe podciągi i są one wpisywane do rejestrów A i B. Z kolei zawartości rejestrów A oraz B są przepisywane do rejestrów A' oraz B' zgodnie z poniższymi równościami

1. $A' := B$,
2. $B' := A \oplus B$.

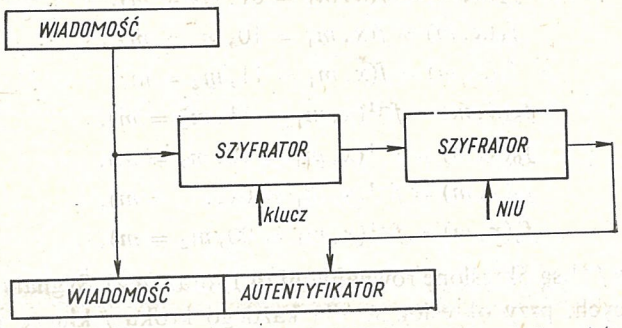
Na tej operacji kończy się pierwszy krok szyfrowania. Dalsze kroki wykonywane są identycznie jak opisany poprzednio. Proces deszyfracji w tym algorytmie polega na wykonaniu odwrotnych operacji w odwrotnej kolejności. Konstrukcja algorytmu pozwala wykorzystać go także w procesie deszyfracji po wykonaniu odpowiedniej permutacji wprowadzanego kryptogramu.

Omówione powyżej dwa algorytmy składają się z 16 kroków szyfrowania i służą do szyfrowania 64-bitowych wiadomości za pomocą 64-bitowego klucza kryptograficznego. Dla obu algorytmów problem określenia klucza kryptograficznego na podstawie znajomości pary (wiadomość, kryptogram) wiąże się z rozwiązywaniem 64 równań boolowskich, z których każde opisuje zależność elementarnego wyjścia kryptogramu od 64 wejść wiadomości i 64 wejść klucza. Z powyższego rozumowania wynika, że określenie klucza wymaga średnio 2^{63} elementarnych operacji przetwarzania. Podsumowując można powiedzieć [34], że oba algorytmy są porównywalne z punktu widzenia czasu wytwarzania kryptogramu oraz liczby elementarnych operacji przetwarzania niezbędnych do złamania szyfru. Niemniej jednak algorytm oparty na modułach szyfrujących w stosunku do algorytmu *DES* ma tę zaletę, że posiada strukturę modularną, co w znakomity sposób ułatwia jego realizację sprzętową.

4.3. Zastosowanie algorytmów szyfrowania

Algorytmy szyfrowania stanowią podstawowy element dla projektantów przy tworzeniu złożonych systemów kryptograficznych. I tak w systemach elektronicznego przesyłania rachunków bankowych [23, 36] występuje konieczność potwierdzania autentyczności

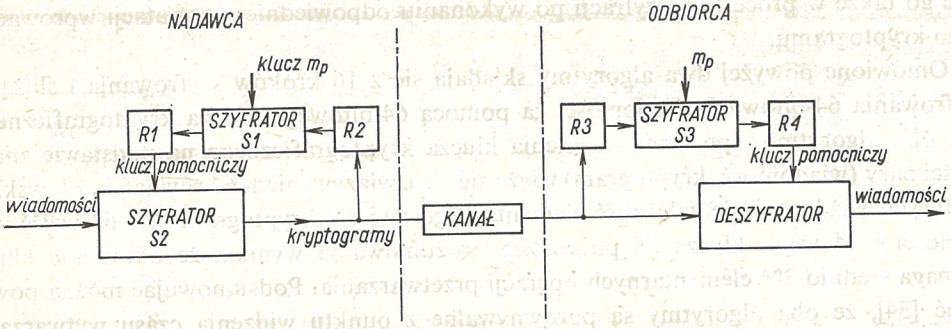
nadawanych wiadomości, przy czym te wiadomości mogą być przesyłane w jawnej postaci. Prowadzi to nas do tzw. szyfrów wykrywających podstawienie fałszywej wiadomości [34]. Potwierdzenie autentyczności danej wiadomości jest możliwe dzięki dołączaniu do przesyłanej wiadomości autentyfikatora (rys. 5). Autentyfikator jest wytwarzany



Rys. 5. Schemat tworzenia autentyfikatora dla wiadomości

z wiadomości poprzez jej dwukrotne szyfrowanie kluczem kryptograficznym oraz numerem identyfikacyjnym użytkownika (*NIU*), który tę wiadomość przesyła.

Szyfratory realizujące algorytm *DES* lub algorytm oparty na *MS*-ach są przykładami blokowych systemów kryptograficznych. Okazuje się, że łatwo można zbudować za pomocą takich szyfratorów rekurencyjny system kryptograficzny [7, 12, 36]. Schemat przykładowego rozwiązania takiego sytemu jest pokazany na rys. 6. Po stronie nadawczej



Rys. 6. Schemat kryptograficznego zabezpieczenia rekurencyjnego

znajdują się dwa szyfratory *S1* i *S2* oraz dwa rejestry *R1* i *R2*. W chwili początkowej rejestr *R2* jest zerowany i jego zawartość jest szyfrowana kluczem podstawowym m_p a wynik wpisywany jest do rejestru *R1* tworząc tzw. klucz pomocniczy m_{p_1}. Pierwsza wiadomość x_1 podawana na wejście szyfratora *S2* jest szyfrowana kluczem pomocniczym m_{p_1} i otrzymujemy kryptogram y_1. Dla i-tej w kolejności nadawania wiadomości (wiadomości x_i) klucz pomocniczy m_{p_i} jest wytwarzany na podstawie kryptogramu y_{i-1} (kryptogramu w chwili (i-1)-wszej) oraz klucza podstawowego m_p. Zatem obowiązują następujące równości:

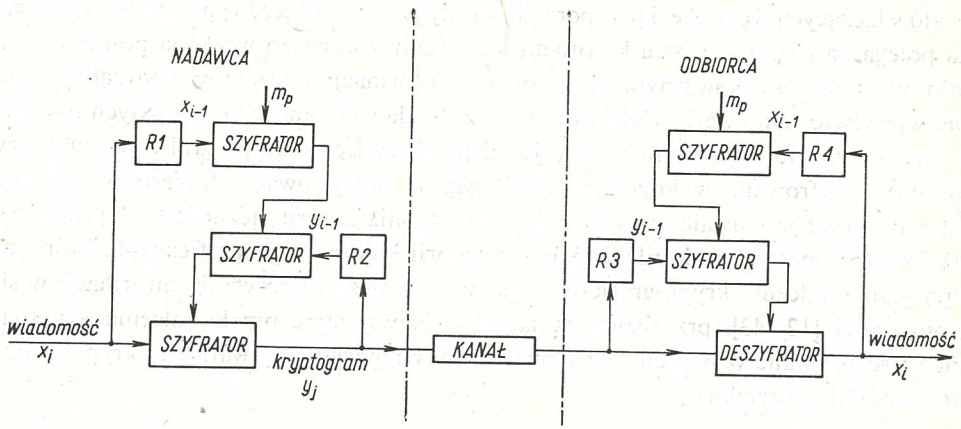
$$y_1 = f(x_1, m_{p_1}), \quad \text{gdzie} \quad m_{p_1} = f(0, m_p),$$
$$y_i = f(x_i, m_{p_i}), \quad \text{gdzie} \quad m_{p_i} = f(y_{i-1}, m_p), \quad i \geqslant 2, \tag{4.3}$$

przy czym f jest funkcją kryptograficzną szyfratorów *S1* oraz *S2*.

Proces deszyfracji odbywa się zgodnie z zależnościami:

$$x_1 = f^{-1}(y_1, m_{p_1}),$$
$$x_i = f^{-1}(y_i, m_{p_i}); \quad i \geqslant 2. \tag{4.4}$$

Powyżej opisany rekurencyjny system kryptograficzny nosi nazwę systemu kryptograficznego ze sprzężeniem zwrotnym kryptogramów [12, 24]. Uogólnieniem tego systemu



Rys. 7. Schemat systemu kryptograficznego ze sprzężeniem zwrotnym wiadomości i kryptogramów

jest rekurencyjny system kryptograficzny ze sprzężeniem zwrotnym wiadomości i kryptogramów (rys. 7). Systemy te są opisane równaniami o postaci

$$y_1 = f\big(x_1, f\big(y_0, f(x_0, m_p)\big)\big),$$

gdzie y_0, x_0 tworzą warunek początkowy ustalony zarówno po stronie nadawczej jak też odbiorczej, oraz

$$y_i = f\big(x_i, f\big(y_{i-1}, f(x_{i-1}, m_p)\big)\big), \quad \text{dla } i \geqslant 2. \tag{4.5}$$

Przedstawione powyżej systemy kryptograficzne ze sprzężeniami zwrotnymi mogą być uproszczone poprzez zastąpienie niektórych szyfratorów elementami realizującymi dodawanie modulo dwa bit po bicie [12, 7, 24].

5. ZARZĄDZANIE KLUCZAMI KRYPTOGRAFICZNYMI W SIECI KOMPUTEROWEJ

Opracowanie znormalizowanych algorytmów szyfrowania w znacznym stopniu ułatwiło stosowanie kryptograficznych metod do zabezpieczenia informacji przesyłanych w sieciach komputerowych. Niemniej jednak pojawiły się nowe problemy związane z rozdziałem kluczy kryptograficznych podczas stosowania przekształceń kryptograficznych.

Wiadomo, że przekształcenia kryptograficzne mogą być stosowane w sieci komputerowej do ochrony informacji przesyłanych pomiędzy:

P1. końcówką a węzłem[1] sieci komputerowej,
P2. węzłami sieci komputerowej,
P3. końcówkami sieci komputerowej.

Zakładamy, że informacje przesyłane pomiędzy końcówką a jej macierzystym węzłem są zabezpieczane dzięki stosowaniu przekształcenia kryptograficznego, przy czym jest ono kontrolowane za pomocą pojedynczego klucza wybieranego oddzielnie dla każdego seansu łączności. W dalszym ciągu przyjmujemy, że również informacje przesyłane pomiędzy węzłami sieci są szyfrowane. Jeżeli droga przesyłania zaszyfrowanych informacji zawiera kilka węzłów tranzytowych, wtedy w każdym z tych węzłów następuje zmiana klucza (przeszyfrowywanie). Konieczność zmiany klucza wynika z faktu, że każdemu z kanałów łączących węzły sieci przyporządkowany jest inny klucz kryptograficzny, a zmiana ta polega na rozszyfrowaniu kryptogramów wchodzących do węzła za pomocą klucza kanału wejściowego i ponownym zaszyfrowaniu informacji kluczem przyporządkowanym kanałowi wyjściowemu węzła. Załóżmy również, że dla wiadomości przesyłanych pomiędzy końcówkami źródłową i docelową stosuje się przekształcenie kryptograficzne, przy czym wiadomości szyfrowane w końcówce źródłowej są deszyfrowane dopiero w końcówce docelowej (klucz jest ustalony przez cały czas trwania seansu łączności). Z powyższych założeń wynika potrzeba określenia kilku kategorii kluczy kryptograficznych, które kontrolują przekształcenia kryptograficzne stosowane podczas przesyłania informacji w sieci komputerowej [12, 43], przy czym zakładamy, że wszystkie przekształcenia kryptograficzne są realizowane w oparciu o znormalizowany algorytm wytwarzania kryptogramów (patrz rozdział poprzedni).

5.1. Zabezpieczenie wiadomości przesyłanych pomiędzy końcówką a jej macierzystym węzłem

W celu sprecyzowania rozważań wprowadzimy teraz kilka kategorii kluczy kryptograficznych, które służą do ochrony wiadomości przesyłanych pomiędzy końcówką a jej macierzystym węzłem. I tak każdej końcówce sieci komputerowej przyporządkowany jest jej klucz pierwotny m_{PK} (klucz pierwotny końcówki). Klucz ten jest przechowywany zarówno w końcówce (w jawnej postaci), jak też w macierzystym węźle końcówki (w postaci zaszyfrowanej). Do szyfrowania kluczy pierwotnych odpowiednich końcówek określonego węzła jest wykorzystywany klucz pierwotny węzła m_{PW}, który jest w nim generowany i przechowywany w węźle w jawnej postaci. Natomiast do szyfrowania wiadomości przesyłanych pomiędzy końcówkami (użytkownikami) sieci komputerowej stosowany jest klucz nazywany kluczem roboczym m_R. Klucz ten jest generowany oddzielnie dla każdego seansu łączności w węźle macierzystym końcówki, z której jest inicjowana transmisja, oraz jest przechowywany w węźle i przesyłany do końcówki w postaci zaszyfrowanej $f_{m_{PK}}(m_R)$[2]. Warto w tym miejscu wyraźnie zaznaczyć, że stosowanie zabezpieczenia

[1] Przez węzeł rozumiemy wielodostępowy system komputerowy, do którego dołączone są końcówki abonenckie.

[2] Oznaczenie $f_m(x) = f(x, m)$, gdzie $f(.,*): X \times M \to Y$, przy czym X jest zbiorem wiadomości, M — zbiorem kluczy, Y — zbiorem kryptogramów, oraz $x \in X$, $m \in M$, $f(x, m) \in Y$.

kryptograficznego ma sens tylko wtedy, gdy klucze nie są znane dla *NU*. Wynika stąd wniosek, że w celu zapewnienia odpowiedniej ochrony wiadomości wymaga się przechowywania kluczy w obszarach pamięci całkowicie zabezpieczonych przed dostępem *NU*. Chodzi tu przede wszystkim o klucze przechowywane w jawnej postaci, czyli o klucz m_{PK} przechowywany w końcówce oraz o klucz m_{PW} przechowywany w węźle.

W celu określenia organizacji i zarządzania kluczami w sieci komputerowej wprowadzimy teraz operacje kryptograficzne, które są zdefiniowane zarówno dla końcówki jak również dla węzła. Operacje te są wykonywane przez odpowiednie programy. I tak dla końcówki można określić następujące operacje kryptograficzne.

OK1. Deszyfracja kryptogramu klucza roboczego (operacja *DFK*)

$$DFK\colon \{f_{m_{PK}}(m_R)\} \to m_R,$$

gdzie *DFK* jest nazwą operacji kryptograficznej określonej na elemencie $f_{m_{PK}}(m_R)$ kryptogram dla klucza m_R uzyskany po zastosowaniu w procesie szyfracji klucza m_{PK}, a wynikiem tej operacji jest wartość klucza m_R w jawnej postaci. Krótko mówiąc, $f_{m_{PK}}(m_R)$ jest daną wejściową operacji, natomiast m_R jest wynikiem jej wykonania. W praktyce operacja ta jest wykonywana następująco. Program realizujący tę operację ma bezpośredni dostęp do klucza m_{PK} i po przeprowadzeniu deszyfracji wynik umieszcza w odpowiednim rejestrze, który nie jest dostępny dla użytkowników. Klucz roboczy pozostaje tam do momentu pojawienia się nowej wartości klucza.

OK2. Szyfracja wiadomości x za pomocą klucza m_R

$$SZFK\colon \{x\} \to f_{m_R}(x).$$

OK3. Deszyfracja kryptogramu wiadomości x przy użyciu klucza m_R

$$DSZFK\colon \{f_{m_R}(x)\} \to x.$$

Z kolei dla węzła określone są operacje, które mają postać:

OW1. Szyfracja wiadomości x za pomocą klucza m

$$SZF\colon \{f_{m_{PW}}(m), x\} \to f_m(x).$$

Elementy w nawiasach oznaczają dane wejściowe, natomiast strzałka wskazuje wynik operacji.

OW2. Deszyfracja kryptogramu $f_m(x)$ w celu określenia wiadomości x

$$DSZF\colon \{f_{m_{PW}}(m), f_m(x)\} \to x.$$

Zanim podamy następną operację kryptograficzną dla węzła, najpierw uzasadnimy konieczność jej zdefiniowania. W tym celu powracamy do określenia klucza roboczego m_R. Jak wiadomo, po pierwsze, klucz ten jest zawsze zmieniany po zakończeniu każdego seansu łączności pomiędzy końcówką a węzłem, po drugie, jest wytwarzany w węźle, po trzecie, jest przechowywany w węźle i przesyłany do końcówki w postaci zaszyfrowanej. Zakładamy, że w węźle zamiast generowania klucza m_R w jawnej postaci wytwarzamy jego kryptogram $f_{m_{PW}}(m_R)$. Do tego celu, jako generator liczb pseudolosowych, można wykorzystać omówione w poprzednim rozdziale algorytmy szyfrowania [17]. Zatem,

jeżeli z generatora liczb pseudolosowych otrzymamy liczbę pseudolosową l_p, to zachodzi równość

$$l_p = f_{m_{PW}}(m_R).$$

Ponieważ do końcówki należy przesłać klucz m_R w postaci kryptogramu $f_{m_{PK}}(m_R)$ (patrz operacja *DFK*), a w węźle jest przechowywany kryptogram $f_{m_{PW}}(m_R)$, więc istnieje konieczność uzyskania kryptogramu $f_{m_{PK}}(m_R)$ na podstawie kryptogramu $f_{m_{PW}}(m_R)$. Z drugiej strony zakładamy, że wszystkie klucze, oprócz klucza m_{PW}, są przechowywane w węźle w postaci zaszyfrowanej. Zatem istnieje kryptogram klucza pierwotnego końcówki w formie $f_{m_{PW}}(m_{PK})$. W przypadku, gdy określimy pożądany kryptogram klucza m_R w postaci $f_{m_{PK}}(m_R)$, to po zastosowaniu w węźle operacji kryptograficznej *DSZF*, otrzymamy

$$DSZF\colon \{f_{m_{PW}}(m_{PK}), f_{m_{PK}}(m_R)\} \to m_R,$$

co przeczy naszemu założeniu, że wszystkie klucze przechowywane w węźle, oprócz klucza m_{PW}, są zapisane w postaci odpowiednich kryptogramów.

W celu uniknięcia tej sprzeczności definiujemy nową kategorię klucza kryptograficznego, który będzie służył wyłącznie do szyfrowania kluczy pierwotnych końcówki, a nazywać go będziemy kluczem pierwotnym węzła m_{PW1}. W praktycznych rozwiązaniach klucz m_{PW1} jest tworzony z klucza m_{PW} poprzez przestawienie wybranych bitów klucza m_{PW}. Stąd wynika wniosek, że uzyskanie kryptogramu $f_{m_{PK}}(m_R)$ z kryptogramu $f_{m_{PW}}(m_R)$ wiąże się z zastosowaniem kryptogramu $f_{m_{PW1}}(m_{PK})$. A więc mamy definicję nowej operacji kryptograficznej.

OW3. Przeszyfrowanie klucza m_R (dla transmisji)

$$PRSZF\colon \{f_{m_{PW1}}(m_{PK}), f_{m_{PW}}(m_R)\} \to f_{m_{PK}}(m_R).$$

W węzłach sieci komputerowej występuje potrzeba przechowywania wiadomości w postaci pakietów. Poprzednio przyjęliśmy, że pakiety (wiadomości) są przechowywane w postaci zaszyfrowanej. Jak wiadomo, pakiety przychodzące do węzłów z końcówek są szyfrowane kluczem m_R. W celu zapewnienia odpowiedniej ochrony kryptograficznej definiujemy nową kategorię kluczy, które będziemy nazywać pierwotnymi kluczami pakietów (m_{PP}). Zakładamy, że klucze m_R służące do szyfrowania pakietów będą szyfrowane za pomocą klucza m_{PP}, oraz że do nagłówka każdego pakietu dołączany jest odpowiedni kryptogram dla klucza roboczego m_R (dołączany kryptogram ma postać $f_{m_{PP}}(m_R)$). Natomiast klucz m_{PP} jest szyfrowany kluczem pierwotnym węzła m_{PW2}[3)], który powstaje, podobnie jak klucz m_{PW1}, poprzez przestawienie wybranych bitów klucza m_{PW}. W przypadku gdy chcemy wysłać odpowiedni pakiet do końcówki, konieczne jest uzyskanie kryptogramu $f_{m_{PW}}(m_R)$ (patrz operacja OW3), podczas gdy klucz m_R jest przechowywany w postaci $f_{m_{PP}}(m_R)$. W tym celu określamy nową operację kryptograficzną wykonywaną w węźle.

OW4. Przeszyfrowanie klucza m_R (dla pakietów)

$$PRSZFP\colon \{f_{m_{PW2}}(m_{PP}), f_{m_{PP}}(m_R)\} \to f_{m_{PW}}(m_R).$$

Jak widać, operacja *PRSZFP* pozwala na uzyskanie $f_{m_{PW}}(m_R)$ z odpowiednich kryptogramów przechowywanych w węźle.

[3)] Uzasadnienie potrzeby stosowania klucza m_{PW2} jest identyczne jak dla klucza m_{PW1}.

5.2. Zabezpieczenie informacji przesyłanej pomiędzy węzłami

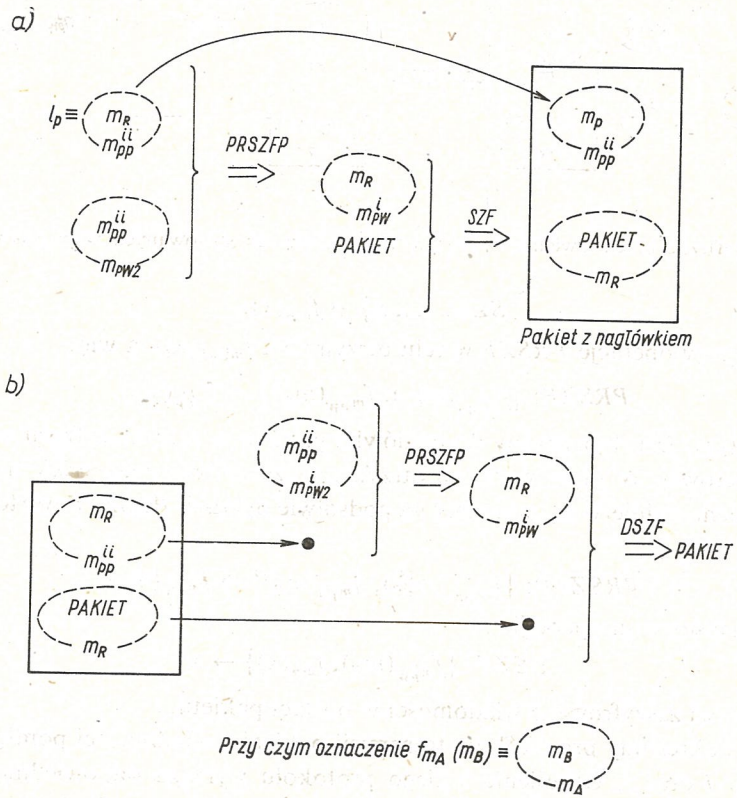
Z poprzednich rozważań wynika, że w i-tym węźle istnieje klucz pierwotny węzła m^i_{PW} oraz jego modyfikacje m^i_{PW1}, m^i_{PW2}. W dalszym ciągu uogólnimy kategorie kluczy zdefiniowane poprzednio (chodzi o klucz m_{PP}) oraz wprowadzimy następne konieczne do przeprowadzenia dalszych rozważań. Mamy zatem

m^{ii}_{PP} — pierwotny klucz dla pakietów, które są szyfrowane i deszyfrowane w i-tym węźle sieci,

m^{ij}_{PP} — pierwotny klucz dla pakietów, które są szyfrowane w i-tym węźle a będą deszyfrowane w j-tym węźle sieci, przy czym $m^{ij}_{PP} \neq m^{ji}_{PP}$. Klucz m^{ij}_{PP} jest znany zarówno w i-tym węźle w postaci $f_{m^i_{PW_1}}(m^{ij}_{PP})$, jak też w węźle j-tym jako $f_{m^j_{PW_2}}(m^{ij}_{PP})$.

W dalszym ciągu pracy podamy protokoły dla stosowania zabezpieczenia kryptograficznego w sieciach komputerowych.

PR1. Protokół szyfrowania dla pakietów, które są gromadzone i przetwarzane tylko w i-tym węźle z pominięciem końcówek (rys. 8).



Rys. 8. Ilustracja graficzna przebiegu procesów; a) szyfrowania, b) deszyfrowania pakietów dla protokołu *PR1*

W tym celu wybieramy liczbę pseudolosową l_p i przyjmujemy, że

$$l_p = f_{m_{PP}^{ii}}(m_R).$$

Jeżeli zastosujemy operację kryptograficzną *PRSZFP*, to możemy uzyskać interesujący nas kryptogram $f_{m_{PW}^{i}}(m_R)$ w następujący sposób

$$PRSZFP\colon \{f_{m_{PW_2}^{i}}(m_{PP}^{ii}),\ l_p\} \to f_{m_{PW}^{i}}(m_R).$$

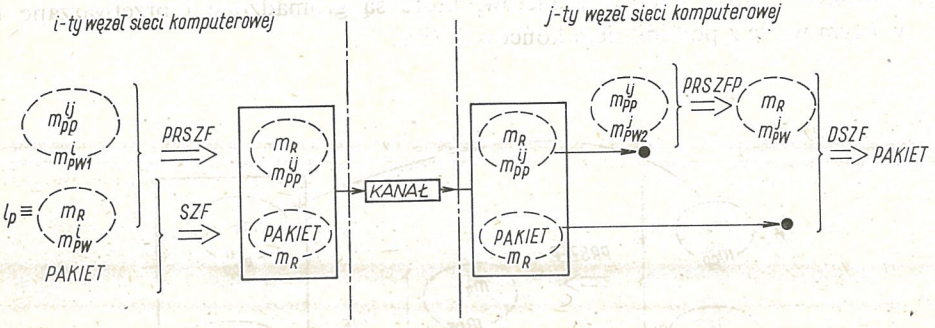
Po zastosowaniu operacji *SZF* możemy otrzymać pakiety w postaci zaszyfrowanej i do nagłówka pakietu wprowadzany jest kryptogram $f_{m_{PP}^{ii}}(m_R)$. Deszyfracja pakietów przebiega poprzez zastosowanie kolejno operacji *PRSZFP* oraz *DSZF*.

PR2. Protokół szyfrowania pakietów w i-tym węźle oraz deszyfrowania w j-tym węźle sieci komputerowej (rys. 9).

Przyjmujemy, że

$$l_p = f_m^i(m_R).$$

Stąd operacja kryptograficzna *SZF* pozwala na zaszyfrowanie pakietu x w postaci



Rys. 9. Graficzne przedstawienie operacji kryptograficznych stosowanych w protokole *PR2*

$$SZF\colon \{l_p, x\} \to f_{m_R}(x).$$

Z kolei stosujemy operację *PRSZF* w celu otrzymania $f_{m_{PP}^{ij}}(m_R)$, a więc

$$PRSZF\colon \{f_{m_{PW_1}^{i}}(m_{PP}^{ij}), f_{m_{PW}}(m_R)\} \to f_{m_{PP}^{ij}}(m_R).$$

Element $f_{m_{PP}^{ij}}(m_R)$ jest dopisywany do nagłówka pakietu w celu deszyfracji w j-tym węźle sieci komputerowej. W j-tym węźle po otrzymaniu zaszyfrowanego pakietu z odpowiednim nagłówkiem, najpierw wyznaczamy na podstawie operacji *PRSZFP* wartość $f_{m_{PW}^{j}}(m_R)$, czyli

$$PRSZFP\colon \{f_{m_{PW_2}^{j}}(m_{PP}^{ij}),\ f_{m_{PP}^{ij}}(m_R)\} \to f_{m_{PW}^{j}}(m_R).$$

Zatem po zastosowaniu operacji

$$DSZF\colon \{f_{m_{PW}^{j}}(m_R), f_{m_R}(x)\} \to x$$

otrzymujemy nie zaszyfrowane wiadomości w postaci pakietu.

Następnie określimy protokół dla transmisji pakietów wiadomości pomiędzy końcówkami węzłów i oraz j. Określenie takiego protokołu wymaga odpowiedniego rozdziału kluczy roboczych pomiędzy końcówkami węzłów i oraz j. W tym celu definiujemy poniżej następujące kategorie kluczy kryptograficznych:

m_T^{ii} — kategoria kluczy przeznaczonych do ochrony klucza roboczego m_R podczas jego transmisji między i-tym węzłem a jego końcówkami. Czyli zbiór wszystkich kluczy m_T^{ii} jest równy zbiorowi kluczy m_{PK} dla końcówek, których i-ty węzeł jest węzłem macierzystym. Stąd mamy $\{m_T^{ii}\} = \{m_{PK1}^i, \ldots, m_{PKn}^i\}$, przy czym m_{PKl}^i $(1 \leqslant l \leqslant n)$ jest kluczem pierwotnym l-tej końcówki i-tego węzła.

m_T^{ij} — kategoria kluczy przeznaczonych do ochrony klucza roboczego podczas jego transmisji pomiędzy węzłami i oraz j, przy czym klucz m_R jest generowany w i-tym węźle oraz jest przesyłany i przechowywany w celu deszyfracji pakietów (wiadomości) w j-tym węźle. Klucz m_T^{ij} jest przechowywany w i-tym węźle w formie $f_{m_{PW_1}^i}(m_T^{ij})$, natomiast w j-tym węźle w postaci $f_{m_{PW_2}^j}(m_T^{ij})$.

PR3. Protokół przesyłania pakietów (wiadomości) z l-tej końcówki i-tego węzła do n-tej końcówki j-tego węzła (rys. 10).

Podobnie jak poprzednio wybieramy liczbę pseudolosową w i-tym węźle oraz zakładamy, że

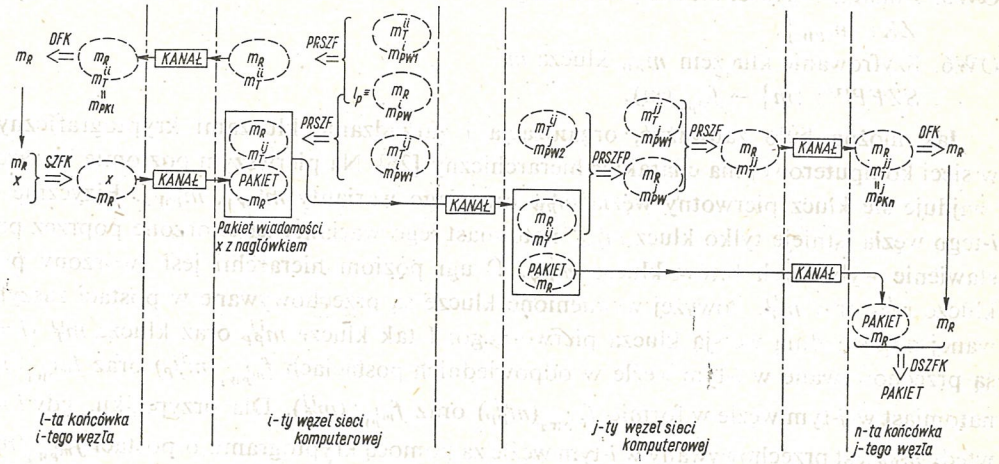
$$l_p = f_{m_{PW}^i}(m_R).$$

W oparciu o wartość l_p oraz przy użyciu operacji

$$PRSZF\colon \{f_{m_{PW_1}^i}(m_T^{ii}), l_p\} \to f_{m_T^{ij}}(m_R) = f_{m_{PKl}^i}(m_R)$$

otrzymujemy $f_{m_{PKl}^i}(m_R)$. Następnie przesyłamy otrzymaną wartość kryptogramu dla klucza roboczego do l-tej końcówki. W końcówce stosujemy operację kryptograficzną DFK i otrzymujemy

$$DFK\colon \{f_{m_{PKl}^i}(m_R)\} \to m_R.$$



Rys. 10. Schemat operacji kryptograficznych wykonywanych przy przesyłaniu wiadomości pomiędzy końcówkami sieci komputerowej (protokół $PR3$)

Z kolei szyfrowanie wiadomości x przy znajomości wartości klucza m_R sprowadza się do zastosowania operacji $SZFK$, czyli

$$SZFK\colon \{x\} \to f_{m_R}(x).$$

Zaszyfrowane wiadomości z końcówki są przesyłane do jej macierzystego węzła (i-ty węzeł) i tam są gromadzone w celu stworzenia odpowiedniego pakietu. Do nagłówka tak otrzymanego pakietu dołączany jest element $f_{m_T^{ij}}(m_R)$, przy czym otrzymujemy go za pomocą operacji

$$PRSZF\colon \{f_{m^i_{PW_1}}(m_T^{ij}), f_{m^i_{PW}}(m_R)\} \to f_{m_T^{ij}}(m_R).$$

Pakiet zaszyfrowanych wiadomości z odpowiednim nagłówkiem jest przesyłany do j-tego węzła. W węźle tym stosujemy operację przeszyfrowywania klucza m_R, stąd

$$PRSZFP\colon \{f_{m^j_{PW_2}}(m_T^{ij}), f_{m_T^{ij}}(m_R)\} \to f_{m^j_{PW}}(m_R),$$

przy czym wiadomo, że element $f_{m^j_{PW_2}}(m_T^{ij})$ jest przechowywany w j-tym węźle. W dalszym ciągu obliczamy wartość $f_{m_T^{ij}}(m_R)$, czyli

$$PRSZF\colon \{f_{m^{ij}_{PW_1}}(m_T^{ij}), f_{m^j_{PW}}(m_R)\} \to f_{m_T^{ij}}(m_R) = f_{m^j_{PKn}}(m_R).$$

Wartość $f_{m^j_{PKn}}(m_R)$ przesyłamy do n-tej końcówki j-tego węzła, gdzie po zastosowaniu operacji *DFK* otrzymujemy wartość klucza m_R. Po zastosowaniu operacji kryptograficznej *DSZFK* dostajemy wiadomości w jawnej postaci.

W rozdziale tym pominięto ważną kwestię generowania klucza pierwotnego węzła oraz jego odpowiedniego zabezpieczenia (klucz ten jest przechowywany, jako jedyny, w jawnej postaci [12, 28]). Nie bez znaczenia jest, pominięty także, problem zmiany klucza m_{PW} podczas normalnej pracy sieci [28]. Problem ten można rozwiązać poprzez wprowadzenie dwóch kolejnych operacji kryptograficznych zdefiniowanych w węzłach, a mianowicie:

OW5. Zmiana wartości klucza pierwotnego węzła
$ZK\colon \{m_{PW}\}$.

OW6. Szyfrowanie kluczem m_{PW} klucza m
$SZFPW\colon \{m\} \to f_{m_{PW}}(m)$.

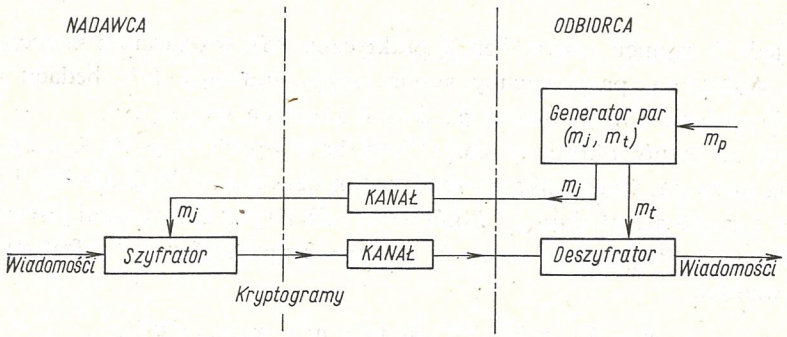
Jak można było zauważyć, organizacja i zarządzanie kluczami kryptograficznymi w sieci komputerowej ma charakter hierarchiczny [26]. Na pierwszym poziomie hierarchii znajduje się klucz pierwotny węzła m^i_{PW} oraz jego warianty m^i_{PW1}, m^i_{PW2}. Fizycznie dla i-tego węzła istnieje tylko klucz m^i_{PW}, natomiast jego warianty są tworzone poprzez przestawienie wybranych bitów klucza m^i_{PW}. Drugi poziom hierarchii jest tworzony przez klucze m^{ij}_{PP} oraz m^{ij}_T. Powyżej wymienione klucze są przechowywane w postaci zaszyfrowanej odpowiednią wersją klucza pierwotnego. I tak klucze m^{ij}_{PP} oraz klucze m^{ij}_T $(i \neq j)$ są przechowywane w i-tym węźle w odpowiednich postaciach $f_{m^i_{PW_1}}(m^{ij}_{PP})$ oraz $f_{m^i_{PW_1}}(m^{ij}_T)$, natomiast w j-tym węźle w formie $f_{m^j_{PW_2}}(m^{ij}_{PP})$ oraz $f_{m^j_{PW_2}}(m^{ij}_T)$. Dla przypadku, gdy $i = j$, wtedy m^{ii}_{PP} jest przechowywany w i-tym węźle za pomocą kryptogramu o postaci $f_{m^i_{PW_2}}(m^{ii}_{PP})$ oraz klucz m^{ii}_T — w formie $f_{m^i_{PW_2}}(m^{ii}_T)$. Trzeci i ostatni poziom zawiera wszystkie klucze robocze m_R, które są chronione podczas ich przechowywania w węzłach, odpowiednim kluczem z drugiego poziomu hierarchii. Oczywiście klucze m_R są bezpośrednio stosowane do ochrony wiadomości, w przypadku gdy ma miejsce transmisja pomiędzy końcówką a węzłem, lub do ochrony pakietów, jeżeli mamy do czynienia z transmisją pomiędzy węzłami.

6. PUBLICZNE SYSTEMY KRYPTOGRAFICZNE

W poprzednim rozdziale podano wielopoziomowy model organizacji i zarządzania kluczami kryptograficznymi w przypadku, gdy do ochrony informacji przechowywanych w sieci komputerowej stosuje się przekształcenie kryptograficzne. Jak można było zauważyć, w celu odpowiedniego zabezpieczenia przesyłanych i przechowywanych informacji konieczne jest zdefiniowanie kilku kategorii kluczy oraz niezbędne jest określenie operacji kryptograficznych do przeprowadzania odpowiednich przekształceń, zarówno wiadomości jak też kluczy. Tak więc problem kryptograficznego zabezpieczenia informacji w sieciach komputerowych wiąże się z rozwiązywaniem stosunkowo skomplikowanych problemów organizacji rozdziału i generowania kluczy kryptograficznych [12, 14, 43]. Problemy te upraszczają się poważnie w przypadku, gdy zabezpieczenie kryptograficzne informacji przesyłanych i przechowywanych w sieci komputerowej jest wprowadzane w oparciu o tzw. publiczne systemy kryptograficzne (*PSK*) [8, 29, 38]. Uproszczenie zarządzania kluczami kryptograficznymi w sieci komputerowej z *PSK* uzyskuje się dzięki możliwości przesyłania klucza kryptograficznego w jawnej postaci. W dalszym ciągu przedmiotem naszego zainteresowania będą wspomniane wyżej publiczne systemy kryptograficzne.

6.1. Określenie publicznych systemów kryptograficznych

Schemat *PSK* jest pokazany na rys. 11. System ten składa się z trzech podstawowych elementów: szyfratora, deszyfratora oraz generatora pary kluczy kryptograficznych



Rys. 11. Ilustracja graficzna publicznego systemu kryptograficznego

(m_j, m_t). Do wejść szyfratora podawany jest klucz kryptograficzny m_j oraz wiadomość x, przy czym klucz m_j jest przekazywany od odbiorcy do nadawcy w jawnej postaci. Na wyjściu szyfratora uzyskujemy kryptogram $y = f(x, m_j)$, gdzie $f(\cdot,*)$ jest funkcją kryptograficzną szyfratora. Następnie kryptogram y jest przekazywany do odbiorcy i tam na podstawie klucza m_t zostaje wydobyta wiadomość zgodnie z równością

$$x = g(y, m_t),$$

przy czym $g(\cdot,*)$ jest funkcją kryptograficzną deszyfratora, oraz m_t jest tajnym kluczem kryptograficznym znanym tylko po stronie odbiorczej ($m_j \neq m_t$). Wytworzenie pary

(m_j, m_t) następuje w generatorze par kluczy. W generatorze tym na podstawie klucza podstawowego m_p, nazywanego także warunkiem początkowym *PSK* [8], określana jest wartość m_j oraz m_t w taki sposób, aby była spełniona równość

$$\bigwedge_{x \in X} x = g(f(x, m_j), m_t).$$

Warto w tym miejscu zauważyć, że w klasycznych systemach kryptograficznych $m_j = m_t$ oraz $g(\cdot, *) \equiv f^{-1}(\cdot, *)$.

Z powyższych rozważań wynikają warunki, jakie musi spełniać *PSK*, a mianowicie:

1. funkcja kryptograficzna $f(\cdot, *)$ szyfratora powinna być zaprojektowana w taki sposób aby znalezienie kryptogramu na podstawie pary: wiadomość x, klucz m_j wymagało małej liczby elementarnych operacji przetwarzania, natomiast wydobycie wiadomości z kryptogramu y przy znajomości klucza m_j wiązało się z wykonaniem dużej liczby elementarnych operacji przetwarzania (np. 2^{64}),
2. funkcja kryptograficzna $g(\cdot, *)$ deszyfratora powinna pozwalać na szybkie określenie wiadomości na podstawie kryptogramu oraz klucza m_t,
3. określenie klucza m_t oraz warunku początkowego m_p przez *NU*, jeżeli zna on klucz m_j, powinno wymagać dużej liczby elementarnych operacji przetwarzania.

Ciągle jeszcze trwają badania nad sposobami konstrukcji publicznych systemów kryptograficznych. Pojawiły się już pierwsze prace na ten temat i poniżej przedstawimy jedną z koncepcji budowy *PSK* opracowaną przez Rivesta, Shamira i Adlemana [38].

6.2. Praktyczne rozwiązanie *PSK*

Zanim jednak zajmiemy się realizacją praktyczną *PSK* w oparciu o koncepcję Rivesta, Shamira i Adlemana, przedstawimy system kryptograficzny (nie będący publicznym systemem kryptograficznym) oparty na dwóch kluczach (m_j, m_t).

Załóżmy, że mamy określony zbiór liczb naturalnych $P = \{0, 1, \ldots, p-1\}$, gdzie p jest liczbą pierwszą, wraz z określonymi w nim działaniami dodawaniem i mnożeniem (dodawanie i mnożenie modulo p). Zbiór P łącznie z tymi działaniami jest ciałem skończonym $(P, +, \cdot)$. Przypuśćmy dalej, że funkcja kryptograficzna szyfratora $f(\cdot, *)$ jest określona wzorem

$$y = f(x, m_j) = x^{m_j} \text{ modulo } p; \quad x, y, m_j \in P, \tag{6.1}$$

przy czym m_j powinno być tak dobrane, aby funkcja kryptograficzna f była różnowartościowa ze względu na x, czyli liczby m_j oraz $(p-1)$ muszą być względnie pierwsze. Proces deszyfracji w tym systemie kryptograficznym polega na zastosowaniu funkcji kryptograficznej deszyfratora $g(\cdot, *)$ wyrażonej wzorem

$$x = g(y, m_t) = y^{m_t} \text{ modulo } p; \quad m_t \in P. \tag{6.2}$$

Biorąc pod uwagę równości określone wzorami (6.1) oraz (6.2) otrzymujemy

$$x^{m_j m_t} = x \text{ modulo } p.$$

Stąd wynika równość

$$x^{m_j m_t - 1} = 1 \text{ modulo } p.$$

Z twierdzenia Fermata wynika kongruencja $x^{p-1} = 1$ modulo p, zatem ze wzoru (6.3) mamy

$$m_j m_t = 1 \text{ modulo } p-1. \tag{6.4}$$

A więc znając m_j można określić w prosty sposób wartość m_t zgodnie z równością (6.4). Przedstawiony system nie jest publicznym systemem kryptograficznym, gdyż znajomość klucza m_j pozwala na bezpośrednie określenie klucza m_t (budowa i zasada pracy *PSK* mogą być znane przez *NU*). Niemniej stanowi on podstawę do budowy *PSK* określonego przez Rivesta et al.

Cytowani powyżej autorzy podali sposób budowy *PSK* w oparciu o pierścień skończony $(\{0, 1, \ldots, r-1\}, +, \cdot)$, gdzie dodawanie i mnożenie jest określone w sensie działań modulo r. Podobnie jak poprzednio przyjmujemy, że funkcje kryptograficzne szyfratora i deszyfratora są określone zgodnie z poniższymi równościami

$$f(x, m_j) = x^{m_j} \text{ modulo } r \tag{6.5}$$

oraz

$$g(y, m_t) = y^{m_t} \text{ modulo } r, \tag{6.6}$$

gdzie $x, y, m_j, m_t \in \{0, 1, \ldots, r-1\}$.

Powiązania pomiędzy kluczami m_j oraz m_t są bardziej złożone niż te, które występują w powyżej przedstawionym systemie kryptograficznym z dwoma kluczami kryptograficznymi. W celu wyznaczenia tych powiązań wybieramy dwie liczby pierwsze p oraz q ($p, q \in \{0, 1, \ldots, r-1\}$), przy czym $pq = r$. Oczywiście liczba r jest publicznie znana, natomiast liczby p oraz q nie są znane dla *NU* i tworzą tzw. warunek początkowy (klucz m_p). Teraz dokonamy wyboru pary kluczy m_j oraz m_t w taki sposób, aby znajomość klucza jawnego m_j nie pozwalała na określenie wartości klucza tajnego m_t. Wybieramy zatem jako pierwszy klucz m_t. Ponieważ funkcja kryptograficzna deszyfratora powinna być funkcją różnowartościową ze względu na zmienną y, więc określamy liczbę m_t w taki sposób, aby była względnie pierwsza w stosunku do liczb $(p-1)$ oraz $(q-1)$. Z twierdzenia Carmichaela wynika, że dla każdej grupy multiplikatywnej istnieje maksymalny okres $\lambda(r)$, dla którego słuszna jest kongruencja

$$\bigwedge_{a \in \{0,1,\ldots,r-1\}} a^{\lambda(r)} = 1 \text{ modulo } r \tag{6.7}$$

oraz

$$\lambda(r) = NMWW(p-1, q-1), \tag{6.8}$$

przy czym *NMWW* — najmniejsza wspólna wielokrotność.

Biorąc pod uwagę równości (6.5) oraz (6.6) mamy

$$x^{m_j m_t - 1} = 1 \text{ modulo } r. \tag{6.9}$$

Porównując wzory (6.7) oraz (6.9) otrzymujemy

$$m_j m_t - 1 = 0 \text{ modulo } \lambda(r),$$

czyli

$$m_j m_t = 1 \text{ modulo } NMWW(p-1, q-1). \tag{6.10}$$

Z kongruencji podanej wzorem (6.10) można określić wartość klucza m_j, gdy znana jest liczba m_t. Z poprzednich określeń wynika, że *NU* w celu określenia wartości przesyłanej wiadomości powinien znać warunek początkowy (klucz podstawowy), który w tym systemie jest tworzony przez parę liczb (p, q), gdyż wtedy może łatwo określić wartość klucza

tajnego z kongruencji (6.10). W przypadku gdy *NU* zna kryptogram i klucz jawny oraz zna budowę *PSK*, to wtedy stoi on przed zadaniem określenia na podstawie liczby r warunku początkowego. Jest to problem wymagający dużej liczby elementarnych operacji przetwarzania dla odpowiednio dużej wartości liczby r.

Na uwagę zasługuje również inna koncepcja konstrukcji *PSK* przedstawiona przez Merklego oraz Hellmana [29]. Budowa takiego *PSK* jest oparta na określeniu odpowiedniego zadania numerycznego znanego w literaturze jako „problem plecaka" (ang. *knapsack problem*).

Dla obu omawianych koncepcji problem określenia wiadomości na podstawie kryptogramu i klucza jawnego wiąże się z rozwiązaniem złożonego zadania numerycznego.

7. OCENA JAKOŚCI KRYPTOGRAFICZNYCH METOD ZABEZPIECZENIA INFORMACJI

Taką ocenę można przeprowadzić dla trzech przypadków.

Przypadek 1. Jeżeli stosowany jest klasyczny system kryptograficzny oraz *NU* ma dostęp zarówno do wiadomości jak też kryptogramów. Wtedy kryterium podstawowym jakości stosowanego przekształcenia kryptograficznego jest średnie prawdopodobieństwo udanej, nielegalnej wymiany wiadomości przy warunku, że *NU* zna n różnych par obserwacji (wiadomość, kryptogram). Czyli kryterium podstawowym jest

$\overline{P} = \overline{P}$ (udanej, nielegalnej wymiany wiadomości | *NU* zna n par wiadomość, kryptogram).

Natomiast kryteriami wtórnymi wynikającymi z kryterium podstawowego są:

— liczba wymaganych operacji elementarnych przetwarzania, dla której $\overline{P} = 1$, czyli $L(\overline{P} = 1)$,

— koszt określenia kryptogramu dla nielegalnej wiadomości,

— czas określenia kryptogramu dla nielegalnej wiadomości,

— koszt określenia klucza kryptograficznego przez *NU*,

— czas wymagany do określenia klucza kryptograficznego przez *NU*.

Przypadek 2. Jeżeli stosowany jest klasyczny system kryptograficzny oraz *NU* ma dostęp do kanału (zna kryptogramy). Kryteriami podstawowymi są wtedy:

— średnie prawdopodobieństwo określenia klucza kryptograficznego przy warunku, że *NU* zna n różnych kryptogramów

$\overline{P}' = \overline{P}(\underset{\sim}{M}^* = m_k \mid$ *NU* zna n kryptogramów),

— średnie prawdopodobieństwo określenia przez *NU* wiadomości przy warunku, że zna on n różnych kryptogramów

$\overline{P}'' = \overline{P}(\underset{\sim}{X}^* = x_i \mid$ *NU* zna n kryptogramów).

Jako kryteria wtórne można przyjąć:

— liczby elementarnych operacji przetwarzania, dla których prawdopodobieństwa $\overline{P}'$ oraz $\overline{P}''$ są równe jeden, czyli

$$L(\overline{P}' = 1) \quad \text{oraz} \quad L(\overline{P}'' = 1),$$

— koszt oraz czas potrzebny do określenia przez *NU* klucza kryptograficznego lub wiadomości.

Przypadek 3. Jeżeli jest stosowany publiczny system kryptograficzny. Kryterium podstawowym są:

— średnie prawdopodobieństwo określenia wiadomości przez *NU*, gdy zna on kryptogram i klucz jawny, a więc
$\bar{P}''' = \bar{P}(\underset{\sim}{X}^* = x_i \mid NU$ zna kryptogram i klucz jawny),

— średnie prawdopodobieństwo określenia klucza tajnego (warunku początkowego) przez *NU* przy założeniu, że zna on klucz jawny
$\bar{P}^{IV} = \bar{P}(\underset{\sim}{M}^* = m_t \mid NU$ zna klucz jawny).

Natomiast kryteriami wtórnymi są:

— liczby operacji $L(\bar{P}''' = 1)$ oraz $L(\bar{P}^{IV} = 1)$,

— koszt i czas potrzebny do określenia wiadomości lub klucza tajnego przez *NU*, gdy zna on kryptogram i klucz jawny.

W ogólnym przypadku wchodzą w grę jeszcze dodatkowe kryteria, które nie były poprzednio rozważane, takie jak:

— opóźnienie transmisji wiadomości powstałe wskutek wprowadzenia przekształcenia kryptograficznego,

— koszt budowy odpowiednich szyfratorów i deszyfratorów,

— liczba komórek pamięci niezbędnych do pamiętania kluczy kryptograficznych oraz programów zarządzających tymi kluczami (to kryterium jest szczególnie ważne w sieciach komputerowych).

8. WNIOSKI KOŃCOWE

W pracy przedstawiono przegląd kryptograficznych metod zabezpieczenia informacji przetwarzanych i przechowywanych w systemach i sieciach komputerowych. Metody te są stosowane do ochrony informacji zarówno przed podstawieniem nielegalnej wiadomości, jak też przed odczytaniem wiadomości z kryptogramu przez nieautoryzowanego użytkownika. Upowszechnienie zastosowań systemów komputerowych spowodowało rozwój klasycznych systemów kryptograficznych, a w szczególności opracowanie algorytmów szyfrowania blokowego, na podstawie których można zarówno budować szyfratory realizujące rekurencyjne przekształcenia kryptograficzne, jak też uzyskiwać autentyfikatory dla wiadomości przesyłanych w jawnej postaci. Mimo faktu, że opracowano już kilka takich algorytmów [9, 34, 42], to w dalszym ciągu nie są znane praktyczne kryteria oceny jakości tych algorytmów (kryteria podane w punkcie 7 są zbyt ogólne). Obecnie prowadzone są badania w celu określenia takich kryteriów [17, 34]. Myślą przewodnią w tych badaniach jest fakt, że każdy algorytm szyfrowania powinien zachowywać się (przy ustalonej wartości klucza kryptograficznego) jak generator liczb losowych. Z tego wynika, że ocena jakości algorytmów powinna opierać się na badaniach własności statystycznych uzyskiwanych kryptogramów.

Kolejnym elementem stymulującym rozwój kryptografii było pojawienie się sieci komputerowych. Jednakże zastosowanie przekształceń kryptograficznych do ochrony informacji w sieciach komputerowych spowodowało pojawienie się wielu nowych problemów. Problemy te są w szczególności związane z generowaniem, zarządzaniem oraz

okresowymi zmianami kluczy kryptograficznych [12, 26, 28, 19]. Przy tym bardzo ważną sprawą jest takie zaprojektowanie ochrony kryptograficznej, aby przypadkowa zmiana wartości klucza kryptograficznego pierwotnego jednego z komputerów sieci nie powodowała utraty całości lub części informacji przetwarzanych lub przechowywanych w tym komputerze. Problemy te poważnie upraszczają się, gdy do ochrony informacji w sieciach komputerowych stosowane są publiczne systemy kryptograficzne. Niestety opracowane dotąd publiczne systemy kryptograficzne wprowadzają duże opóźnienie przesyłanych wiadomości, co w poważnym stopniu ogranicza zasięg ich stosowania. Dlatego w dalszym ciągu prowadzone są intensywne badania nad opracowaniem takich publicznych systemów kryptograficznych, które zapewniałyby stosunkowo dobrą ochronę informacji przy jednocześnie małym opóźnieniu w transmisji wiadomości. Na zakończenie warto zauważyć, że jak dotąd nie opracowano matematycznego modelu zabezpieczenia kryptograficznego stosowanego w sieci komputerowej, w oparciu o który byłoby możliwe określenie globalnego wskaźnika jakości użytych metod kryptograficznych w sieci komputerowej.

LITERATURA

1. R. Bayer, J.K. Metzger, *On the Encipherment of Search Trees and Random Access Files*, ACM Transactions on Database Systems, vol. 1, No. 1, March 1976, pp. 37—52.
2. D.K. Branstad, *Encryption Protection in Computer Data Communications*, NBS Report, 1975.
3. H.S. Bright, R.L. Enison, *Cryptography Using Modular Software Elements*, AFIPS Press, National Computer Conference, 1976, pp. 113—123.
4. P.S. Browne, *Computer Security — a Survey*, Proceedings National Computer Conference, 1976, pp. 53—62.
5. H.R. Burris, *Computer Network Cryptography Engineering*, Proceedings National Computer Conference, 1976, pp. 91—96.
6. G.D. Cole, *Design Alternatives for Computer Network Security*, NBS Special Publication 500—21, vol. 1, January 1978.
7. D.W. Davies, *New Techniques for Cryptography in Networks*, Proceedings of Eurocomp, 1978.
8. D.W. Davies, W.L. Price, G.I. Parkin, *An Evaluation of Public Key Cryptosystems*, National Physical Laboratory Report CTU 1, March 1979.
9. *Data Encryption Standard*, Federal Information Processing Standard, Publication 46, National Bureau of Standards, January 1977.
10. W. Diffie, M.E. Hellman, *Multiuser Cryptographic Techniques*, Proceegings of National Computer Conference, vol. 45, June 1976, pp. 109—112.
11. W. Diffie, M.E. Hellman, *New Directions in Cryptography*, IEEE Transactions on Information Theory, vol. IT-22, No. 6, November 1976, pp. 644—654.
12. W.F. Ehrsam, S.M. Matyas, C.H. Meyer, W.L. Tuchman, *A Cryptographic Key Management Scheme for Implementing the Data Encryption Standard*, IBM Systems Journal, vol. 17, No. 2, 1978, pp. 106—126.
13. N. Enticknap, *Should You Bet Your Company*?, Data Processing, July-August 1976, pp. 41—42.
14. J.K. Everton, *A Hierarchical Basis for Encryption Key Management in a Computer Communication Network*, SPERRY-RAND Corporation Report, 1978.
15. H. Feistel, *Cryptographic Coding for Data Bank Privacy*, IBM Research Report RC-2827, March 1970.
16. H. Feistel, *Cryptography and Computer Privacy*, Scientific American, vol. 228, No. 5, May 1973.
17. J. Gait, *A New Nonlinear Pseudorandom Number Generator*, IEEE Transactions on Software Engineering, vol. SE-3, No. 5, September 1977, pp. 359—363.

18. E. N. Gilbert, F. J. MacWilliams, N. J. A. Sloane, *Codes which Detect Deception*, BSTJ, vol. 53, No. 3, March 1974, pp. 405—424.
19. H. M. Gladney, *Administrative Control of Computing Service*, IBM Systems Journal, vol. 17, No. 2, 1978, pp. 151—179.
20. E. Gudes, H. S. Koch, F. A. Stahl, *The Application of Cryptography for Data Base Security*, National Computer Conference, 1976, pp. 97—107.
21. M. A. Harrison, W. L. Ruzzo, J. D. Ullman, *Protection in Operating Systems*, Communications of the ACM, vol. 19, No. 8, August 1976, pp. 461—471.
22. D. K. Hsiao, D. S. Kerr, S. E. Madnick, *Computer Security — Problems and Solutions*, The Ohio State University, March 1978.
23. D. Kaufman, K. Auerbach, *A Secure National System for Electronic Funds Transfer*, National Computer Conference, 1976, pp. 129—138.
24. S. T. Kent, *Encryption-based Protection for Interactive User/Computer Communication*, 5th Data Communications Symposium, September 1977, pp. 7—13.
25. B. W. Lampson, *Dynamic Protection Structures*, Proceedings Fall Joint Computer Conference, 1969, pp. 27—38.
26. R. E. Lennon, *Cryptography Architecture for Information Security*, IBM Systems Journal, vol. 17, No. 2, 1978, pp. 138—151.
27. R. J. Lipton, L. Snyder, *A Linear Time Algorithm for Deciding Subject Security*, Journal of the Association for Computing Machinery, vol. 24, No. 3, July 1977, pp. 455—464.
28. S. M. Matyas, C. H. Meyer, *Generation, Distribution, and Installation of Cryptographic Keys*, IBM Systems Journal, vol. 17, No. 2, 1978, pp. 126—138.
29. R. C. Merkle, M. E. Hellman, *Hiding Information and Signatures in Trap-Door Knapsacks*, IEEE Transactions on Information Theory, IT-24, September 1978, pp. 525—530.
30. N. Minsky, *Cooperative Authorization in Computer Systems*, Rutgers Technical Report, Department of Computer Science, Rutgers University, April 1978.
31. R. M. Needham, M. D. Schroeder, *Using Encryption for Authentication in Large Networks of Computers*, Communications of the ACM, vol. 21, No. 12, December 1978, pp. 993—999.
32. M. J. Orceyre, R. H. Courtney Jr, *Considerations in the Selection of Security Measures for Automatic Data Processing Systems*, NBS Special Publication 500—33, June 1978.
33. D. B. Parker, *Computer Abuse Perpetrators and Vulnerabilities of Computer Systems*, National Computer Conference, 1976, pp. 65—73.
34. J. Pieprzyk, *Zastosowanie metod kryptograficznych do zabezpieczenia danych w systemach komputerowych*, Praca doktorska, Instytut Podstaw Informatyki PAN, Warszawa 1979.
35. J. Pieprzyk, *Uogólnienie szyfrów Vigenera*, Archiwum Automatyki i Telemechaniki, No. 4, 1979, s. 454—469.
36. J. Pieprzyk, D. Rutkowski, *Zastosowanie kryptografii do zabezpieczenia informacji w sieciach komputerowych*, I Konferencja Naukowa nt. Systemy Telekomunikacyjne, Kołobrzeg, 21—23 kwietnia 1980, s. 148—152.
37. V. S. Pless, *Encryption Schemes for Computer Confidentiality*, IEEE Transactions on Computers, vol. C-26, No. 11, November 1977, pp. 1133—1136.
38. R. L. Rivest, A. Shamir, L. Adleman, *A Method for Obtaining Digital Signatures and Public-Key Cryptosystems*, Communications of the ACM, vol. 21, No. 2, February 1978, pp. 120—126.
39. J. H. Saltzer, M. D. Schroeder, *The Protection of Information in Computer Systems*, Proceedings of the IEEE, vol. 63, No. 9, September 1975, pp. 1278—1308.
40. J. Seidler, *Teoria kodów*, PWN, Warszawa 1965.
41. C. E. Shannon, *Communications Theory of Secrecy Systems*, BSTJ, vol. 28, No. 4, October 1949, pp. 656—715.
42. J. L. Smith, W. A. Notz, P. R. Osseck, *An Experimental Application of Cryptography to a Remotely Accessed Data System*, Proceedings of the ACM, 1972, pp. 282—297.
43. D. J. Sykes, *The Management of Encryption Keys*, Honeywell Information Systems Report, May 1979.
44. H. Weiss, *Computer Security an Overview*, Datamation, vol. 20, No. 1, January 1974, pp. 42—47.

45. C. Wood, E.B. Fernandez, *Decentralized Authorization in a Database System*, Proceedings of 5th International Conference on Very Large Data Bases, 1979, pp. 352—360.
46. H. M. Wood, *The Use of Password for Controlling Access to Remote Computer Systems and Services*, National Computer Conference, 1977, pp. 465—478.

J. PIEPRZYK, D. RUTKOWSKI

CRYPTOGRAPHIC METHODS OF INFORMATION PROTECTION

Summary

The work is a review of cryptographic methods, which are applied in computer systems and networks for the purpose of information protection. Although cryptographic methods have been known and applied since antiquity, recently definite animation is observed in research, and notable development of cryptography theory has followed. For that reason, the presentation of the latest trends in cryptography is stressed in this work. For this purpose the classification is given of cryptographic methods, together with its short description. Then two protection methods are presented of information transmitted between the terminal and computer. Next are discussed: enciphering algorithms, both the DES algorithm and the algorithm constructed on the basis enciphering modules. One can regard enciphering algorithms as basic elements used to design of complex cryptosystems, e.g. the stream cipher. Also it is possible to bring in the information protection of computer network with the aid of enciphering algorithms. Such an idea has been presented in the work, and the key management problem has been discussed. Finally a method of designing public key cryptosystem is shown. These systems are particularly useful for ensuring cryptographic protection of information in computer networks.

J. PIEPRZYK, D. RUTKOWSKI

MÉTHODES CRYPTOGRAPHIQUES DE LA PROTECTION DE L'INFORMATION

Résumé

Dans l'étude on passe en revue les méthodes cryptographiques de la protection de l'information, appliquées dans les systèmes et les réseaux d'un ordinateur. Quoique les méthodes cryptographiques aient été déjà connues dans l'antiquité, ce n'est que dans les dernières années qu'on a pu observer un intérêt plus poussé vers les recherches, ce qui a attiré, par la suite, un important développement des théories de cryptographie. C'est pourquoi, l'étude s'est posé pour but de présenter les tendances les plus récentes qui se manifestent dans le développement de la cryptographie. Après avoir classé les méthodes de la protection cryptographique et les ayant décrites, on a présenté deux méthodes de la protection des informations transmises entre le terminal et le système de l'ordinateur. Ensuite on a décrit les algorithmes de chiffrement, à savoir l'algorithme américain DES et l'algorithme basé sur les modules de chiffrement. Les algorithmes de chiffrement peuvent être traités comme éléments fondamentaux avec lesquels on peut construire des systèmes cryptographiques composés, tels que systèmes cryptographiques de récurrence. La protection de l'information dans un réseau de l'ordinateur peut être aussi assurée à l'aide d'un algorithme de chiffrement. Cette dernière conception a été présentée dans l'étude, et on a discuté les problèmes se rapportant à la disposition des clés cryptographiques. Vers la fin l'étude traite le problème fort intéressant de la projection des systèmes cryptografiques publiques, qui sont spécialement utilés pour la protection cryptographique de l'information dans les réseaux de l'ordinateur.

J. PIEPRZYK, D. RUTKOWSKI

KRYPTOGRAPHISCHE METHODEN FÜR INFORMATIONSSICHERUNG

Zusammenfassung

Die Problematik der Bearbeitung betrifft die Übersicht über kryptographische Sicherungsmethoden für Informationen, die bei Computersystemen und -netzen angewandt werden. Obwohl kryptographische Methoden bereits im Altertum gebraucht worden waren, so kann doch erst in den letzten Jahren eine merkliche Belebung bei der Aufnahme diesbezüglicher Forschungen beobachtet werden, und demzufolge eine erhebliche Entwicklung der kryptographischen Theorie. Deswegen wurde auch hier besonderes Interesse den neuesten Tendenzen in der Entwicklung der Kryptographie geschenkt. Nach Angabe der Klassifikation kryptographischer Sicherungsmethoden und deren kurzer Beschreibung wurden zwei Schutzmethoden präsentiert für die vom Endglied ins Computersystem zu übertragenden Informationen. Im weiteren wurden Chiffrealgorithmen diskutiert, und zwar der amerikanische DES-Algorithmus und der auf chiffrierenden Modulen basierende Algorithmus. Die Chiffrealgorithmen können als Grundelemente behandelt werden, woraus zusammengesetzte kryptographische Systeme gebaut werden können, wie z.B. rekurrierende kryptographische Systeme. Man kann auch den Informationsschutz im Computernetz in Anlehnung an die Chiffrealgorithmen einführen. Derartige Konzeption wurde in vorliegender Bearbeitung dargestellt, auch wurden Probleme diskutiert, die sich aus der Verwaltung von kryptographischen Schlüsseln ergeben. Zum Schluß werden interessante Probleme der Projektierung öffentlicher kryptographischer Systeme berührt, die sich besonders für die Ausarbeitung kryptographischer Informationssicherung in Computernetzen eignen.

Ю. ПЕПШИК, Д. РУТКОВСКИ

КРИПТОГРАФИЧЕСКИЕ МЕТОДЫ ОХРАНЫ ИНФОРМАЦИИ

Резюме

Приведен обзор криптографических методов охраны информации применяемых в компютерных системах и сетях. Не смотря на то, что криптографические методы применялись с древних времен, заметное оживление интереса к исследовательским работам в этой области проявилось лишь только в последнее время, чего следствием явилось значитеьное развитие теории криптографии. По этому поводу асобое влимание уделено новейшим тенденциям развития криптографии. После склассифицирования методов криптографической охраны и их краткого обсуждения предствавлены два метода охраны информации передаваемых между терминалем и компютерной системой. Далее обсуждены алгоритмы шифовки, а в частности американсний алгоритм ДЕС и алгоритм обоснованный на шифрующих модулях. Алгоритмы шифрования можно рассматривать как основные элементы из которых можно составлять сложные криптографические системы, как на пример рекуренционные криптографические системы. Также и охрану информации в компютерной сети можно реализировать опираясь на алгоритмы шифрования. Такая концепция представлена в труде с обсуждением проблем управления криптографическими ключами. В заключение приведена интересная проблема проектирования общественных криптографических систем, особо пригодных для систем криптографической охраны информации в компютерных сетях.

ROZPRAWY ELEKTROTECHNICZNE 1982, 28, z. 1-2, ss. 315-325

Metoda uśredniania i rozpoznawania sygnałów biomedycznych

ZBIGNIEW M. WÓJCIK (WARSZAWA)

Centrum Badań Kosmicznych Polskiej Akademii Nauk

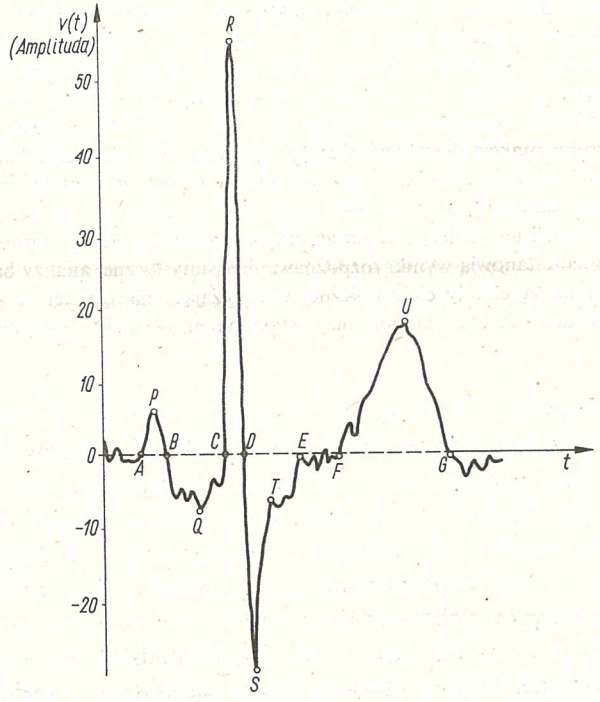
Otrzymano 13.3.1980

Praca przedstawia metodę i układ elektroniczny dla automatycznego uśredniania próbkowanego sygnału analogowego, jak również metodę i układy elektroniczne dla rozpoznawania tych uśrednionych sygnałów. Uśrednianie polega na wyznaczaniu pewnej średniej ważonej wyników próbkowania każdego punktu pomiarowego badanego sygnału, rozłożonych symetrycznie względem tego punktu. Dzięki zastosowaniu wag, wpływ sąsiednich wartości mierzonego sygnału na uśrednioną (wynikową) wartość badanego punktu pomiarowego maleje w miarę oddalania się od tego punktu. Wyniki uśredniania są zgodne z podstawowym warunkiem reprezentowania zdarzeń [9, 4, 11, 12, 7]. W procesie rozpoznawania badane są wartości funkcji charakterystycznych wzrostu, malenia i zera badanego sygnału [13, 8]. Wartości jeden tych funkcji stanowią alfabet pewnej gramatyki formalnej. Cechy badanego sygnału stanowią zbiór symboli końcowych tej gramatyki, otrzymywanych w wyniku zbioru produkcji tej gramatyki. Dla gramatyki tej istnieje pewien automat. Stany końcowe tego automatu stanowią wyniki rozpoznawania i statystycznej analizy badanych sygnałów. Przedstawione są układy elektroniczne, wyznaczające najbardziej podstawowe cechy sygnałów. Przedstawiona metoda opisana została na przykładzie rozpoznawania elektrokardiogramów.

Terminologia, stosowana w pracy, odnosi się do tematyki rozpoznawania dowolnych sygnałów analogowych, (np. sygnału mowy [13, 8]), choć w dalszym ciągu artykułu omawiany będzie konkretny przykład analizy elektrokardiogramów. Wyniki tej pracy mogą być bowiem stosowane również daleko poza zagadnieniami dotyczącymi elektrokardiogramów. Pewnych wyjaśnień wymaga termin „rozpoznawanie". Zgodnie z definicją podaną w pracach [4, 7, 9, 11, 12], rozpoznawanie jest procesem polegającym na transformacji badanego zdarzenia (obrazu, elektrokardiogramu) w zbiór symboli, niezmiennych w przestrzeni obserwacji tego zdarzenia (tzn. spełniających podstawowy warunek reprezentowania zdarzeń). Symbole, będące słowami języka naturalnego, spełniają ten warunek. Dlatego ciąg symboli typu „wzrost — pik dodatni — malenie — zero — pik ujemny — zero", będący wynikiem pewnego procesu przetwarzania sygnału elektronicznego, stanowi wynik rozpoznawania tego sygnału, pod warunkiem zgodności wyników zastosowanych algorytmów z intuicyjnym znaczeniem słów: wzrost, pik, malenie, zero. Interpretacja rozpoznanych sygnałów polega natomiast na porównaniu wyników rozpoznawania (np. ciągu wyrazów języka naturalnego wraz z wartościami liczbowymi intensywności cech, reprezentowanych poszczególnymi słowami) z wzorcami. Wynika stąd, że proces rozpoznawania elektrokardiogramu może być przeprowadzony przez elektronika, będącego

laikiem w dziedzinach biomedycyny, natomiast interpretacja elektrokardiogramu musi być dokonana przez właściwego specjalistę — biomedyka, który potrafi skonstruować właściwe wzorce dla prawidłowych i nieprawidłowych przebiegów, albo przez układ elektroniczny, zbudowany zgodnie z jego wskazówkami. Artykuł ten nie dotyczy procesów interpretacji sygnałów, ale kończy się na procesie ich rozpoznawania.

Diagnozę arytmii serca ułatwiają lekarzowi wyniki precyzyjnych pomiarów elektrokardiogramów, wykonane przez wyspecjalizowany przyrząd. Rysunek 1 przedstawia przykład elektrokardiogramu, wykreślonego w układzie odniesienia czas-amplituda. Na wykresie zaznaczono charakterystyczne punkty, tzw. załamki, brane pod uwagę przy pomiarach. Ponieważ czujnik przyrządu może być w różny sposób przymocowany do ciała osoby badanej, jak również każdy człowiek posiada pewną charakterystyczną budowę ciała, bezwzględne wartości mierzonej amplitudy są mało istotne dla ustalenia diagnozy. Analizowane są przede wszystkim względne wysokości pików (tj. ekstremów) elektrokardiogramu, ich orientacja (tzn. czy wyróżnione na rys. 1 piki mają amplitudy dodatnie,

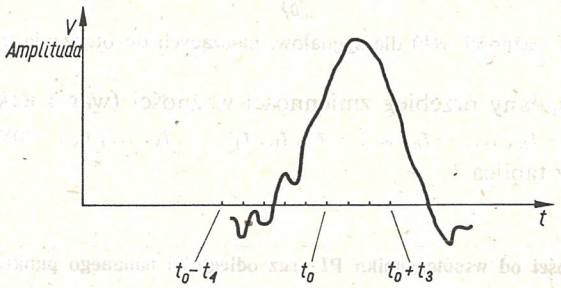


Rys. 1. Powiększony, pojedynczy cykl elektrokardiogramu, uśredniony dwunastokrotnie [1]

czy ujemne). Mierzone są szerokości pików oraz ich powierzchnie. Badana jest symetria pików oraz ich dokładna pozycja (tzn. ich położenie na osi czasu). Zmierzone wielkości są porównywane z wartościami, uznanymi jako prawidłowe oraz z wielkościami, zmierzonymi wcześniej, z tego samego elektrokardiogramu. Automatyczna analiza elektrokardiogramów jest złożona, dlatego przedstawione zostaną dalej tylko jej podstawowe fragmenty.

Wejściowy sygnał, otrzymany z czujnika, jest silnie zakłócony, głównie przez pracę rozmaitych organów badanej osoby. Pierwszym zabiegiem, pozwalającym znacznie zmniejszyć te zakłócenia, jest uśrednianie przebiegu. Zwykłe całkowanie sygnału przez filtr analogowy ma ograniczone zastosowanie, ponieważ dla dowolnej chwili czasu t_o brany jest pod uwagę jedynie przebieg, jaki był do tej chwili t_o. Na wyniki uśredniania nie ma natomiast wpływu sygnał, jaki miał miejsce po tym czasie t_o. Uśrednianie takie nie jest więc symetryczne względem tej chwili t_o i jest dlatego sprzeczne z podstawowym warunkiem reprezentowania zdarzeń [7, 9]. Całkowanie sygnału przez klasyczny, analogowy układ całkujący prowadzi np. do zmiany pozycji pików na osi czasu; piki te są przesuwane w prawo. W przypadku, gdy piki nie są równe i symetryczne (co jest normalne w przypadku elektrokardiogramów), są one przesuwane o różną wartość (wielkość przesunięcia zależy od powierzchni piku do momentu t_o). Prawidłowego uśrednienia sygnału można dokonać zgodnie z następującym wzorem

$$v'(t_o) = \frac{w_o v(t_o) + \sum_{k=1}^{K} w_k \left(v(t_o - t_k) + v(t_o + t_k)\right)}{w_o + 2 \sum_{k=1}^{K} w_k}, \tag{1}$$



Rys. 2. Graficzna interpretacja dyskretnych chwil t_o, $t_o - t_4$ oraz $t_o + t_3$ na osi czasu

gdzie:

k — indeks chwili próbkowania sygnału, otrzymanego z czujnika, liczony od chwili próbkowania t_o w obu kierunkach osi czasu (rys. 2),

$2K$ — liczba próbkowań sygnału, uwzględniana przy obliczaniu uśrednionego sygnału $\{v'(t_o), v'(t_1), v'(t_3), \ldots, v'(t_j), \ldots\}$; wartość K należy dobrać doświadczalnie;

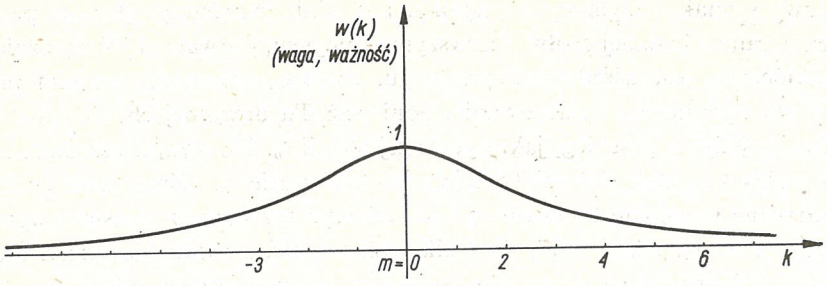
K, k — liczby całkowite dodatnie;

w_o — ważność sygnału $v(t_o)$; $w_o = 1$,

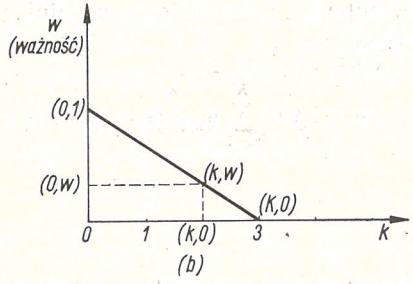
w_k — ważność sygnału $v(t_o - t_k)$ lub sygnału $v(t_o + t_k)$, otrzymanego z czujnika, odległego od chwili t_o o czas t_k (rys. 2); poszczególne ważności $\{w_1, w_2, \ldots, w_k, \ldots, w_K\}$ można dobrać zgodnie ze wzorami (2) i (3) (rys. 3 i 4)

$$w_k = w(k) = \exp(-PI \cdot k^2), \tag{2}$$

gdzie PI jest parametrem, dobieranym dla danego układu pomiarowego i dla każdego parametru K oddzielnie; dla dużego k wartość stałej PI jest odpowiednio mała, tak aby



Rys. 3. Wykres Gaussa ważności $w(k)$ dla sygnałów $\{v(t_o-t_K), \ldots, v(t_o-t_k), \ldots, v(t_o-t_1), v(t_o), v(t_o+t_1), \ldots, v(t_o+t_k), \ldots, v(t_o+t_K)\}$, uwzględnianych przy obliczaniu uśrednionego sygnału $v'(t_o)$



Rys. 4. Liniowy wykres ważności $w(k)$ dla sygnałów, należących do otoczenia $0(t)$ sygnału $v(t)$

był zapewniony pożądany przebieg zmienności ważności (wagi) $w(k)$ w całym badanym otoczeniu $0(t) = \{-t_K, \ldots, -t_k, \ldots, -t_1, t_0, t_1, \ldots, t_k, \ldots, t_K\}$ chwili próbkowania t_0 sygnału $v(t)$ — patrz tablica 1.

Tablica 1

Tablica wag w zależności od współczynnika PI oraz odległości badanego punktu od punktu centralnego

k	PI	W
1	0,18	0,835
2	0,18	0,487
3	0,18	0,198
4	0,18	0,056
5	0,18	0,011
1	0,3	0,741
2	0,3	0,301
3	0,3	0,067
4	0,3	0,008
1	0,6	0,549
2	0,6	0,091
3	0,6	0,005
1	1	0,368
2	1	0,018

Powyższy wzór można otrzymać z rozkładu Gaussa

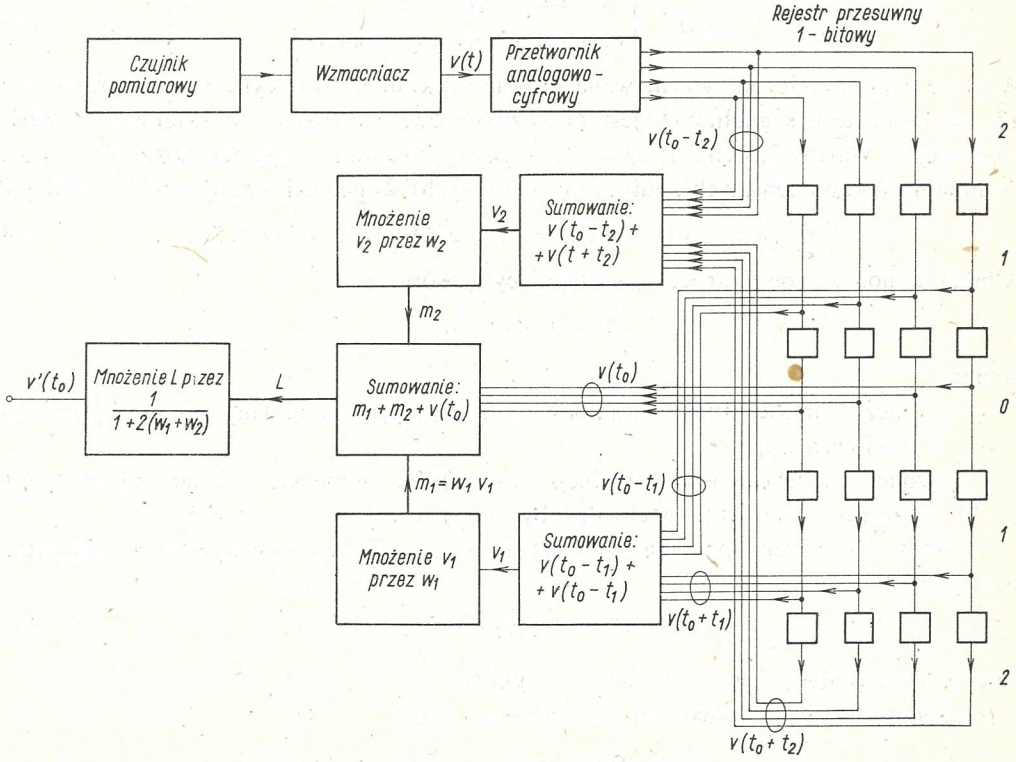
$$w(k) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(k-m)^2}{2\sigma^2}\right),$$

podstawiając $w(0) = w(m) = 1 = w_o$. Dla warunku tego odchylenie standardowe $\sigma = 1/\sqrt{2\pi}$. Ponieważ z podstawowego warunku reprezentowania zdarzeń wynika, że badane otoczenie $0(t)$ musi być symetryczne, $m = 0$ i w związku z tym otrzymujemy wzór (2), przy czym π jest traktowane jako parametr *PI*. Ważność (waga) $w(k)$ może być także funkcją liniową (rys. 4). Przyjmujemy wtedy, że $w(K) = w(-K) = 0$, tzn. że wpływ (waga) ostatniego elementu t_K (lub $-t_K$) badanego otoczenia $0(t)$ na uśrednioną wartość $v'(t_0)$ analizowanego przebiegu jest równy zeru. Wpływ ten rośnie liniowo w miarę zbliżania się do elementu centralnego t_0 i $w(0) = 1$ dla tego elementu t_0. Z rys. 4 wynika, że $w(k)$ wynosi

$$w(k) = \frac{K-r}{K}. \tag{3}$$

W sposób analogiczny do wyrażenia (1) można zapisać wzór na dowolną inną, uśrednioną wartość $v'(t_j)$ sygnału, otrzymaną na wyjściu czujnika.

Wyspecjalizowany układ elektroniczny symetrycznego uśredniania sygnału $v(t)$ jest przedstawiony na rys. 5. Sygnał $v(t)$, otrzymany z czujnika za pomocą klasycznego prze-



Rys. 5. Układ automatycznego uśredniania sygnału $v(t)$, uwzględniający podstawowy warunek reprezentowania zdarzeń

twornika analogowo-cyfrowego, jest kodowany za pomocą czterech bitów i jest przekazywany na wejścia czterech szeregowych rejestrów przesuwnych, zbudowanych z pojedynczych rejestrów jednobitowych. Na wyjściach i wejściach tych rejestrów są otrzymywane sygnały $v(t_o+t_k)$, $v(t_o)$ oraz $v(t_o-t_k)$, $k = 1, 2$. Waga (ważność) sygnału $v(t_o)$ wynosi 1. Cały układ może pracować w czasie rzeczywistym. Bity w rejestrach są przesuwane generatorem, sterującym pracą przetwornika analogowo-cyfrowego. Otrzymany, uśredniony prawidłowo sygnał

$$\{v'(t_1), v'(t_2), v'(t_3), \ldots v'(t_j), \ldots\} = v'(t) \tag{4}$$

jest przekazywany na zestaw układów, wyznaczających wartość funkcji charakterystycznych wzrostu $w(t_j)$ i malenia $m(t_j)$ sygnału $v'(t)$:

$$w(t_j) = \begin{cases} 1, \text{ jeśli } v'(t_j) > v'(t_j - t_{j-1}) \\ 0, \text{ w przeciwnym przypadku} \end{cases} \tag{5}$$

$$m(t_j) = \begin{cases} 1, \text{ jeśli } v'(t_j) \leqslant v'(t_j - t_{j-1}) \\ 0, \text{ w przeciwnym przypadku} \end{cases} \tag{6}$$

Procesu porównywania można dokonać za pomocą pojedynczego, klasycznego komparatora cyfrowego [10].

Wyznaczana jest także wartość $z(t_j)$ funkcji charakterystycznej zera

$$z(t_j) = \begin{cases} 1, \text{ jeśli } |v'(t_j)| < x_o \\ 0, \text{ w przeciwnym przypadku} \end{cases} \tag{7}$$

Wartość 1 tej funkcji jest otrzymywana za pomocą komparatora cyfrowego w przypadku, gdy moduł wartości sygnału $v'(t_j)$ jest mniejszy od pewnej założonej, bliskiej zera wartości progowej x_o. Wartości jeden funkcji $w(t_j)$, $m(t_j)$, $z(t_j)$ tworzą tzw. alfabet (lub zbiór symboli terminalnych (pomocniczych), lub zbiór zmiennych) Σ pewnej gramatyki formalnej G

$$\Sigma = \{w(t_j) = 1, \quad m(t_j) = 1, \quad z(t_j) = 1\}, \quad j = 1, 2, \ldots \tag{8}$$

Alfabet Σ można zapisać krócej w następujący sposób [1]

$$\Sigma = \{/ \cdot \backslash \cdot 0\}, \tag{9}$$

gdzie:

/ oznacza wartość 1 funkcji charakterystycznej $w(t_j)$, tzn. badany sygnał $v(t)$ wzrastał w chwili czasu t_j;

\ oznacza, że analizowany przebieg $v(t)$ malał w momencie $t(j)$, tzn. że $m(t_j) = 1$;

0 oznacza, że sygnał $v(t)$ w chwili t_j był bliski zeru, tzn. że $z(t_j) = 1$.

Alfabet Σ jest elementem następującej gramatyki G, deterministycznej i bezkontekstowej [1]

$$G = \{V, \Sigma, P, W\}. \tag{10}$$

Oto opis pozostałych elementów tej gramatyki G:

V jest zbiorem symboli końcowych (nieterminalnych):

$$V = \{W, \langle\text{fala}\rangle, \langle\text{pik}+\rangle, \langle\text{pik}-\rangle, \langle\text{wzrost}\rangle, \langle\text{malenie}\rangle, \langle\text{zero}\rangle\}, \tag{11}$$

W jest symbolem początkowym gramatyki G;

P jest zbiorem produkcji tej gramatyki; do zbioru P należą między innymi następujące elementy:

$$\begin{aligned}
&\{W \to \langle\text{zero}\rangle \mid \langle\text{fala}\rangle \mid \langle\text{zero}\rangle,\\
&W \to \langle\text{fala}\rangle \mid \langle\text{zero}\rangle,\\
&W \to \langle\text{zero}\rangle \mid \langle\text{fala}\rangle,\\
&W \to \langle\text{fala}\rangle,\\
&W \to \langle\text{zero}\rangle,\\
&\langle\text{fala}\rangle \to \langle\text{pik}+\rangle,\\
&\langle\text{fala}\rangle \to \langle\text{pik}-\rangle,\\
&\langle\text{fala}\rangle \to \langle\text{malenie}\rangle,\\
&\langle\text{fala}\rangle \to \langle\text{wzrost}\rangle,\\
&\langle\text{wzrost}\rangle \to \langle\text{wzrost}\rangle \mid /,\\
&\langle\text{wzrost}\rangle \to /,\\
&\langle\text{wzrost}\rangle \to \langle\text{wzrost}\rangle \mid \langle\text{zero}\rangle \mid \langle\text{wzrost}\rangle,\\
&\langle\text{pik}+\rangle \to \langle\text{wzrost}\rangle \mid \langle\text{malenie}\rangle,\\
&\langle\text{zero}\rangle \to \langle\text{zero}\rangle \mid 0,\\
&\langle\text{zero}\rangle \to 0,\\
&\langle\text{pik}-\rangle \to \langle\text{malenie}\rangle \mid \langle\text{wzrost}\rangle,\\
&\langle\text{malenie}\rangle \to \langle\text{malenie}\rangle \mid \backslash,\\
&\langle\text{malenie}\rangle \to \backslash,\\
&\langle\text{malenie}\rangle \to \langle\text{malenie}\rangle \mid \langle\text{zero}\rangle \mid \langle\text{malenie}\rangle\},
\end{aligned} \tag{12}$$

gdzie | oznacza złożenie (połączenie w czasie) dwóch elementów (patrz także wzory (17) i (18) oraz rys. 6).

Językiem tej gramatyki są między innymi następujące reprezentacje lewego i prawego zbocza piku [1]:

$$L = /\,|\,/\,|\{/\,|0\}^*|\,/, \tag{13a}$$

$$P = \backslash\,|\,\backslash\,|\{\backslash\,|0\}^*|\,\backslash, \tag{13b}$$

gdzie $\{/\,|0\}^*$ oznacza, że elementy $/$, 0 mogą się powtarzać dowolną ilość razy, np. $L = /\,|\,/\,|\,/\,|$; $L = /\,|\,/\,|\,/\,|0|\,/$;
$L = /\,|\,/\,|\,/\,|0|0|0|\,/$; $L = /\,|\,/\,|\,/\,|\,/\,|\,/\,|0|\,/\,|\,/\,|\,/$; itd.

Dodatni pik (tzn. $\langle$pik+$\rangle$) oraz ujemny pik (czyli $\langle$pik−$\rangle$) mogą być reprezentowane przez wyrażenia:

$$\langle\text{pik}+\rangle = L|P, \tag{14a}$$

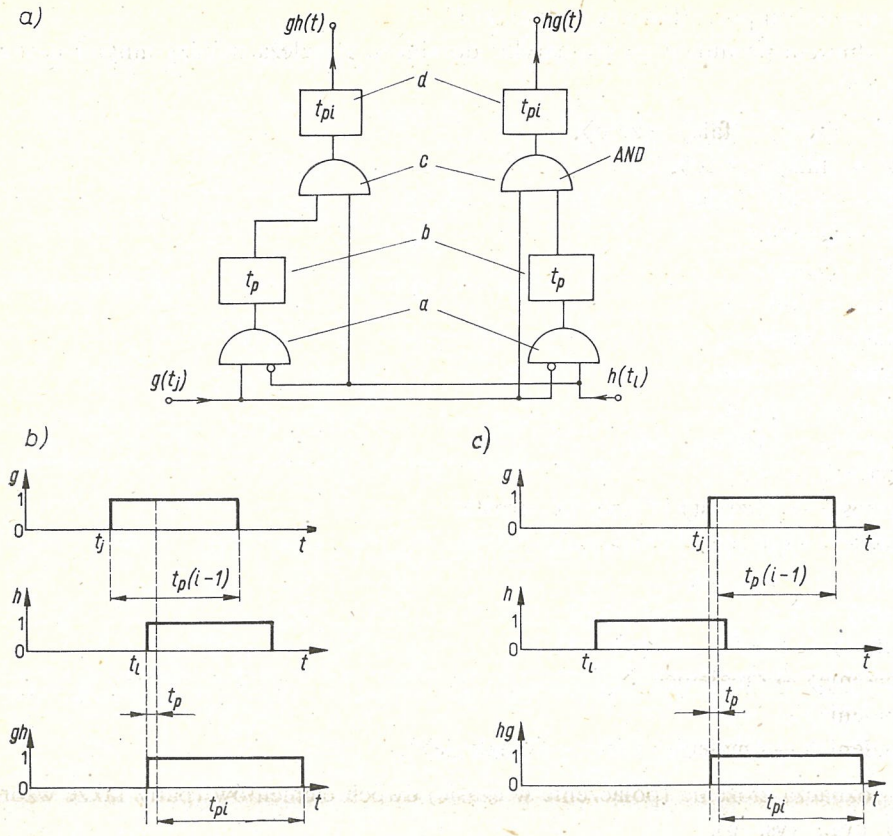
$$\langle\text{pik}-\rangle = P|L. \tag{14b}$$

Cały badany przebieg (tzn. $\langle$fala$\rangle$) może być reprezentowany w następujący sposób

$$\langle\text{fala}\rangle = 0^*|\ \ldots\ |\langle\text{pik}+\rangle|0^*|\langle\text{pik}-\rangle|\ \ldots\ |0^*. \tag{15}$$

Dla gramatyki G istnieje deterministyczny i skończony automat A

$$A = \{v'(t), \Sigma, S, F, I, z\}, \tag{16}$$



Rys. 6. Schemat blokowy (a) układu przeznaczonego do rozpoznawania porządku (kolejności) dwóch sygnałów, reprezentujących cechy $h(t_l)$ oraz $g(t_j)$, pojawiające się w momentach czasowych t_l oraz t_j; b), c) wykresy ilustrujące pracę układu z rys. (a)

za pomocą którego można automatycznie analizować wszystkie słowa, (tj. wyniki produkcji P) generowane (wytwarzane) przez tę gramatykę G, przy czym:

Σ jest alfabetem wejściowym (pomocniczym) — wzory (8) i (9);

z jest tzw. symbolem początkowym; $z \in S$;

S jest zbiorem wszystkich możliwych stanów automatu A;

F jest zbiorem stanów końcowych tego automatu;

I jest listą instrukcji (przejść) automatu.

W przypadku badań elektrokardiogramów, stany S są utworzone z dowolnych słów, generowanych przez gramatykę G. Do zbioru analizowanych stanów S należą więc między innymi zbiór V symboli końcowych gramatyki G oraz wyrażenia (13) ÷ (15), reprezentujące cechy i charakterystyczne fragmenty elektrokardiogramów. Zbiór stanów końcowych F jest zbiorem wyników procesu automatycznej analizy elektrokardiogramów. Na przykład, do zbioru F należą: wyniki pomiarów położeń (na osi czasu) wszystkich charakterystycznych punktów, zaznaczonych na rys. 1, np. punktów P, Q, T, E, wyniki pomiarów powierzchni pików P oraz U, odległość (na osi czasu) piku P od piku U, wyniki statystycznej analizy otrzymanych rezultatów dla całego badanego przebiegu.

Zbiór I jest zbiorem algorytmów, umożliwiających otrzymanie żądanych wyników. Dziedziną tych algorytmów jest zbiór stanów S oraz badany sygnał $v'(t)$ (wzór (1)), przeciwdziedziną — zbiór F. Na przykład, jedna z instrukcji, należąca do zbioru I, zapisuje w pamięci moment t_j oraz amplitudę sygnału $v'(t_j)$ otrzymaną dla chwili t_j, dla której otrzymywany jest ⟨pik+⟩. Element ⟨pik+⟩ może być rozpoznawany automatycznie w układzie, przedstawionym na rys. 6. Na rysunku tym na wejście, oznaczone symbolem $h(t_l)$ wprowadzany jest sygnał, reprezentujący wartości $w(t_j)$ funkcji charakterystycznej wzrastania przebiegu $v'(t)$ — wzór (5) — natomiast na wejście, oznaczone symbolem $g(t_j)$, jest wprowadzany sygnał reprezentujący wartości 0 oraz 1 funkcji $m(t_j)$ malenia przebiegu $v'(t)$ — wzór (6). Na wyjściu $hg(t)$ w momencie t_j przejścia piku dodatniego przez maksimum ukazuje się wartość cyfrowa 1. Moment ten jest sygnałem do zapisania w pamięci komputera tej chwili t_j oraz amplitudy $v'(t_j)$ w celu dalszej analizy.

Elementy z rysunku 6, zaznaczone symbolami a oraz c, realizują iloczyny logiczne. Elementy b zapewniają koincydencję sygnałów wejściowych dla bloków c. Układy d przechowują wartości logiczne 1 sygnałów, otrzymanych z wyjść układów c, przez krótkotrwałe odcinki czasu t_p. Na wyjściu $gh(t)$ otrzymywane są wartości 0 albo 1 funkcji charakterystycznej $gh(t)$ określonej w następujący sposób

$$gh(t) = \begin{cases} 1, \text{ jeśli } g(t_j) \prec h(t_l), \\ 0, \text{ w przeciwnym przypadku.} \end{cases} \tag{17}$$

Relacja $\prec$ oznacza, że wartość funkcji $gh(t)$ jest równa jeden, jeśli cecha g pojawiła się przed cechą h, przy czym pomiędzy chwilą t_j pojawienia się cechy g a chwilą t_l pojawienia się cechy h nie upłynął czas dłuższy od t_p [8]. Chwila pojawienia się wartości 1 funkcji h jest oznaczona indeksem t. Natomiast na wyjściu $hg(t)$ otrzymywane są wartości 0 albo 1 funkcji charakterystycznej $hg(t)$, zdefiniowanych w następujący sposób

$$hg(t) = \begin{cases} 1, \text{ jeśli } g(t_j) \prec h(t_l), \\ 0, \text{ w przeciwnym przypadku.} \end{cases} \tag{18}$$

LITERATURA

1. S. L. Horowitz, *Peak Recognition in Waveform*, in the ,,Syntactic Pattern Recognition, Applications", ed. by Fu K. S., Springer-Verlag, 1977.
2. M. Nowakowska, *Analiza dynamiki spostrzegania*, zbiór pt. ,,Problemy psychologii matematycznej", red. J. Kozielecki, s. 121—152, PWN, 1971.
3. *Pattern Recognition Theory and Application*, ed. by Eu K. S. and Whinston A. B., NATO Advanced Study Institutes Series, 1977.
4. Z. M. Wójcik, *Badania sztucznej inteligencji jako skuteczny środek poznania świata*, Studia Filozoficzne, 4, 1979, Warszawa.
5. W. Turski, *Struktury danych*, WNT, 1971.
6. Z. M. Wójcik, *Algorytm automatycznego kwantowania obrazów graficznych*, Rozprawy Elektrotechniczne, 1, 1980.
7. Z. M. Wójcik, *Cybernetyczny model znaczenia języka naturalnego*, Postępy Cybernetyki, No 1, 1980, Warszawa.
8. Z. M. Wójcik, *A Mathematical Model for the Determination of Speech Signal Parameters with Aid of the Instant Memory*, Archives of Acoustics, 2, 2, 1977, pp. 121—137.

9. Z. M. Wó cik, *Zagadnienia budowy sztucznej inteligencji*, Wiadomości Telekomunikacyjne, WKŁ, 9, 1978.
10. J. Pieńkos, J. Turczyński, *Układy scalone TTL*, WKŁ, 1977.
11. Z. M. Wójcik, *Badania współczesnych narzędzi człowieka za pomocą podstawowego warunku reprezentowania zdarzeń*, Prakseologia, 3, 1980, Warszawa.
12. Z. M. Wójcik, *A model of pattern recognition*, Proc. of the Fourth Intern. Symp. SYSTEM-MODELLING-CONTROL (org. by Polish Cybernetical Society), Zakopane, Oct. 8—13, 1979.
13. Z. M. Wójcik, *Ideja kratkowremjennogo zapominajuszczego ustrojstwa dla raspoznawanija zwukow reczi*, Trudy zasjedanija roboczej grupy 2 KNWWT po mjetodam raspoznawanija, klassifikacji i poiska informacji, Jadwisin, 2—8 czerwiec 1975, Instytut Organizacji i Kierowania PAN i MNSzWiT, Warszawa.

Z. M. WÓJCIK

METHOD OF AVERAGING AND OF RECOGNITION OF BIOMEDICAL SIGNALS

Summary

The paper presents a method and an electronic system for automatic averaging of a sampled analog signal, as well as a method and electronic systems for the recognitin of these averaged signals. The averaging consists in determining a weighted mean of the results of sampling at each measuring point of the signal under consideration, distributed symmetrically in relation to the point. Due to the weighting, the influence of neighboring values of the signal being measured decreases as the point grows more and more remote. The results of the averaging remain in accordance with the elementary condition of representing events [9, 4, 11, 12, 7]. In the process of recognition the values of characteristic functions of growth, decrease and zero of the signal under consideration are investigated [13, 8]. The values of these functions make up the alphabet of a formal grammar. The features of the signal under consideration constitute a set of final symbols of the grammar, being a result of the set of productions of the grammar. A certain automation exists for this grammar. The results of recognition and of statistical analysis of the signals under consideration constitute the final states of the automation. Electronic systems determining the most elementary characteristics of the signals are presented. The method presented is described with the recognition of electrocardiograms as an example.

Z. M. WÓJCIK

MÉTHODE DE DÉTERMINATION DES MOYENNES ET L'IDENTIFICATION DES SIGNAUX BIOMÉDICAUX

Résumé

L'étude contient la méthode et le système électronique de détermination automatique de la moyenne du signal analogique d'échantillon ainsi que la méthode et les systèmes électroniques de l'identification de ces signaux réduits à la moyenne. La détermination de la moyenne consiste en désignation d'une certaine moyenne pondérée des résultats d'échantillonnage de chaque point de mesure de signal examiné, posés symétriquement envers ce point. Grâce à l'emploi des balances, l'influence des valeurs voisines du signal mesuré sur la valeur moyenne (la résultante) du point de mesure examiné diminue avec l'éloignement de ce point. Les résultats de la détermination de la moyenne sont conformes à la principale condition de représentation des événements [9, 4, 11, 12, 7]. Au cours du processus de l'identification sont examinées les valeurs des fonctions caractéristiques de croissance, de diminution et de la valeur nulle du signal examiné [13, 8]. Les valeurs de l'unité de ces fonctions constituent un alphabet d'une certaine grammaire formelle. Les caractéristiques du signal examiné font un ensemble de symboles finals de cette grammaire. Il existe pour cette grammaire un certain automate. Les états finals de cet automate constituent les résultats d'identi-

fication et de l'analyse statistique des signaux examinés. On a présenté aussi les systèmes électroniques, déterminant les principaux caractéristiques des signaux. La méthode présenté a été décrite d'après l'identification des électrocardiogrammes.

Z. M. WÓJCIK

METHODE FÜR DIE MITTLUNG UND ERKENNUNG VON BIOMEDEZINISCHEN SIGNALEN

Zusammenfassung

In der Bearbeitung wird die Methode und das elektronische System für die automatische Mittlung des geprüften Analogsignals dargelegt, als auch die Methode und elektronische Systeme für die Erkennung dieser gemittelten Signale. Die Mittlung beruht auf der Bestimmung eines gewissen Mittelwertes der Prüfergebnisse für jeden einzelnen Meßpunkt des geprüften Signals. Die Meßpunkte sind im Verhältnis zu dem gegebenen Punkt symmetrisch verteilt. Zufolge der Anwendung von Waagen verringert sich der Einfluß der Nachbarwerte des gemessenen Signals auf den gemessenen (resultierenden) Wert des geprüften Meßpunktes je größer der Abstand von diesem Punkt wird. Die Ergebnisse der Mittlung stimmen mit den Grundbedingungen der Vertretung der Ereignisse überein [9, 4, 11, 12, 7]. Während des Erkennungsprozesses werden die Werte der Funktionen der Charakteristischen des Aufstiegs, der Abnahme und des Nullpunktes des geprüften Signals untersucht [18, 8]. Die Eins-Werte dieser Funktionen bilden das Alphabet einer gewissen formellen Grammatik. Die Eigenschaftem des geprüften Signals bilden eine Sammlung der Schlußsymbole dieser Grammatik, die im Ergebnis der Sammlung der Produktion dieser Grammatik erhalten werden. Für diese Grammatik gibt es ein gewisses Automat. Die Schlußzustände dieses Automates bilden die Ergebnisse der Erkennung und der statistischen Analyse der geprüften Signale. Es werden elektronische Systeme gezeigt, die die wichtigsten Grundeigenschaften der Signale bestimmen. Die dargestellte Methode wurde aufgrund eines Beispiels der Erkennung von Elektrokardiogrammen beschrieben.

З. М. ВУЙЦИК

МЕТОД УСРЕДНЕНИЯ И ИДЕНТИФИКАЦИИ БИОМЕДИЦИНСКИХ СИГНАЛОВ

Резюме

Представлен метод и электронная схема для автоматического усреднения опробовываемого аналогового сигнала, а также метод и электронная схема для идентификации усредненных сигналов. Усреднение состоит в определении некоторой средней взвешенной результатов опробований каждой точки измерения исследуемого сигнала расположенной симметрично по отношению к исходной точке. Благодаря применению взвешивания, влияния соседних значений измеряемого сигнала на усредненное (результативное) значение исследуемой точки измерения уменьшаются по мере отдаления от исходной точки. Результаты усреднения согласны с основным условием репрезентации явлений [9, 4, 11, 12, 7]. В процессе идентификации исследуются значения характеристических функций роста, уменьшения и нуля исследуемого сигнала [13, 8]. Единичные значения этих функций являются алфавитом некоторой формальной грамматики. Свойства исследуемого сигнала представляют собою множество конечных символов этой грамматики. Для этой грамматики существует определенный автомат. Конечные состояния этого автомата являются результатами идентификации и статистического анализа исследуемых сигналов. Представлены электронные схемы определяющие самые основные свойства сигналов. Представленный метод описан на основании диагнозирования электрокардиограмм.

TREŚĆ

J. Sochoń: Podstawy analizy komputerowej wieloparametrowych metod pomiaru wielkości nieelektrycznych . . . 3

M. Stabrowski: Numeryczne rozwiązywanie bardzo dużych pasmowych niesymetrycznych układów równań liniowych . . . 19

J. Sawicki: Innere Schaltungen nebst Konstruktionshinweisen für die Blindleistungsmessumformer mit starrer Phasenverschiebung . . . 31

L. Frąckowiak, S. Jackowski: Charakterystyki sprawności silników prądu stałego zasilanych z tyrystorowych układów prostownikowych . . . 49

R. Sikora, W. Lipiński, M. Gramz, K. Gawrylczyk, R. Pałka, S. Gratkowski, M. Ziółkowski, K. Pawluk: Pole magnetyczne i związane z nim parametry całkowe w strefie czołowej maszyn elektrycznych . . . 67

J. Prokop: Wpływ tłumienia wewnętrznego na transmitancję macierzową maszyny prądu stałego . . . 97

Z. Dmochowski: Badanie przekładników wielkoprądowych do stalowniczych pieców łukowych . . . 115

S. Lesiński: Prognostication of Contact Durability of Switches on the Basis of Probabilistic Model of Contact Wear . . . 137

J. Bajorek: Szybki algorytm przekształcenia Laplace'a dla analizy przebiegów przejściowych w liniach elektroenergetycznych . . . 155

H. Szukalski: Rozkład potencjałów i prądów w podziemnych przewodach ułożonych w pobliżu uziomów stacji elektroenergetycznych . . . 171

J. Smyczek: Wyznaczanie mocy szumów układu dynamicznego z nieliniową rezystancją . . 179

K. Buczek: Metoda analizy układów tyrystorowych wysokiej częstotliwości na przykładzie falownika wielotaktowego . . . 185

Z. Wąsowicz, P. Włoka: Zastosowanie udarów termicznych do przyspieszonego oszacowania charakterystyk niezawodności mikroelektronicznych hybrydowych układów grubowarstwowych . . . 201

S. Krawczyk: Własności nierównowagowe kondensatora typu MIS w warunkach oświetlenia . . . 219

W. Nakwaski: Sprawność sprzężenia lasera złączowego ze światłowodem włóknistym za pomocą mikrosoczewki . . . 233

J. Kaczmarek: Analiza wpływu prądu obciążenia na sterujący strumień magnetyczny w kontaktronie . . . 247

B. Wolszczak: Zastosowanie metody korekcji macierzy odwrotnej w programach projektowania układów elektronicznych . . . 261

H. Trzaska: Metody pomiaru gęstości mocy w polu bliskim . . . 269

J. Pieprzyk, D. Rutkowski: Kryptograficzne metody zabezpieczenia informacji . . . 285

Z. M. Wójcik: Metoda uśredniania i rozpoznawania sygnałów biomedycznych . . . 315

CONTENTS — TABLE DES MATIÈRES — INHALT

J. Sochoń: Foundations of Computer Analysis of Multiparameter Methods of Measuring Nonelectrical Quantities 17
Principes de l'analyse numérique des méthodes à paramètres multiples de mesure des grandeurs non électriques 17
Grundlagen einer rechnergestützten Analyse einer Mehrparametermessmethode für nichtelektrische Grössen 18
M. Stabrowski: Numerical Solution of Very Large Banded Unsymmetric Linear Equations Systems 28
Solutions numériques pour de très grands systèmes d'équations linéaires à bandes et asymétriques 28
Numerische Lösung sehr grosser unsymmetrischer Bandsysteme der Lineargleichungen . . . 29
J. Sawicki: Input Structures and Principles of Construction of Reactive Power Transducers with CONSTANT Phase Shift 46
Systèmes des circuits et les principes de construction des transducteurs de mesure de puissance réactive à déphasage constant 46
Układy torowe i zasady konstrukcji przetworników pomiarowych mocy biernej o stałym przesunięciu fazowym 47
L. Frąckowiak, S. Jackowski: Efficiency Curves of Direct-Current Motors Fed from Thyristor Rectifier Circuits 65
Caractéristiques de rendement des moteurs à courant continu alimentés par les montages des redresseurs à thyristor 65
Wirkungsgradcharakteristiken der durch Thyristorgleichrichtersysteme gespeisten Gleichstrommotoren 65
R. Sikora, W. Lipiński, M. Gramz, K. Gawrylczyk, R. Pałka, S. Gratkowski, M. Ziółkowski, K. Pawluk: Magnetic Field and Related Integral Parameter of the End-Zone of Electrical Machines 94
Champ magnétique et liés avec lui paramètres intégraux dans la zone frontale des machines électriques 94
Magnetfeld und die von ihm bestimmten Integralparameter der Endzone elektrischer Maschinen 94
J. Prokop: Influence of Inside Damping on Matrix Transfer Function of Direct Current Machine 112
Influence de l'atténuation intérieure sur la transmittance de matrice d'une machine à courant continu 113
Einfluss der inneren Dämpfung auf die Matrixtransmitanz der Gleichstrommaschine 113
Z. Dmochowski: Studies of Current Transformers for Large Currents Used in Steel Arc Furnaces 135
Examen des transformateurs de courant fort installés dans les fours d'aciérie à arc 135
Untersuchungen an Hochstromwandlern der Stahllichtbogenöfen 135
S. Lesiński: Prognozowanie trwałości styków łączników w oparciu o probabilistyczny model zużycia styków 152
Prévision de la durabilité des contacts des connecteurs sur la base du modèle probabiliste de l'usure des contacts 152
Prognosierung der Kontaktlebensdauer von Schaltgeräten in Anlehnung an das probabilistische Modell der Abbrand der Kontakte 152
J. Bajorek: Fast Algorithm of Laplace Transform for Calculating Transmission Lines Transients 168
Algorithme rapide de transformation de Laplace pour l'analyse des réponses transitoires dans les lignes d'énergie 168
Schnellalgorithmus der Laplace-Transformation zwecks Durchführung einer Analyse der Übergangsvorgänge in Starkstromleitungen 168

H. Szukalski: Distribution of Potentials and Currents in Underground Conductors Near Earth Electrodes in Substations 177
Répartition des potentiels et des courants dans les câbles se trouvant à proximité des prises de terre des postes électroénergetiques 177
Verteilung der Potentiale und Ströme in den in der Nähe der Erder elektromagnetischer Stationen unterirdisch verlegten Leitungen 178
J. Smyczek: Calculation of Noise Power in Dynamic Circuit with Nonlinear Resistance . . 183
Détermination de la puissance des bruits d'un circuit dynamique à résistance non linéaire . . 183
Bestimmung der Rauschleistung eines dynamischen Systems mit nichtlinearer Resistanz 184
K. Buczek: Method of Analysis of High Frequency Thyristor Systems by Example of Multitact Inverter 199
Méthode de l'analyse des systèmes thyristors de haute fréquence sur l'example d'un onduleur multiple 199
Analysenmethode für Hochfrequenz-Thyristorsysteme am Beispiel eines mehrgliedrigen Wechselrichters 199
Z. Wąsowicz, P. Włoka: Application of Thermal Shocks for Accelerated Estimation of Reliability Characteristics of Thick Film Microelectronic Hybrid Circuits 216
Utilisation du choc thermique pour accélérer l'estimation des caractéristiques de fiabilité des circuits micro-électroniques hybrides à couche épaisse 216
Anwendung von Wärmestössen zur beschleunigten Bewertung der Zuverlässigkeitscharakteristiken mikroelektronischer Dickschicht-Hybridenschaltungen 216
S. Krawczyk: Non-Equilibrium Properties of MIS Capacitor Under Illumination Conditions 231
Propriétés du déséquilibre de la capacité du condensateur type MIS dans les conditions de l'éclairage 231
Die ungleichwertigen Eigenschaften des MIS-Kondensators in Beleuchtungsverhältnissen . . 231
W. Nakwaski: The efficiency of Coupling Light From a Laser Diode Into an Optical Fiber With the Aid of a Lens 244
Rendement du couplage d'un làser à jonction au conduit de lumière fibreux à l'aide d'une microlentille 244
Leistungsfähigkeit der Kopplung eines Verbindungsglases mit einem fasrigen Lichtleiter mittels einer Mikrolinse 245
J. Kaczmarek: Analysis of Load Current Influence on Dry-Reed Control Magnetic Flux 258
Analyse de l'influence du courant de charge sur le flux magnétique de commandé dans le relais à lames souples 258
Analyse des Laststromeinflusses auf den Steuermagnetfluss im Schutzrohrkontakt 258
B. Wolszczak: Application of Inverse Matrix Correction Method in Electronic Circuit Design Programs 267
Application de la méthode de correction de la matrice inverse dans les programmes des projets des circuits électroniques 268
Anwendung der Korrektionsmethode einer Umkehrmatrix für Projektionsprogramme elektronischer Systeme 268
H. Trzaska: Methods of the Near Field Power Density Measurements 282
Méthode de mesure de la densité de puissance dans le champ proche de la source de rayonnement électromagnétique 282
Messmethoden für die Leistungsflussdichte in einem Nahfeld 282
J. Pieprzyk, D. Rutkowski: Cryptographic Methods of Information Protection . . . 312
Méthodes cryptographiques de la protection de l'information 312
Kryptographische Methoden für Informationssicherung 313
Z. M. Wójcik: Method of Averaging and of Recognition of Biomedical Signals 324
Méthode de détermination des moyennes et l'identification des signaux biomédicaux 324
Methode für die Mittlung und Erkennung von biomedizinischen Signalen 325

СОДЕРЖАНИЕ

Я. Сохонь: Основы компютерного анализа многопараметровых методов измерения неэлектрических величин 18
М. Стабровски: Численное решение очень больших ленточных несимметрических систем линейных уравнений 29
Е. Савицки: Цепные схемы и принципы конструкции измерительных преобразователей реактивной мощности с постоянным фазовым сдвигом 47
Л. Фронцковяк, С. Яцковски: Характеристики коэффициента полезного действия двигателей постоянного тока питаемых от тиристорных выпрямителей . . . 66
Р. Сикора, В. Липиньски, М. Грамз, К. Гаврыльчик, Р. Палка, С. Гратковски, М. Зюлковски, К. Павлюк: Магнитное поле и связанные с ним интегральные параметры в торцевой зоне электрических машин 95
Я. Прокоп: Влияние внутреннего демпфирования на матричную передаточную функцию машины постоянного тока 113
З. Дмоховски: Исследование высокотоковых трансформаторов тока для дуговых сталеплавильных печей 135
С. Лесиньски: Прогнозирование долговечности контактов коммутационных аппаратов на основании вероятностной модели износа контактов 153
Е. Байорек: Скорый алгоритм преобразования Лапласса для анализа переходных процессов в электроэнергетических линиях 169
Х. Шукальски: Распределение токов и напряжений в подземных проводниках проложенных вблизи заземлений электроэнергетических станций 178
Е. Смычек: Определение мощности шумов динамической системы из нелинейной резистанции 184
К. Бучек: Способ анализа тиристорных системов высокой частоты на примере многотактного инвертора 200
З. Вонсович, П. Влока: Применение термических ударов для ускоренного определения характеристик надежности толстопленочных гибридных микросхем 217
С. Кравчик: Неравновесные свойства конденсатора типа MIS в условиях освещения 231
В. Накваски: К.п.д. сцепления с помощью линзы инжекционного лазера со стеклянным волоконным световодом 245
Е. Качмарек: Анализ влияния тока нагрузки на управляющий магнитный поток проходящий через геркон 259
Б. Вольщак: Применение метода коррекции обратной матрицы в программах анализа электронных схем 268
Х. Тшаска: Методы измерения плотности потока энергии в ближней зоне 283
Ю. Пепшик, Д. Рутковский: Криптографические методы охраны информации . . 313
З. М. Вуйцик: Метод усреднения и идентификации биомедицинских сигналов 325

WYTYCZNE DLA AUTORÓW

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

1. *Tematyka, charakter i oryginalność artykułów.* „Rozprawy Elektrotechniczne" są przeznaczone do ogłaszania prac naukowych wchodzących w zakres szeroko pojętej elektrotechniki, zwłaszcza jeżeli traktują one omawiane zagadnienia w sposób możliwie kompletny i zmierzają do osiągnięcia celu informacyjno-dydaktycznego. W związku z tym publikowane prace mogą być w mniejszym lub większym stopniu pracami nieoryginalnymi, jeśli chodzi o treść w nich zawartą, co nie zwalnia Autorów od obowiązku ujęcia tematu w sposób oryginalny. Twórczy wkład Autora może polegać wówczas np. na: porównaniu opisywanych teorii, metod lub systemów, krytycznej ich ocenie, wyciągnięciu własnych wniosków co do celowości takiego lub innego działania, zastosowaniu własnej klasyfikacji zagadnień, prognostyce itp. W szczególności „Rozprawy" ogłaszają takie prace naukowe, które zawierają: analizę porównawczą teorii, metod lub systemów, syntetyczne ujęcia określonego zagadnienia naukowego, omówienie obecnego stanu wybranego działu elektrotechniki, omówienie jego postępów lub tendencji rozwojowych itp.

2. *Wymagania ogólne.* Artykuły należy nadsyłać w maszynopisie, w dwóch egzemplarzach. Maszynopis powinien być napisany jednostronnie przez czarną kalkę, na maszynie do pisania z niezabrudzonymi i nieuszkodzonymi znakami. Dopuszcza się odręczne czytelne uzupełnienie lub pisanie atramentem albo 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 maszyn do pisania koloru białego, formatu A-4; numeracja ciągła na wszystkich kartkach. W artykułach należy stosować Międzynarodowy Układ Jednostek SI.

3. *Sposób pisania tekstu.* Tekst w maszynopisie powinien być pisany bez używania wyróżnień, a w szczególności nie dopuszcza się spacjowania (rozstrzelania), podkreślania i pisania tekstów dużymi literami. Proponowane wyróżnienia Autor może wprowadzić do maszynopisu (zwykłym ołówkiem) za pomocą przyjętych znaków adiustacyjnych (podkreślenie linią przerywaną — spacjowanie, podkreślenie linią ciągłą — pogrubienie, podkreślenie wężykiem — kursywa itp.). Na jednej kartce maszynopisu powinno być 30 wierszy po około 60 znaków łącznie z odstępami. Marginesy każdej kartki 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 w maszynopisie 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.

5. *Sposób pisania wzorów matematycznych.* Rozmieszczenie znaków, cyfr, liter i odstępów powinno być zbliżone do rozmieszczenia elementów druku. Wskaźniki i wykładniki 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, 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 streszczenia w językach obcych: angielskim, francuskim, niemieckim i rosyjskim (w 2 egz.). W razie niemożności przygotowania streszczeń w językach obcych Autor powinien przynajmniej podać niezbędną do wykonania tłumaczenia obcojęzyczną terminologię.

8. *Literatura.* Po zakończeniu artykułu należy podać wykaz literatury wymieniając 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.

Uwagi

1. 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.

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

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

Cena zł 200,—

Rozprawy Elektrotechniczne

Kwartalnik

CZASOPISMO
ROZPRAWY ELEKTROTECHNICZNE

Warunki prenumeraty
Cena prenumeraty krajowej
półrocznie zł 200,—
rocznie zł 400,—

Prenumeratę **na kraj** przyjmują Oddziały RSW „Prasa—Książka—Ruch" oraz urzędy pocztowe i doręczyciele w terminach:
— do 25 listopada na I półrocze roku bieżącego i na cały rok następny,
— do 10 czerwca na II półrocze roku bieżącego.
Jednostki gospodarki uspołecznionej, instytucje, organizacje i wszelkiego rodzaju zakłady pracy zamawiają prenumeratę w miejscowych Oddziałach RSW „Prasa—Książka—Ruch", w miejscowościach zaś, w których nie ma Oddziałów RSW — w urzędach pocztowych.
Czytelnicy indywidualni opłacają prenumeratę wyłącznie w urzędach pocztowych i u doręczycieli.
Prenumeratę ze zleceniem wysyłki **za granicę** przyjmują RSW „Prasa—Książka—Ruch", Centrala Kolportażu Prasy i Wydawnictw, ul. Towarowa 28, 00-958 Warszawa, konto NBP XV Oddział w Warszawie Nr 1153-201045-139-11 — w terminach podanych dla prenumeraty krajowej.
Prenumerata ze zleceniem wysyłki za granicę jest droższa od prenumeraty krajowej o 50% dla zleceniodawców indywidualnych i o 100% dla zleceniodawców instytucji i zakładów pracy.
Bieżące i archiwalne numery można nabyć lub zamówić we Wzorcowni Wydawnictw Naukowych PAN-Ossolineum-PWN, Pałac Kultury i Nauki (wysoki parter) 00-901 Warszawa oraz w księgarniach naukowych „Domu Książki".

A subscription order stating the period of time, along with the subscriber's name and address can be sent to your subscription agent or directly to Foreign Trade Enterprise Ars Polona — Ruch, 00-068 Warszawa, 7 Krakowskie Przedmieście, P. O. Box 1001, Poland. Please send payments to the account of Ars Polona — Ruch in Bank Handlowy S. A., 7 Traugutt Street, 00-067 Warszawa, Poland.

Rozpr. Elektrot. T. 28, z. 1—2, 1—332, Warszawa 1982

Index 37483

POLSKA AKADEMIA NAUK
KOMITET ELEKTROTECHNIKI

PLISSN 0035-938

ROZPRAWY ELEKTROTECHNICZNE

TOM XXVIII—ZESZYT 3—4

KWARTALNIK

PAŃSTWOWE WYDAWNICTWO NAUKOWE
WARSZAWA 1982

POLSKA AKADEMIA NAUK
KOMITET ELEKTROTECHNIKI

ROZPRAWY ELEKTROTECHNICZNE

TOM XXVIII-ZESZYT 3-4

KWARTALNIK

PAŃSTWOWE WYDAWNICTWO NAUKOWE
WARSZAWA 1982

RADA REDAKCYJNA

Przewodniczący

prof. dr hab. JANUSZ GROSZKOWSKI — członek rzeczywisty PAN

Członkowie

prof. dr hab. LUDWIK BADIAN — członek korespondent PAN, prof. TADEUSZ CHOLEWICKI — członek rzeczywisty PAN, doc. dr ZDZISŁAW KACHLICKI, prof. dr hab. TADEUSZ KACZOREK, prof. dr LEONARD KNOCH, doc. dr hab. ANDRZEJ KUSY, prof. JAN MANITIUS, prof. dr WITOLD NOWICKI, prof. dr WITOLD ROSIŃSKI — członek korespondent PAN, prof. dr JERZY SEIDLER — członek korespondent PAN, prof. dr RYSZARD SIKORA, prof. dr hab. JANUSZ TUROWSKI, prof. dr STEFAN WĘGRZYN — członek rzeczywisty PAN

REDAKCJA

Redaktor naczelny

prof. dr WITOLD NOWICKI

Zastępca redaktora naczelnego

prof. dr hab. TADEUSZ KACZOREK

Sekretarz odpowiedzialny

red. JANINA MATHEISEL

ADRES REDAKCJI

00-661 Warszawa, Plac Jedności Robotniczej 1
Politechnika, Instytut Telekomunikacji, Gmach Elektroniki, pok. 483

Dyżury Redakcji: środy i piątki, godz. 11-13
tel. 28 89 91 lub 210 07-737

Telefony domowe: redaktor naczelny — 17 37 03
zastępca redaktora naczelnego — 44 88 34
sekretarz odpowiedzialny — 33 84 97

PAŃSTWOWE WYDAWNICTWO NAUKOWE
Warszawa, Miodowa 10

Nakład 570	Oddano do składania 6.V.1982 r.
Ark. druk. 19,75+0,50 luźna wkładka Ark. wyd. 24,50	Podpis do druku w kwietniu 1983 r.
Papier druk. b. sat. kl. III, 90 g, 70×100	Druk ukończono w lipcu 1983 r.
Zamówienie nr 2953/82	Cena zł 200,— M-6

Drukarnia im. Rewolucji Październikowej, Warszawa

ROZPRAWY ELEKTROTECHNICZNE 1982, 28, z. 3-4, ss. 303—312

538.3

Nonlinearity Problem in Electrodynamics

ADAM MAISON

The paper received 7.5.1980

The main ideas of de Broglie constituting a critical review of purely probabilistic interpretation of wave mechanics and quantum physics are presented. References are made to simplified theory of electricity formulated by Planck in 1922 and to Einstein's theory of general relativity. The theory of polar molecule of Debye is recalled in conclusion.

1. INTRODUCTION

The nonlinearity problem in electronic engineering is as old as electronics but it was usually latent and treated as approximately obyeing linear equations. The linear approximations accompanied the whole development of electronic engineering and theoretical physics including the general relativity of Einstein.

The Journal of Applied Physics [1], ran a very topical paper concerning the nonlinearity problem in electronic engineering, particularly the varavtors. Nevertheles, the nonlinearity problem is more fundamental than that of engineering It concerns the most fundamental problems of human science on nature. The linear approximations were very fruitful asthe methods of approach to physical reality but in electrical and electro nic engineering they were treated not as approximations but as unquestionable canons.

In order to facilitate the historical review of the nonlinearity problem in electrodynamics, the author will restrict himself only to the works of the two eminent physicists namely, Max Planck and Louis de Broglie without any presumption of completness in the presentation of their grand scientific works.

2. SIMPLIFIED ASSUMPTIONS OF MAX PLANCK OF 1922 [2]

Max Planck deduced Maxwell's equations from the principle of the conservation of energy and from the principle of contiguous action introducing the Poynting's law of energy flux and the notions of volume and surface divergences. Neither the Poynting vector nor the notion of divergence exist in Maxwell's treatise of 1873.

Planck introduces the vectors

$$\varepsilon E = D, \tag{1}$$

$$\varkappa E = J \ (\varkappa \text{ is the space density of charge}), \tag{2}$$

$$\mu H = B \tag{3}$$

and so Maxwell's equations assume the form

$$\dot{D} = c\,\mathrm{curl}\ H - 4\pi J, \tag{4}$$

$$\dot{B} = -c\,\mathrm{curl}\,E \tag{5}$$

which were distinguished by the fact that they contain no constant which refers to the particular nature of the body and so have a universal character. This form of equations can also be applied to nonhomogeneous and non-isotropic bodies so long as they are continuously extended and may be assumed to be at rest. Within the range thus specified these equations have been found to hold extremally well wherever they have been applied. They form, so to speak, the solid framework into which all the individual peculiarities conditioned by the special properties of various bodies can be fitted.

On the basis of the above given Maxwell's equations, M. Planck considers the ecletrostatic field without contact potentials, electrostatic field with contact potentials the magnetostatic field, ponderomotive actions in the statical field, stationary electromagnetic field, quasi—stationary processes in the case of unclosed currents and dynamical processes in stationary bodies.

Finally, Planck considers the limits of the electrodynamics of Maxwell and Hertz and criticizes Lorentz's equations remembering the experiments of Michelson and Morley which excluded any influence of the earth's motion on the relative velocity of propagation of light. Further, Planck cites Einstein's opinion that Lorentz's equations of electrodynamics are just as completely invariant as the exact fundamental equations of mechanics for transformations from a stationary coordinate system to one moving with uniform velocity with respect to it, but that we may not use Galilei's transformation for this purpose. Instead, we must use the Lorentz transformation which may be regarded as a generalization of the former transformation since it becomes identical with it when $c \to \infty$.

Planck concludes that it is not possible to go further into Einstein's principle of relativity, as the purpose of his book is to serve as an introduction to the theory of electricity. It should be emphasized that Planck was first to fully recognize Einstein's view on relationship between space and time, but he passed over in silence the fact that Einstein in his general theory of relativity says that the space — time metrics obey non-linear equations and linear ones only approximately. So, Einstein introduced a linear approximation in his general relativity being, however, fully conscious that this is only an approximation facilitating mathematical expression of covariance for four-dimensional space but does not express strictly the physical reality.

Planck, in his book, recalls the principle of contiguous action and the principle of action at a distance and states that both principles are by no means regarded as co-ordinated; rather the principle of action at a distance is of a more general nature, whereas that of contiguous action is more special. It is owing to this that there have been several different theories of action at a distance in electrodynamics, as says Planck, but only one of contiguous action — namely, that of Maxwell. But contrary to Planck, it should be emphasized that the principle of contiguous action was formulated else where in 1836 by Faraday who introduced the term „dielectric" into science on electromagnetism. Maxwell adopted Faraday's conception and developed it. He introduced the term „displacement current" which is a fundamental assumption in his theory of electrodynamics

although it is only approximative in comparison with Faraday's views on solid dielectrics. Planck states that the theory of Maxwell owes its sovereign position over all others not to its greater „correctness", but rather to its greater definiteness and simplicity. The theory of Maxwell, though approximative, is accepted as fundamental in electrical and electronic engineering as well as in theoretical physics.

It is true that Planck's book from 1922 influenced predominantly the views of electronic engineers and theoretical physicists. In electronic engineering, Maxwell's equations were accepted owing to Heaviside, the unit propagation vector of electromagnetic wave k being admitted as fundamental despite of its only approximate significance. In theoretical physics, the Ψ wave which is Maxwell's electromagnetic wave was adopted as fundamental owing to Planck despite of the obscurities and inadequacies that emerge from this assumption by considering the wave mechanics. But, it should be reminded that Planck passed over in silence Planck's constant h introduced by him into science else where in 1900. This means that Planck in his book of 1922 does not connect Planck's constant h with the science on electromagnetism.

Louis de Broglie was the first to introduce Planck's constant h into electrodynamics in 1923 creating the basis of the wave mechanics. He was as strongly criticized as Planck was when he published his conception in 1900. However, the united conceptions of the two grand physicists that of 1900 and that of 1923 were very fruitful in the development of wave mechanics or quantum mechanics and the problem of nonlinearity in electrodynamics.

3. LONG — DRAWN CONTROVERSIES BETWEEN THE PHYSICISTS

3.1. Solway Congress in Brussels (October 1927)

At the Solway Congress, while a few of the „old guard" (Lorentz, Einstein, Langevin, Schrödinger) insisted on finding a causal interpretation of wave mechanics — without, however, coming in favor of de Broglie's proposal — Bohr and Born, along with their young disciples (Heisenberg, Dirac etc.) came out categorically in favour of the new purely probabilistic interpretation that they had developed and they did not even discuss the point of view of de Broglie. Pauli was the only one to present a definite objection to the pilot-wave approach of de Broglie and he did so by examining the case of a collision between a particle and a rotator, which Fermi studied a short while before.

The question of causal interpretation of wave mechanics finds its full explanation in the work of de Broglie of 1964 [3]. In this book de Broglie undertakes a thorough criticism of the „orthodox" purely probabilistic interpretation of wave mechanics which prevailed during the Solvay Congress of October 1927. He shows that the usual formalism, though strict in appearance and generally leading to precise conclusions, does not provide a profound and convincing explanation of the physical reality.

The criticism of the orthodox interpretation of wave mechanics undertaken by de Broglie is dazzling although his theory of the double solution does not spell Finality. There is no finality in the science on nature.

3.2. Localication Principle and Spectral Resolution

Louis de Broglie emphasises that the world of atoms and particles escapes our direct observation. We understand it only through the intermediary „observable particle localisation" i.e. phenomena in which particle acting at the microphysical level initiates an observable effect by a chain reaction.

Another approach, which enters into the Theory of Transformations consists in comparing particle localisation with a contraction of the associated wave train to extremely small dimensions, represented approximately by a Dirac function $\delta(R-R_o)$. But such a contraction cannot be achieved experimentally and offers nothing but the abstract image of wave train reduced to infinitely small dimensions.

The wave equations used today depend on the species of the particle under consideration. For particles with zero spin, for which relativity corrections may be considered to be negligible, a simple wave function Ψ will suffice and Schrödinger's wave equation can be written

$$\frac{h}{2\pi i}\frac{\partial \Psi}{\partial t} = \left[-\frac{h^2}{8\pi^2 m}\Delta + V(x, y, z, t)\right]\Psi \tag{6}$$

or

$$\frac{h}{2\pi i}\frac{\partial \Psi}{\partial t} = H\Psi, \tag{7}$$

where

$$H = -\frac{h^2}{8\pi^2 m}\Delta + V.$$

If, however, we wish to take into account relativity corrections, then we must use the Klein-Gordon equation which in absence of a field may be written

$$\Box\Psi + \frac{8\pi^2}{h^2} m_o^2 c^2 \Psi = 0. \tag{8}$$

For a particle with spin, it is assumed that the function Ψ has four componenents Ψ_k $(k = 1, 2, 3, 4,)$ which satisfy Dirac's system of four equations. In field-free space, these equations are of the form

$$\frac{h}{2\pi i}\frac{\partial \Psi}{\partial t} = \left(\alpha_1 \frac{\partial}{\partial x} + \alpha_2 \frac{\partial}{\partial y} + \alpha_3 \frac{\partial}{\partial z} + \alpha_4 m_o c\right)\Psi_k \quad (k = 1, 2, 3, 4), \tag{9}$$

where $\alpha_1, \alpha_2, \ldots$ are the Dirac matrix operators.

All these wave equations have a common property of great importance. They enable us to define two real quantities — namely, a density ϱ and a flux ϱv, which can be considered as being characteristic of a hypothetical conservative fluid which is associated with the propagation of the wave.

This fluid is conservative because, by virtue of the equation or equations of propagation, the quantities ϱ and ϱv satisfy the continuity equation

$$\frac{\partial \varrho}{\partial t} + \operatorname{div}(\varrho v) = 0, \tag{10}$$

which ensures that the quantity $\int \varrho d\tau$ is independent of time. This fact enables us to „normalise" ϱ by putting $\int \varrho d\tau = 1$. In the case of Schrödinger equation we have

$$\varrho = \Psi \cdot \Psi^* = |\Psi|^2; \quad \varrho v = \frac{h}{4\pi i m} (\Psi \operatorname{grad} \Psi^* - \Psi^* \operatorname{grad} \Psi), \tag{11}$$

where the asterisk denotes a complex conjugate.

If we put $\Psi = ae^{\frac{2\pi i}{h}\varphi}$ with real a and φ, then we can also write

$$\varrho = a^2; \quad \varrho v = -\frac{a^2}{m} \operatorname{grad} \varphi; \quad v = \frac{1}{m} \operatorname{grad} \varphi. \tag{12}$$

Analogus expressions can be found for ϱ and ϱv in the case of the Klein-Gordon equation and the Dirac equation.

Not long after the publication of Schrödinger's original paper in the spring of 1926, an effort was made to understand the nature and uses of the wave function Ψ. It was Max Born who in 1926—1927, was first to discover the probabilistic interpretation of the wave function Ψ. Louis de Broglie presented his own views in the form of two distinct principles: the principle of localisation and the principle of spectral resolution.

Formally, the principle of localization can be absorbed into a generalised form of the principle of spectral resolution. By using the well-known Dirac δ-function, the localisation of a particle at a point in space defined by a position vector R_o can be represent by the eigenfunction $\delta(R-R_o)$ corresponding to the eigenvalue R_o. In view of the properties of the Dirac function the explicit formula can then be written

$$\Psi(R) = \int \Psi(R_o)\, \delta(R-R_o) d\tau, \tag{13}$$

where $d\tau = dR_o = dx_o\, dy_o\, dz_o$.

In applying the principle of generalized spectral resolution, it can be stated that the probability of a particle appearing in an element of volume $d\tau$ in the vicinity of the point R_o is $|\Psi(R_o)|^2 d\tau$. Thus as a result of this elegant mathematical argument, the principle of localisation appears to have been absorbed into the generalized principle of spectral resolution. This elegant demonstration, which is the basis of the transformation theory is rather deceptive and gives a distorded view of physics. Thus, the principle of localisation seems to be Broglie to be in reality independent of the principle of spectral resolution, for the same reason that the observable localisation processes play the fundamental role in microphysics.

In concluding this usual formalism of wave mechanics de Broglie mentioned the Heisenberg uncertainty relationship. These equations are, mathematically, a consequence of the law of probability for two canonically conjugate variables such as x and p_x, and the laws themselves follow from the localisation principle and from the principle of spectral resolution.

One method of deducing the uncertainty relationship is based on the use of the Fourier integral. Thus, consider a wave train which, for example, at time $t = 0$, has a Fourier integral of the form

$$\Psi(x, y, z, 0) = \iiint c(p_x, p_y, p_z) e^{-\frac{2\pi i}{h}(p_x x + p_y y + p_z z)} dp_x dp_y dp_z. \tag{14}$$

Let us suppose that the wave train has maximum linear dimensions Δx, Δy and Δz along the three rectangular coordinate axes. In order to represent it by a Fourier integra, it is necessary that the triple integral in (14) acts on the intervals Δp_x, Δp_y and Δp_z (the c's being non-zero) such that

$$\Delta x \Delta p_x \geqq h; \quad \Delta y \Delta p_y \geqq h; \quad \Delta z \Delta p_z \geqq h. \tag{15}$$

Since according to the localisation principle, the particle may appear at any point of the wave train, and since the measurement of its momentum can only yield values for p_x, p_y and p_z which lie within Δp_x, Δp_y and Δp_z respectively, it follows that $\Delta x, \Delta y, \Delta z$, $\Delta p_x, \Delta p_y, \Delta p_z$, are the „uncertainties" in the values of the coordinates and momenta and the inequalities of equation (15) may be referred to as the „uncertainty relationships',

3.3. Meaning of q—representation and p—representation in the configuration space and physical space

The usual interpretation of wave mechanics is purely probabilistic and no attempt is made to look beyond the laws of probability and the idea of a hidden reality, on which the laws of probability are based is rejected. This positivistic interpretation is based on the assertion that everything which is unobservable is non-existent and should have no place in theoretical physics. This assertion seems to de Broglie to be highly debatable. Those who accept without reservation the purely probabilistic interpretation speak constantly of electrons, atoms and atomic nuclei, although these basic units of matter have never been directly observed.

Bohr and Heisenberg seem to have proved that the measurement of a physical quantity introduces a modification of the situation which exists prior to measurement, because of the existence of the quantum of action. There is certainly no doubt that the measurement of the momentum of a particle yields the value for the momentum which may be very different from that which existed prior to measurement.

The usual theory describes the stationary states of quantized system by assigning to them the approximate eigenfunctions, but is completely incapable of describing the rapid transitions which take place from one stationary state to another, quantum transitions which accompany the emission of radiation in Bohr's theory of the atom. A long time ago Bohr said that the description of these quantum transitions has „transcended" the framework of space — time, which is purely and simply a refusal to offer an explanation. Schrödinger also was able to say with humour: „the present theory describes minutely the stationary states which are of no interest because nothing takes place here, but it remains silent about the intermediate states". This remark emphasises an insolvency which has a very profound significance.

De Broglie postulates the Principle of the double Solution which states: „To every regular solution $\Psi = ae^{\frac{2\pi i}{h}\varphi}$ of the wave equation of wave mechanics, there corresponds a solution of the form

$$u = fe^{\frac{2\pi i}{h}\varphi}, \tag{16}$$

having the same phase φ, but with an amplitude f involving a point singularity which is, in general, in motion." But, as in general relativity, the field equations should be non-linear and the motion of the particle should result from this non-linearity.

The solution of the non-linear equation could be written as

$$u = u_o + v, \tag{17}$$

where $u_o \ll v$ within the entire linear domain outside the singular region.

Actually since the Ψ-wave which is a representation of probability, must be constructed on the state of particle, we can define it, as being everywhere proportional to the v-wave so that

$$\Psi = Cv, \tag{18}$$

where C is normalisation constant.

Moreover, the Ψ-wave would not conform to a non-linear equation because the principle of superposition appears to be a necessary condition of the normal statistical interpretation. The objective v-wave can determine physical phenomena such as interference, diffraction and quantized states of atomic systems; the Ψ-wave is only a representation of probability and is subjective in character and can be normalised at will.

De Broglie reminds the statement of the theory of transformation. In wave mechanics one regards every physical quantity as corresponding to a linear and Hermitian operator A, which has a series of complex eigenfunctions φ_i. These eigenfunctions form a complete orthogonal set so that a wave function Ψ can always be expanded in the form

$$\Psi = \sum_i c_i \varphi_i, \tag{19}$$

where c_i are the complex generalised Fourier coefficients. They are given by

$$c_i = \int \varphi_i^* \Psi d\tau. \tag{20}$$

If the spectrum of c_i is continuous, then equation (19) still applies, but the φ_i are now the „eigen-differentials" of the spectrum. These eigendifferentials represent groups of waves whose introduction can be justified physically by nothing that monochromatic waves are abstractions and that only groups of waves have a physical reality.

The quantities c_i are the „coordinates" of Ψ in the Hilbert space relative to a system of reference defined by the basic functions φ_i. If the coefficients c_i are known, then Ψ is known. If we transfer from the set of basic function φ_i to another set of basic functions φ_i' then, since the entire theory is linear, we have

$$\varphi_i = \sum d_{ki} \varphi_k' \tag{21}$$

and consequently

$$\Psi = \sum_i c_i \varphi_i = \sum_k c_k' \varphi_k', \tag{22}$$

where

$$c_k' = \sum d_{ki} c_i.$$

Moreover, it is assumed that the eigen functions φ_i which correspond to the position vector R_o of the particle are the Dirac functions $\delta(R - R_o)$.

We have already seen that this hypothesis is physically dubious because observable particle localisation results from a short range effect between the particle and another microphysical system, which give rise o an obsevable phenomenon by a chain proces, and not by a reduction of the wave function to a Dirac δ function.

If we accept, as is customary, the doubtful postulate in question, we can write

$$\Psi(R) = \int \Psi(R_o)\,\delta(R-R_o)\,dR \tag{23}$$

and we can consider $\Psi(R_o)$ as being the coefficient c in expansion of Ψ in terms of eigenfunctions of position. Morever, we note that if the φ_i are functions of x, y and z, and the coefficients c_i can be time dependent, then where (19) is substituted into the wave equation

$$\frac{h}{2\pi i}\frac{\partial \Psi}{\partial t} = H\Psi, \tag{24}$$

where H is the Hamiltonian operator, we have, after multiplying by φ_i^* and integrating with respect to $d\tau$,

$$\frac{h}{2\pi i}\frac{\partial c_j}{\partial t} = \sum H_{ji}c_i, \tag{25}$$

where $H_{ji} = \int \varphi_j^* H \varphi_i\, d\tau$ is the matrix element corresponding to the operator H in the φ_j system. Equation (25) which is Dirac's equation for the variation of constants, can be considered equivalent to the wave equation.

This elegant formalism suggests that all „representations" of the Ψ function by complete and orthonormal system of eigenfunctions corresponding to physical quantities are equivalent from a physical point of view. In particular the „q-representation" should be just as significant as the „p-representation" i.e. the momentum representation. However, the q-representation is the only objective representation, whilst the p-representation the abstract representation in momentum space — exists only in the mind of the theoretician.

The two representations — q and p — are by no means equivalent. This conclusion is the consequence of the obvious fact that three dimensional space is a physical reality and is essential framework of our experiment, whilst momentum space is only abstract mathematical representation.

The probability density exists in the initial state prior to the action of any measuring device which measures the quantity A, with the eigenfunctions φ_i. The $|c_i|^2$ cannot have the significance of probabilities existing objectively in the initial state. The nature of the measuring process will, in general, influence $|c_i|^2$. De Broglie mentions the very curious fact that a theory which relies on the concept that every process involving a measurement necessarily disturbs the state of the system, should place ϱ and $|c_i|^2$ on an equal footing and thus ignore its own fundamental assumption.

Finally, de Broglie criticizes the statistical structure of wave mechanics. This statistical structure is in fact, completely at variance with the classical calculus of probabilities. By neglecting the unavoidable efect which every process of measurement of the momentum exercises on the initial state, it disregards the fact that the location probability $|\Psi|^2$ and the momentum probability $|c_p|^2$ cannot simultaneously be real in the initial state. Now,

the very fact that we can restore the classical statistical structure by using the hidden variables shows that the theorem of Neuman according to which it is impossible to give an interpretation of the laws of probabilities of wave mechanics by a picture which introduces hidden variables and which thus permits a sharply defined localisation and momentum to be assigned to the particle is the pseudo-theorem even if the picture proposed in the theory of the double solution does not conform with physical reality.

Criticizing the theory of transformation, de Broglie criticizes the mathematical formalism of its own theory of the double solution and that of Einstein's work on general relativity. He emphasizes that it may be that during the lifetime of transient processes of very short duration the non-linear terms become very significant even outside singular regions. De Broglie recognizes that the theory of the double solution was suggested to him by Einstein's work on general relativity. In the idealised theory of Einstein, there should be no terms in the equations which represent field sources (like the charge and electric current terms in the second part of the Maxwell-Lorentz equations). If field source terms were not excluded, the differential field equations would be inadequate to determine the evolution of the total field, even if the initial and boundary conditions were given. De Broglie admits that he departs somewhat from Einstein's concepts by requiring that the field should be a wave field, so that the fundamental formulaw $W = h\nu$ and $\lambda = h/p$ can be retained. In fact, the mentioned formulae result from Einstein's idealised notion of photon and are the chief contribution of de Broglie and Einstein to science on electrodynamics although they are incompatible with the usual interpretation of quantum mechanics.

4. CONCLUSION

The theory of the double solution of de Broglie with approximation equal to tha of Einstein's general relativity is nearer to physical reality than the theory of non-isotropic bodies proposed by Planck. All these three theories remain silent about the theory of polar molecule presented by Debye in 1912 and honored by Nobel Prize in 1936.

The microwave spectroscopy technique used by radioastronomers caused that the OH radical was detected in the space among the stars and in 1968 ammonia, NH_3 was also detected. So, the forgotten by physicists and electrical and electronic engineers polar molecules of Debye exist even in space between the stars.

The theory of nonlinearity presented by Gupta [1] distinguished for very elegant mathematical formalism seems, however, to be highly debatable, all the more as the fundamentall principles of wave mechanics concerning the meaning of Ψ wave are also debatable as results of the views of Einstein and de Broglie.

REFERENCES

1. M. S. Gupta, *Thermodynamic limits to nonlinearity: Lossless energy — storing systems*, J. Appl. Phys. 50, 569 (1979).
2. M. Planck, *Theory of Electricity and Magnetism*, The Macmillan Company, 1957 (First German Edition, Berlin 1922).
3. L. de Broglie, *The Gurrent Interpretation of Wave Mechanics: A Critical Study*, Elsevier Publishing Company, 1964.

A. MAISON

ZAGADNIENIA NIELINIOWOŚCI W ELEKTRODYNAMICE

Streszczenie

Przedstawiono podstawowe pojęcia de Broglie'a stanowiące krytyczny przegląd czysto probabilistycznej interpretacji mechaniki falowej i fizyki kwantowej. Powołano się na uproszczoną teorię elektryczności sformułowaną przez Plancka w 1922 r. i na ogólną teorię względności Einsteina. We wnioskach odniesiono się do teorii spolaryzowanych molekuł Debye'a.

A. MAISON

PROBLÈMES DE NON-LINÉARITÉ DANS L'ÉLECTRODYNAMIQUE

Résumé

On a présenté les principales notions de Broglie qui constituent une révision critique de l'interprétation probabiliste de la mécanique ondulatoire et de la physique quantique. On a cité la théorie simplifiée d'électricité formulée par Planck en 1922 et la théorie générale de la relativité d'Einstein. Dans les conclusions on s'est rapporté à la théorie des molécules polarisées de Debye.

A. MAISON

PROBLEME DER UNLINEARITÄT BEI DER ELEKTRODYNAMIK

Zusammenfassung

Es wurden die Grundbegriffe von de Broglie dargelegt, die eine kritische Übersicht über die rein probabilistische Interpretation der Wellenmechanik und der Quantenphysik bilden. Man hat sich auf die vereinfachte, von Planck im Jahre 1922 formulierte Elektrizitätstheorie berufen sowie auf die Relativitätstheorie von Einstein. In den Schlußfolgerungen wurde auf die Theorie polarisierter Molekülen von Debey Bezug genommen.

А. МАЙСОН

ВОПРОСЫ НЕЛИНЕЙНОСТИ В ЭЛЕКТРОДИНАМИКЕ

Резюме

Представлены основные понятия Де-Бройля являющиеся критическим обзором чисто пробабилистической интерпретации волновой механики и квантовой физики. Приведены ссылки на упрощенную теорию электричества формулированную Планком в 1922 г. и на общую теорию относительности Эйнштейна. В заключениях обсуждены отношения к теории споляризованных молекул Дебей'я.

ROZPRAWY ELEKTROTECHNICZNE 1982, 28, z. 3-4, ss. 313—324

538.122:538.221

Synteza pola magnetycznego w obecności idealnych elementów ferromagnetycznych

KAZIMIERZ ADAMIAK (KIELCE)

Instytut Elektrotechniki Politechniki Świętokrzyskiej

Otrzymano 20.6.1980

Omówione zostały metody syntezy pola magnetycznego w skończonych, trójwymiarowych obszarach powietrznych w przypadku, gdy w przestrzeni otaczającej ten obszar znajdują się idealne elementy ferromagnetyczne. Założone pole magnetyczne uzyskuje się przez dobór:

— rozkładu gęstości prądu
— geometrii uzwojenia
— kształtu ferromagnetyku.

W przypadku pierwszym wykonano obliczenia numeryczne przykładu syntezy pola jednorodnego w obszarze leżącym na płaszczyźnie symetrii układu płaskorównoległego.

1. WPROWADZENIE

Synteza pola magnetycznego, zagadnienie odwrotne do analizy, polega na poszukiwaniu konstrukcji wytwarzającej, w pewnym obszarze, pole o założonym rozkładzie. Przeprowadzenie syntezy pola jest zadaniem trudnym i dlatego, mimo dużego znaczenia aplikacyjnego, zostało rozwiązane jedynie w najprostszych przypadkach. We wszystkich dotychczasowych publikacjach na temat syntezy pól magnetycznych przyjmowano założenie o jednorodności magnetycznej środowiska, tzn. zakładano jednakową przenikalność magnetyczną w każdym punkcie rozpatrywanej przestrzeni. W praktyce oznacza to, że rozważa się pola w obszarach powietrznych wytwarzane za pomocą uzwojeń bezrdzeniowych. Założenie takie jest do przyjęcia tylko przy wytwarzaniu bardzo słabych i bardzo silnych (indukcja wyższa od indukcji nasycenia ferromagnetyków) pól. W przypadku pierwszym nie ma potrzeby stosowania ferromagnetyków, ponieważ słabe pola można uzyskać stosując tylko uzwojenia bezrdzeniowe. W przypadku drugim konieczne jest zastosowanie uzwojeń nadprzewodzących. Elementy ferromagnetyczne są w takim przypadku niepożądane, ponieważ pracując w nasyceniu, zmniejszałyby wartość indukcji magnetycznej w użytecznym obszarze. Dla przypadków pośrednich zastosowanie ferromagnetyków jest jednak jak najbardziej celowe.

Dopuszczając obecność elementów ferromagnetycznych należy rozróżnić dwa przypadki:

—rozkład pola w rozpatrywanym obszarze powietrznym zależy prawie wyłącznie od

kształtu ferromagnetyku, nie zależy natomiast praktycznie od ukształtowania uzwojeń wzbudzających (np. pole w szczelinie nabiegunników),

— rozkład pola zależy od kształtu zarówno ferromagnetyku, jak też i uzwojeń wzbudzających.

W przypadku pierwszym synteza może polegać tylko na poszukiwaniu geometrii ferromagnetyku. Po wprowadzeniu potencjału skalarnego pola magnetycznego problem może być sprowadzony do rozwiązania odwrotnego zagadnienia brzegowego typu Dirichleta dla równania Laplace'a [1].

W tej pracy rozważany będzie przypadek drugi. Problem zostanie sprowadzony do rozwiązania układu dwu równań całkowych, z których jedno będzie całkowym równaniem Fredholma I rodzaju. Z powodu złego uwarunkowania do rozwiązania zastosowana zostanie metoda regularyzacji [2].

2. MODEL MATEMATYCZNY ZAGADNIENIA

Niech w trójwymiarowej przestrzeni Euklidesa R^3 dane będą trzy rozłączne obszary: Ω, D i S. W obszarze Ω dokonywana będzie synteza pola magnetycznego, obszar D stanowi uzwojenie z prądem, a S jest elementem ferromagnetycznym. Zakłada się, że przenikalność magnetyczna μ wynosi

$$\mu = \begin{cases} \mu_o & P \in R^3 \setminus S, \\ \infty & P \in S; \end{cases}$$

Ω, D i S mogą być obszarami wielospójnymi, np.

$$S = \bigcup_{i=1}^{n} S_i,$$

$$\bigwedge_{i,j} S_i \cap S_j = \phi.$$

Pole magnetyczne jest opisywane równaniem różniczkowym cząstkowym

$$\nabla^2 \mathbf{A} = \begin{cases} 0 & P \in R^3 \setminus D, \\ -\mu_o \mathbf{J} & P \in D, \end{cases} \tag{1}$$

gdzie:

$\mathbf{A}$ — potencjał wektorowy pola magnetycznego

$$\mathbf{B} = \mathrm{rot}\mathbf{A}, \tag{2}$$

$\mathbf{B}$ — wektor indukcji pola magnetycznego,

$\mathbf{J}$ — wektor gęstości prądu w obszarze D.

Spełniony być musi również warunek brzegowy

$$\left.\frac{\partial \mathbf{A}}{\partial \mathbf{n}}\right|_{\Gamma} = 0; \tag{3}$$

$\Gamma = \overline{S} \setminus S$ jest brzegiem obszaru S.

Ferromagnetyk zajmujący obszar S można zastąpić okładem prądowym $\boldsymbol{\delta}(R)$, $R \in \Gamma$, ułożonym na brzegu Γ. Wartość tego okładu wyznacza się z warunku (3). Słuszność takiego postępowania wynika z faktu, iż w jednym i drugim przypadku pole jest opisywane tym samym równaniem (1) z tym samym warunkiem brzegowym (3). Ponieważ rozwiązanie zagadnienia brzegowego typu Neumanna dla równań Laplace'a i Poissona jest jednoznaczne, wobec tego rozkład pola w obu przypadkach jest identyczny.

W dowolnym punkcie obszaru Ω indukcja magnetyczna $\mathbf{B}$ wyraża się zależnością

$$\mathbf{B}(P) = \frac{\mu_o}{4\pi}\left[\int_D \frac{\mathbf{J}(Q)\times\mathbf{r}_{PQ}}{|\mathbf{r}_{PQ}|^3}\,dV + \int_\Gamma \frac{\boldsymbol{\delta}(R)\times\mathbf{r}_{PR}}{|\mathbf{r}_{PR}|^3}\,dS\right]; \tag{4}$$

$\mathbf{r}_{ST}$ — wektor z początkiem w punkcie S i końcem w punkcie T,

$\mathbf{C}\times\mathbf{E}$ — iloczyn wektorowy wektorów $\mathbf{C}$ i $\mathbf{E}$.

Zależność (4), nazywana też prawem Biota-Savarta, wynika z ogólnego rozwiązania równania (1) i związku (2). Warunek brzegowy (3) prowadzi do równania

$$\int_D \frac{(\mathbf{J}(Q)\times\mathbf{r}_{RQ})_t}{|\mathbf{r}_{RQ}|^3}\,dV + \int_\Gamma \frac{(\boldsymbol{\delta}(R)\times\mathbf{r}_{R'R})_t}{|\mathbf{r}_{R'R})|^3}\,dS = 0, \tag{5}$$

$(\mathbf{E})_t$ — składowa styczna wektora $\mathbf{E}$.

Równania (4) i (5) w pełni opisują pole magnetyczne w rozpatrywanej przestrzeni, a więc na ich podstawie można również dokonać syntezy pola. Należy w tym celu przyjąć funkcję $\mathbf{B}(P)$, a poszukiwać innych parametrów konstrukcji. Na rozkład pola magnetycznego mają wpływ: $\mathbf{J}(Q)$, D oraz S. Poszukiwanie rozwiązania poprzez zmianę wszystkich trzech parametrów byłoby zbyt skomplikowane, dlatego dwa z nich zadaje się, a poszukuje trzeciego. Otrzymuje się w ten sposób 3 następujące zadania:

— poszukiwanie $\mathbf{J}(Q)$, $Q \in D$, przy założonych D i S,

— poszukiwanie D, przy założonych $\mathbf{J}(Q)$, $Q \in D$ i S; ze względów praktycznych zakłada się przeważnie $\mathbf{J}(Q) = \mathbf{J}_o = \text{const}$,

— poszukiwanie S, przy założonych $\mathbf{J}(Q)$ i D.

Wymienione sposoby zostaną omówione bardziej szczegółowo na przykładzie syntezy pola w obszarze leżącym na płaszczyźnie symetrii układu płaskorównoległego.

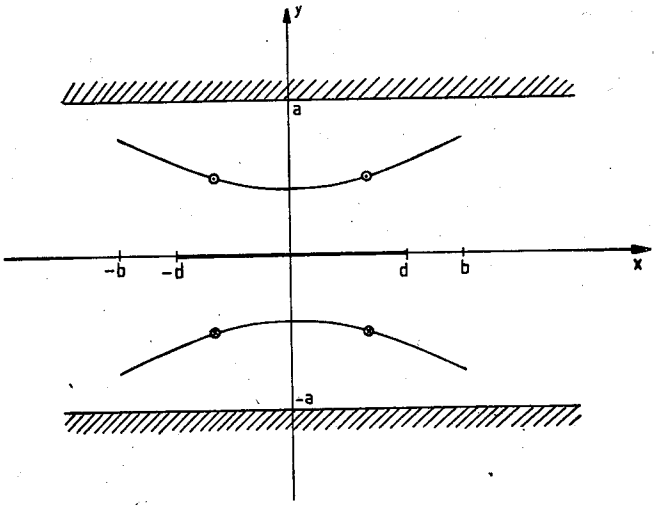
3. SYNTEZA POLA METODĄ POSZUKIWANIA ROZKŁADU GĘSTOŚCI PRĄDU

Do obliczeń przyjęty został model płaskorównoległy, symetryczny względem osi x i antysymetryczny względem osi y z uzwojeniami nieskończenie cienkimi (rys. 1). Ferromagnetyk stanowią dwie półprzestrzenie:

$$\Omega = \{(x, y); \quad y = 0 \wedge x \in (-d, d)\},$$

$$D = \{(x, y); \quad (y = f(x) \vee y = -f(x)) \wedge x \in (-b, b)\},$$

$$S = \{(x, y); \quad |y| \geqslant a\}.$$



Rys. 1. Model przyjęty do syntezy pola magnetycznego

W rozpatrywanym przypadku równania (4) i (5) przyjmą postać

$$b_o(x) = \int_{-b}^{b} \frac{f(t)J(t)\,dt}{f^2(t)+(x-t)^2} + \int_{-\infty}^{\infty} \frac{a\delta(t)\,dt}{a^2+(x-t)^2}, \tag{6a}$$

$$b_o(x) = \frac{2\pi}{\mu_o} B(x), \quad \mathbf{B} = B(x)\mathbf{1}_x,$$

$$x \in (-d, d), \quad \mathbf{J} = J(t)\mathbf{1}_z, \quad \boldsymbol{\delta} = \delta(t)\mathbf{1}_z,$$

$$2\pi\delta(x)+2a\int_{-\infty}^{\infty} \frac{\delta(t)\,dt}{4a^2+(x-t)^2} = \int_{-b}^{b} J(t)\left[\frac{a-f(t)}{[a-f(t)]^2+(x-t)^2} - \right.$$

$$\left. - \frac{a+f(t)}{[a+f(t)]^2+(x-t)^2}\right] dt, \tag{6b}$$

$$x \in (-\infty, \infty).$$

Jest to układ dwu równań całkowych Fredholma: równanie (6a) jest równaniem I rodzaju, (6b) — równaniem II rodzaju. Danymi w tym układzie są parametry a, b, d oraz funkcje $f(t)$ i $b_o(x)$; $J(t)$ jest poszukiwanym rozwiązaniem określającym rozkład gęstości prądu w uzwojeniu, a $\delta(t)$ — pomocniczą niewiadomą określającą rozkład gęstości prądu na płaszczyznach $|y| = a$ równoważny półprzestrzeniom ferromagnetycznym.

Równania całkowe, jak też i ich układy, rozwiązuje się najczęściej, ze względu na prostotę i uniwersalność, metodą algebraizacji. Polega ona na zastąpieniu funkcji ciągłych funkcjami schodkowymi, dzięki czemu całkowanie sprowadza się do sumowania i w efekcie zamiast równań całkowych otrzymuje się układy równań algebraicznych. Rozwiązanie układów otrzymanych z równań II rodzaju nie nastręcza żadnych problemów obliczeniowych, natomiast rozwiązanie tą metodą równań I rodzaju napotyka na trudności.

Jest to związane z tym, że otrzymany układ równań algebraicznych jest przeważnie źle uwarunkowany, tzn. wyznacznik główny układu jest bliski zeru (choć różny od zera). Powoduje to niestabilność obliczeń spowodowaną błędami zaokrągleń. Aby tego uniknąć należy dokonać regularyzacji równania [2, 3].

W sposób identyczny, jak wyżej opisany, otrzymuje się równania pozwalające dokonać syntezy pola za pomocą uzwojeń o skończonych wymiarach, składających się z pewnej liczby sekcji ze zmniejszającą się gęstością prądu [3, 4].

4. SYNTEZA POLA METODĄ POSZUKIWANIA KSZTAŁTU UZWOJEŃ

Sposób syntezy pola podany powyżej posiada dość istotną wadę, ponieważ prowadzi do niewygodnych w praktyce rozwiązań konstrukcyjnych: uzwojenia cienkiego, przez co wytwarzane pole jest słabe, lub sekcyjnego, kłopotliwego w eksploatacji ze względu na skomplikowane zasilanie. Wady tej nie posiadają uzwojenia o stałej gęstości prądu i odpowiednio dobranej geometrii [5, 6]. Dla przykładu, syntezę pola za pomocą uzwojenia nieskończenie cienkiego, z równomierną gęstością prądu, lecz odpowiednio ukształtowanego, można przeprowadzić rozwiązując równanie (6) z niewiadomą $f(t)$ i założonym $J(t) = J_0 = \text{const}$. Otrzymuje się wówczas

$$b_0(x) = \int_{-b}^{b} \frac{f(t)\,dt}{f^2(t)+(x-t)^2} + \int_{-\infty}^{\infty} \frac{a\delta(t)\,dt}{a^2+(x-t)^2}, \tag{7a}$$

$$b_0(x) = \frac{2\pi}{\mu_0 J_0} B(x), \qquad x \in (-d, d),$$

$$2\pi\delta(x) + 2a \int_{-\infty}^{\infty} \frac{\delta(t)\,dt}{4a^2+(x-t)^2} = \int_{-b}^{b} \left[\frac{a-f(t)}{[a-f(t)]^2+(x-t)^2} - \frac{a+f(t)}{[a+f(t)]^2+(x-t)^2} \right] dt, \tag{7b}$$

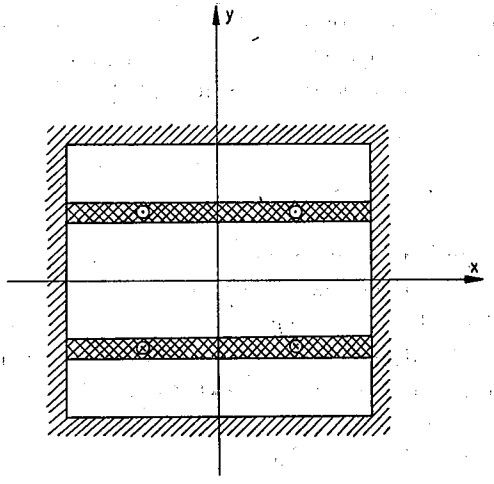
$$x \in (-\infty, \infty).$$

Układ równań (7) jest liniowy ze względu na niewiadomą $\delta(t)$ i nieliniowy ze względu na $f(t)$. Rozwiązywanie układów nieliniowych jest zadaniem trudniejszym od rozwiązywania układów liniowych. Najczęstszą metodą jest zastosowanie procesów iteracyjnych [7], odpowiednio zmodyfikowanych, aby ustrzec się przed niestabilnością obliczeń [8].

5. SYNTEZA POLA METODĄ POSZUKIWANIA KSZTAŁTU FERROMAGNETYKU

Uzwojenie wzbudzające jest, w tej metodzie, z góry założone i to w postaci najwygodniejszej zarówno do wykonania jak też i w eksploatacji. Jest to najczęściej cewka o przekroju prostokątnym z jednorodną gęstością prądu. Synteza polega na znalezieniu równania opisującego kształt ferromagnetyku. Problem, dla pola jednorodnego, posiada rozwią-

zanie trywialne, mianowicie, uzwojenie całkowicie otoczone ferromagnetykiem, przedstawione na rys. 2. Oczywiście ze względu na niedostępność obszaru z polem jednorodnym konstrukcja taka rzadko znajduje zastosowanie praktyczne.



Rys. 2. Konstrukcja wytwarzająca jednorodne pole magnetyczne

Równania (4) i (5) sprowadzają się, w rozpatrywanym przypadku, do układu równań różniczkowo-całkowych

$$b_o(x) = J_o \int_{-b}^{b} \frac{f(t)\,dt}{f^2(t)+(x-t)^2} + \int_{-\infty}^{\infty} \frac{\delta(t)g(t)\,dt}{g^2(t)+(x-t)^2}, \tag{8a}$$

$$b_o(x) = \frac{2\pi}{\mu_0} B(x), \qquad x \in (-d, d),$$

$$J_o \int_{-b}^{b} \left[\frac{\sin r_1(x,t)}{\sqrt{[g(x)-f(t)]^2+(x-t)^2}} - \frac{\sin r_2(x,t)}{\sqrt{[g(x)+f(t)]^2+(x-t)^2}} \right] dt =$$

$$= \int_{-\infty}^{\infty} \left[\frac{\delta(t)\sin r_3(x,t)}{\sqrt{[g(x)+g(t)]^2+(x-t)^2}} - \frac{\delta(t)\sin r_4(x,t)}{\sqrt{[g(x)-g(t)]^2+(x-t)^2}} \right] dt, \tag{8b}$$

$$r_i(x,t) = \operatorname{arctg} g'(x) + \operatorname{arctg} s_i(x,t),$$

$$s_1(x,t) = \frac{f(t)-g(x)}{t-x},$$

$$s_2(x,t) = \frac{f(t)+g(x)}{t-x},$$

$$s_3(x,t) = \frac{g(t)+g(x)}{t-x},$$

$$s_4(x,t) = \frac{g(t)-g(x)}{t-x},$$

liniowego ze względu na $\delta(t)$ i J_o, oraz nieliniowego ze względu na $g(t)$. Dodatkową, w porównaniu z (7), trudnością pojawiającą się przy rozwiązywaniu otrzymanego układu równań jest fakt występowania pod znakiem całki pochodnej niewiadomej $g'(x)$. Stosując do obliczeń metodę różnic skończonych, pochodną aproksymuje się wyrażeniem

$$g'(x) = \frac{g(x+\Delta x)-g(x-\Delta x)}{2\Delta x}. \tag{9}$$

Powoduje to konieczność zadania warunków brzegowych: $g(x)$ lub $g'(x)$ przy $x \to \infty$.

6. WYNIKI NUMERYCZNE SYNTEZY POLA JEDNORODNEGO

Układ równań (6) został rozwiązany za pomocą emc ODRA 1305. Założono, że $b_o(x)$ jest funkcją parzystą, wobec tego $J(t)$ i $\delta(t)$ będą również funkcjami parzystymi. Układ równań (6) można więc zapisać w równoważnej postaci

$$b_o(x) = \int_0^b \left[\frac{1}{f^2(t)+(x-t)^2} + \frac{1}{f^2(t)+(x+t)^2}\right] f(t)J(t)\,dt + \\ + \int_0^\infty a\delta(t)\left[\frac{1}{a^2+(x-t)^2} + \frac{1}{a^2+(x+t)^2}\right] dt, \tag{10a}$$

$$x \in (0, d),$$

$$2\pi\delta(x) + 2a\int_0^\infty \delta(t)\left[\frac{1}{4a^2+(x-t)^2} + \frac{1}{4a^2+(x+t)^2}\right] dt = \\ = \int_0^b J(t)\left[\frac{a-f(t)}{[a-f(t)]^2+(x-t)^2} + \frac{a-f(t)}{[a-f(t)]^2+(x+t)^2} - \right. \\ \left. - \frac{a+f(t)}{[a+f(t)]^2+(x-t)^2} - \frac{a+f(t)}{[a+f(t)]^2+(x+t)^2}\right] dt, \tag{10b}$$

$$x \in (0, \infty).$$

W równaniach (10) występują całki niewłaściwe z granicami całkowania $(0, \infty)$. Aby obliczenia numeryczne były możliwe, konieczne jest pozbycie się tej niewłaściwości. Jednym ze sposobów jest założenie zamiast ∞ pewnej wartości skończonej. Wadą jest niemożliwość analitycznego oszacowania tej wartości. Przyjęcie granicy zarówno zbyt małej jak i zbyt dużej może być przyczyną poważnych błędów. Należałoby więc stosować metodę kolejnych prób.

Innym podejściem do tego problemu jest zamiana zmiennych zgodnie z zależnością

$$s = 1 - e^{-t}. \tag{11}$$

Przy zmianach t w granicach od 0 do ∞, s będzie zmieniać się w przedziale (0, 1)

$$\int_0^{\infty} K(x,t)\,\delta(t)\,dt = \int_0^{1} K(x,s)\,\delta(s)\,\frac{ds}{1-s}. \tag{12}$$

Funkcja podcałkowa po prawej stronie wyrażenia (12) zawiera osobliwość dla $s \to 1$. Dlatego do obliczeń wykorzystano zmodyfikowaną postać równania (12) biorąc zamiast prawej strony wyrażenie

$$\int_0^{1-\varepsilon} \frac{K(x,s)}{1-s}\,\delta(s)\,ds; \tag{12a}$$

ε — jest pewną małą liczbą dodatnią, której wartość została wyznaczona iteracyjną metodą prób.

Zastosowanie metody różnic skończonych pozwala sprowadzić układ równań (10) do układu równań algebraicznych

$$\begin{aligned} &\sum_{i=1}^{N} a_{ij}J_i + \sum_{i=N+1}^{N+M} c_{ij}\delta_i = b_j, \quad j = 1,\ldots,N,\\ &\sum_{i=1}^{N} d_{ij}J_i + \sum_{i=N+1}^{N+M} e_{ij}\delta_i = 0, \quad j = N+1,\ldots,N+M, \end{aligned} \tag{13}$$

gdzie:

$$a_{ij} = \left[\frac{1}{f^2(u_i)+(x_j-u_i)^2} + \frac{1}{f^2(u_i)+(x_j+u_i)^2}\right] f(u_i)+\alpha_{ij},$$

$$c_{ij} = \frac{a}{1-s_i}\left[\frac{1}{a^2+(x_j-t_i)^2} + \frac{1}{a^2+(x_j+t_i)^2}\right],$$

$$d_{ij} = \frac{2a}{1-s_i}\left[\frac{1}{4a^2+(y_j-u_i)^2} + \frac{1}{4a^2+(y_i+u_i)^2}\right]+\beta_{ij},$$

$$e_{ij} = \frac{a+f(s_i)}{[a+f(s_i)]^2+(y_j+t_i)^2} + \frac{a+f(s_i)}{[a+f(s_i)]^2+(y_j+t_i)^2} -$$

$$- \frac{a-f(s_i)}{[a-f(s_i)]^2+(y_j-t_i)^2} - \frac{a-f(s_i)}{[a-f(s_i)]^2+(y_j+t_i)^2},$$

$$b_j = b_o(x_j),$$

$$u_i = (i-0{,}5)h_1, \quad h_1 = b/N,$$

$$x_j = (j-0{,}5)h_2, \quad h_2 = d/N,$$

$$s_i = (i-0{,}5)h_3, \quad h_3 = 1/M,$$

$$t_i = \ln\frac{1}{1-s_i}$$

$$y_j = (j-0{,}5)h_3,$$

$$\alpha_{ij} = \begin{cases} 0 & i \neq j, \\ \alpha & i = j, \end{cases}$$

$$\beta_{ij} = \begin{cases} 0 & i \neq j, \\ 2\pi & i = j; \end{cases}$$

α jest tzw. parametrem regularyzacji, którego wartość znajduje się metodą kolejnych iteracji.

Obliczenia numeryczne omawianego problemu przeprowadzono dla: $b_0(x) = 1{,}0 =$ $=$ const, $a = 1{,}0$, $d = 1{,}0$ oraz $f(t) = c =$ const i różnych kombinacji pozostałych wymiarów geometrycznych. Wektor rozwiązań $\{J_i\}$ posiada bardzo dużo elementów zerowych. Elementy niezerowe są nieliczne (tabl. 1). Optymalna wartość parametru α wynosi we wszystkich przypadkach $\alpha = 10^{-5}$. Ze względu na ograniczoną pojemność pamięci maszyny założono $N+M = 100$. Najdokładniejsze wyniki otrzymano dla $N = M = 50$.

Tablica 1

Rozkłady gęstości prądu w uzwojeniu wytwarzającym jednorodne pole magnetyczne

a)

$b = 1{,}0$, $c = 1{,}0$,
$\alpha = 10^{-5}$, $N = 50$,
$M = 50$, $\Delta = 0{,}44\%$

i	J_i
1	−2,4613
29	87,6280
40	−138,9721
41	−16,7269
42	−10,5333
50	130,5938

b)

$b = 1{,}0$, $c = 1{,}0$,
$\alpha = 10^{-5}$, $N = 40$,
$M = 60$, $\Delta = 0{,}55\%$

i	J_i
1	8,1502
13	−2,2564
14	−16,5344
23	74,0903
32	−113,2107
40	93,6228

Dokładność otrzymanych wyników porównano za pomocą błędu zdefiniowanego w następujący sposób

$$\Delta = \max_{x\in(0,b)} b(x) - \min_{x\in(0,b)} b(x); \tag{14}$$

$b(x)$ — rzeczywisty przebieg indukcji magnetycznej dla obliczonego rozkładu gęstości prądu.

Wartości Δ dla różnych wariantów wymiarów b i c zestawiono w tabl. 2. Zależność Δ od c (przy $b = 1{,}0$) jest niemonotoniczna, natomiast przy rosnącym b (dla $c = 1{,}0$) Δ szybko maleje do zera.

Tablica 2

Błąd jednorodności pola magnetycznego przy zmianie parametrów b i c

b	1,0							1,25	1,5	2,0	3,0
c	0,5	0,6	0,7	0,8	0,85	0,9	0,95	1,0			
Δ	0,38	1,07	0,63	0,47	0,38	0,56	0,54	0,44	0,13	0,04	0,01	0,00

Otrzymane wyniki, zarówno rozkład gęstości prądu jak też i dokładność realizacji pola jednorodnego, są zbliżone do otrzymanych przy syntezie pola na osi symetrii solenoidu walcowego [4].

7. WERYFIKACJA OTRZYMANYCH WYNIKÓW

Przyjęty do obliczeń model zagadnienia pozwala sprawdzić otrzymane wyniki przez rozwiązanie równania Laplace'a z warunkami brzegowymi Neumanna.

Potencjał wektorowy pola magnetycznego w obszarach bezprądowych jest opisywany równaniem (1); **A** posiada tylko jedną składową, wzdłuż osi z

$$\mathbf{A} = A(x, y)\mathbf{1}_z. \tag{15}$$

Wobec tego równanie (1) sprowadza się do

$$\frac{\partial^2 A(x, y)}{\partial x^2} + \frac{\partial^2 A(x, y)}{\partial y^2} = 0. \tag{16}$$

Rozwiązaniem równania (16) są:

— w obszarze $D_1 = \{(x, y);\ |y| \leqslant c\}$

$$A_1 = \int_0^\infty A(\lambda)\operatorname{sh}\lambda y \cos\lambda x\, d\lambda, \tag{17a}$$

— w obszarze $D_2 = \{(x, y);\ c \leqslant y \leqslant a\}$

$$A_2 = \int_0^\infty [B(\lambda)\operatorname{sh}\lambda y + C(\lambda)\operatorname{ch}\lambda y]\cos\lambda x\, d\lambda; \tag{17b}$$

$A(\lambda)$, $B(\lambda)$ i $C(\lambda)$ są nieznanymi stałymi całkowania, które należy wyznaczyć z warunków brzegowych. Wynikają one z założenia o nieskończonej przenikalności ferromagnetyka, dzięki czemu zeruje się składowa styczna wektora indukcji, oraz z ciągłości rozwiązania na granicy obszarów D_1 i D_2

$$\left.\frac{\partial A_2}{\partial y}\right|_{y=a} = 0, \tag{18a}$$

$$A_1|_{y=c} = A_2|_{y=c}, \tag{18b}$$

$$\left.\left(\frac{\partial A_1}{\partial y} - \frac{\partial A_2}{\partial y}\right)\right|_{y=c} = \begin{cases} 0 & |x| > b, \\ -\mu J(x) & |x| \leqslant b. \end{cases} \tag{18c}$$

Po wykonaniu obliczeń otrzymuje się

$$A_1 = -\frac{\mu}{\pi b}\int_0^\infty \left(\int_0^b J(t)\cos\lambda t\, dt\right)\lambda^{-1}\operatorname{sh}\lambda c(\operatorname{th}\lambda a - \operatorname{cth}\lambda c). \tag{19}$$

Składowa wzdłuż osi x wektora indukcji magnetycznej wyrazi się zależnością

$$B_x = \frac{\partial A_1}{\partial y} = -\frac{\mu}{\pi b}\int_0^\infty \left(\int_0^b f(t)\cos\lambda t\, dt\right)\operatorname{sh}\lambda c(\operatorname{th}\lambda a - \operatorname{cth}\lambda c)\operatorname{ch}\lambda y\cos\lambda x\, d\lambda. \tag{20}$$

Wartości funkcji $J(t)$, wyznaczone metodą równań całkowych, zostały podstawione, przy założeniu $y = 0$, do wzoru (20). Przebieg otrzymanych w ten sposób funkcji $B(x)$ jest bardzo zbliżony (z dokładnością większą od 1%) do obliczonych przez sprawdzenie

układu równań (13). Pewna różnica jest nieunikniona ze względu na błąd aproksymacji funkcji ciągłych funkcjami schodkowymi przy zastępowaniu równań całkowych równaniami algebraicznymi.

8. ZAKOŃCZENIE

Przedstawiona w pracy metoda syntezy pól magnetycznych w obszarach zawierających elementy ferromagnetyczne jest metodą uniwersalną, ponieważ można ją zastosować w konstrukcjach o dowolnej geometrii. Jedynym uproszczeniem jest założenie o nieskończonej przenikalności magnetycznej ferromagnetyku. Dla współczesnych materiałów ferromagnetycznych jest to warunek łatwy do spełnienia, jeśli tylko ferromagnetyk nie jest nasycony.

LITERATURA

1. K. Adamiak, *Odwrotne zagadnienia brzegowe typu Dirichleta dla równonia Laplace'a*, Rozprawy Elektrotechniczne (w druku).
2. M. M. Ławrentiew, *Ob niekotorych niekorektnych zadaczach matematiczeskoj fiziki*, Nowosibirsk 1962.
3. K. Adamiak, *Synthesis of Homogeneous Magnetic Field in Internal Region of Cylindrical Solenoid*, Arch. f. Elektr., 62, 75—79 (1980).
4. K. Adamiak, *Synteza pola magnetycznego w obszarach dwuwymiarowych za pomocą uzwojeń bezrdzeniowych*, Rozpr. Elektr. (w druku).
5. K. Adamiak, *Determination of the Ampere Turn Disposition in a Solenoid Generating Homogeneous Magnetic Field*, Scientia Electrica, 24, 48—60 (1978).
6. K. Adamiak, *Tichonov's Regularization Method Applied to the Magnetic Field on a Solenoid Axis Synthesis*, Arch. f. Elektr., 59, 19—22 (1977).
7. J. M. Ortega, W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, New York 1970.
8. A. N. Tichonow, W. B. Głasko, *Primienienie metoda regularizacji w nieliniejnych zadaczach*, Żurn. Wyczisl. Matem. i Matem. Fiz., 5, 463—473 (1965).

K. ADAMIAK

MAGNETIC FIELD SYNTHESIS IN PRESENCE OF IDEAL FERROMAGNETIC ELEMENTS

Summary

The paper discusses the methods of magnetic field synthesis in finite, three dimensional aerial areas when there are ideal ferromagnetic elements in the space surrounding the area. The magnetic field assumed is obtained by the selection of:
— current density distribution in the winding,
— winding geometry,
— shape of ferromagnetic.
In the first case, numerical calculations were made of an example of homogenous field synthesis in an area lying in the plane of symmetry of a plano-parallel system.

K. ADAMIAK

SYNTHÈSE DU CHAMP MAGNÈTIQUE EN PRÈSENCE DES ÈLÈMENTS FERROMAGNÈTIQUES IDÉALS

Résumé

On a traité les méthodes de la synthèse du champ magnétique dans les espaces finis tridimensionnels dans le cas où dans la zone environnant cet espace se trouvent des éléments ferromagnétiques idéals. Le champ magnétique admis est déterminé par le choix de:
— répartition de densité du courant,
— géométrie de l'enroulement,
— forme du corps ferromagnétique.
Pour le premier cas on a fait des calculs numériques d'un exemple de la synthèse du champ homogène dans l'étendue située dans le plan de symétrie du système plan-parallèle.

K. ADAMIAK

SYNTHESE EINES MAGNETISCHEN FELDES UNTER ANWESENHEIT VON IDEALEN FERROMAGNETISCHEN ELEMENTEN

Zusammenfassung

Erörtert werden die Methoden für die Synthese eines Magnetfeldes im endlichen, dreidimensionalen Luftraum in solchem Fall, wenn sich in dem diesen Raum umgebenden Bereich ideale ferromagnetische Elemente befinden. Das gegebene Magnetfeld wird durch die Wahl folgender Elemente gewonnen:
— Zerlegung der Stromdichte,
— Wicklungsgeometrie,
— Form des ferromagnetischen Stoffes.
Für den ersten Fall wurde eine numerische Berechnung eines Beispiels für die Synthese eines homogenen Feldes in demjenigen Raum durchgeführt, der auf der Symmetriefläche des flachparallelen Systems liegt.

К. АДАМЯК

СИНТЕЗ МАГНИТНОГО ПОЛЯ В ПРИСУТСТВИИ ИДЕАЛЬНЫХ ФЕРРОМАГНИТНЫХ ЭЛЕМЕНТОВ

Резюме

Обсуждены методы синтеза магнитного поля в конечных трехмерных воздушных областях в случаях когда в пространстве окружающем эти области находятся идеальные ферромагнитные элементы. Требуемое магнитное поле получается путем подбора:
— распределения плотности тока в обмотке,
— геометрии обмотки,
— формы ферромагнетика.
В первом случае проведены числовые расчеты примера синтеза однородного поля в области расположенной в плоскоети симметрии плоскопараллельной системы.

ROZPRAWY ELEKTROTECHNICZNE 1982, 28, z. 3-4, ss. 325—341

621.3.011.21

Podstawy rekurencyjnej metody obliczania admitancji międzywęzłowych

STANISŁAW JANKOWSKI (WARSZAWA)

Instytut Podstaw Elektroniki Politechniki Warszawskiej

Otrzymano 16.10.1980

W artykule przedstawiono podstawy rekurencyjnej metody obliczania admitancji międzywęzłowych dla ciągu liniowych pasywnych sieci elektrycznych. Wykazano, że metodę można stosować wyłącznie w przypadku sieci zwanych w pracy sieciami warstwowymi. Sieci te powstają przez agregację w kolejnych krokach rekurencyjnych pewnych sieci, zwanych warstwami, w taki sposób, aby połączenia gałęziowe występowały tylko między kolejnymi warstwami, np. n-tą i $(n+1)$-szą. Sieci warstwowe mogą być regularne lub nieregularne topologicznie, planarne lub nieplanarne. Struktura blokowa macierzy admitancyjnych sieci warstwowych pozwala zredukować podczas obliczeń macierz admitancyjną sieci n-tej do macierzy admitancyjnej dwóch ostatnich warstw. A zatem algorytmy rekurencyjne obliczania admitancji międzywęzłowych są efektywne dla sieci zawierających nawet wiele tysięcy węzłów i gałęzi. W pracy podano również przykłady ciągów sieci warstwowych. Metoda rekurencyjna może być stosowana do badania wielkich sieci o losowych rozkładach admitancji gałęziowych w celu modelowania elektrycznych właściwości niejednorodnych ośrodków ziarnistych.

1. WPROWADZENIE

Przedstawiona w pracy rekurencyjna metoda obliczania admitancji międzywęzłowych prowadzi do efektywnych algorytmów, które mogą być stosowane dla określonej w pracy klasy sieci elektrycznych, zawierających nawet wiele tysięcy węzłów i gałęzi. Metoda ta polega na rozszerzeniu postępowania znanego w teorii linii łańcuchowej. Postępowanie takie pozwala wyznaczyć właściwości transmisyjne linii oraz admitancje wejściową i wyjściową dla linii złożonej z $(n+1)$ ogniw przez wykorzystanie obliczeń dla linii zbudowanej z n ogniw. W ten sposób można kontynuować obliczenia dla dowolnie wielu ogniw [2].

Postępowanie rekurencyjne może być stosowane wtedy, gdy istnieje pewien ciąg obiektów $\{A_n\}$, przy czym element n-ty ciągu jest określony za pomocą k poprzedzających go elementów

$$A_n = G(A_{n-1}, A_{n-2}, \ldots, A_{n-k}), \quad n > k, \tag{1}$$

gdzie G oznacza funkcję określoną na elementach ciągu $\{A_n\}$; liczba k nazywana jest rzędem zależności rekurencyjnej.

Dobre określenie zależności (1) wymaga podania wartości k elementów początkowych: $A_1, A_2, \ldots, A_k$.

Idea przedstawionej w pracy metody jest następująca. Przyjmuje się istnienie pewnego ciągu liniowych pasywnych sieci elektrycznych $\{S_n\}$. Sieć S_{n+1} otrzymuje się przez dołączenie do sieci S_n pewnej sieci ΔS_{n+1}, zwanej warstwową. Warunki, określające strukturę sieci ciągu $\{S_n\}$ podano w punkcie 2.

W pracy przyjęto, że rozpatrywane sieci będą analizowane metodą potencjałów węzłowych, wobec tego ciągowi sieci $\{S_n\}$ odpowiadać będzie ciąg macierzy admitancyjnych $\{Y_n\}$, których szczególna postać, wynikająca ze struktury sieci, prowadzi do rekurencyjnego obliczania wyznaczników i minorów głównych tych macierzy, niezbędnych do wyznaczenia pewnych admitancji międzywęzłowych.

Rekurencyjny algorytm obliczania tych admitancji jest możliwy przy założeniu, że wszystkie sieci ciągu $\{S_n\}$ mają tak samo usytuowany węzeł odniesienia R, natomiast drugi węzeł wyróżniony P może być wybrany na 2 sposoby:

a) jest to węzeł należący do sieci początkowej S_o i nie zmienia się w zależności od numeru sieci,

b) w n-tym kroku rekurencyjnym węzeł ten należy do przyłączonej w tym kroku sieci ΔS_n.

2. REKURENCYJNE OKREŚLENIE CIĄGU SIECI ELEKTRYCZNYCH $\{S_n\}$

Zgodnie z ogólną postacią zależności rekurencyjnej (1) należy określić zasady tworzenia ciągu sieci elektrycznych $\{S_n\}$. Niech siecią początkową ciągu będzie pewna sieć S_o o dowolnej topologii, złożona z (m_o+1) węzłów i g_o gałęzi. Liczba węzłów niezależnych m_o sieci S_o może być równa dowolnej liczbie naturalnej, począwszy od 1 aż do takiej wartości, dla której obliczanie admitancji zastępczej między węzłem odniesienia R, a węzłem P o numerze m_p $(m_p \leqslant m_o)$ może być wykonane przez zastosowanie dowolnej znanej metody.

Następnie tworzy się nową sieć elektryczną S_1 przez dołączenie do sieci S_o pewnej sieci ΔS_1, zawierającej r_1 węzłów i Δg_1 gałęzi, przy czym dołączenie sieci ΔS_1 można przeprowadzić w sposób dowolny. Należy jednak zachować ten sam węzeł odniesienia R, jak w sieci S_o.

Przyjmuje się, że numery węzłów sieci S_o nie zmieniają się, toteż numery węzłów dołączonych sieci ΔS_1 są zawarte w granicach od (m_o+1) do (m_o+r_1). Liczba węzłów niezależnych sieci S_1 równa się

$$m_1 = m_o + r_1 . \tag{2}$$

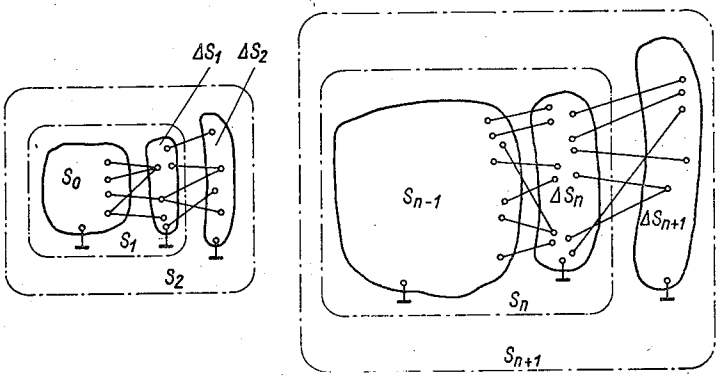
Sieć S_2 ciągu $\{S_n\}$ powstaje przez dołączenie do sieci S_1 pewnej nowej sieci ΔS_2, złożonej z r_2 węzłów i Δg_2 gałęzi. Jednak począwszy od sieci S_2 dołączenie nowej sieci ΔS_2 nie może być przeprowadzone w sposób dowolny.

W celu otrzymania związków rekurencyjnych istotne jest spełnienie wymagania, aby dołączenie sieci ΔS_2 do sieci S_1 zostało dokonane w taki sposób, aby połączenia gałęziowe występowały tylko między węzłami należącymi do sieci ΔS_2 i węzłami należącymi do sieci ΔS_1, dołączonej w poprzednim kroku. Oznacza to, że nie występują bezpośrednie połączenia między węzłami należącymi do sieci początkowej S_o a węzłami nowej sieci ΔS_2, z wyjątkiem węzła odniesienia R, który jest wspólny dla wszystkich sieci ciągu.

Sieć S_{n+1} ciągu $\{S_n\}$ dla $n \geqslant 2$ otrzymuje się przez dołączenie do sieci S_n pewnej sieci ΔS_{n+1} w taki sposób, aby nie występowały połączenia bezpośrednie między należącymi do sieci S_{n-1} a węzłami sieci ΔS_{n+1}. A więc połączenia są możliwe wyłącznie między węzłami sieci ΔS_n i węzłami sieci ΔS_{n+1}. Obrazowo można powiedzieć, że tworzenie ciągu sieci $\{S_n\}$ odbywa się przez dołączenie pewnych „warstw" w kolejnych krokach rekurencyjnych. Dlatego otrzymywane w sposób wyżej opisany sieci ciągu $\{S_n\}$ będą nazywane sieciami warstwowymi, a sieci dołączone ΔS_n — warstwami.

W opisanej metodzie tworzenia sieci warstwowych rekurencja sięga do numeru sieci mniejszej o 1 — jest więc rekurencją pierwszego rzędu. Istnieje również możliwość określenia rekurencyjnej metody tworzenia sieci warstwowych rzędu wyższego — nie będzie ona jednak omawiana w pracy.

Na rys. 1 przedstawiono rekurencyjne powstawanie ciągu elektrycznych sieci warstwowych $\{S_n\}$.



Rys. 1. Rekurencyjne powstawanie ciągu elektrycznych sieci warstwowych $\{S_n\}$

3. STRUKTURA MACIERZY ADMITANCYJNYCH SIECI WARSTWOWYCH

W pracy przyjęto, że sieci warstwowe $\{S_n\}$ będą analizowane metodą potencjałów węzłowych, a więc macierze podstawowe tych sieci będą macierzami admitancyjnymi. Utworzą one pewien ciąg macierzy $\{Y_n\}$. Metodę potencjałów węzłowych wybrano z powodu dobrej i prostej interpretacji tych macierzy w zastosowaniach fizycznych. Rozważania prowadzono w ramach metody liczb zespolonych.

Sformułowane uprzednio warunki dotyczące struktury sieci ciągu $\{S_n\}$ mają oczywiście wpływ na postać ciągu odpowiednich macierzy $\{Y_n\}$. Jak wiadomo, macierz admitancyjna jest tzw. macierzą podstawową danej sieci wtedy, gdy do analizy jest stosowana metoda potencjałów węzłowych. Macierz admitancyjna, odpowiadająca liniowej sieci biern· j, jest symteryczna, przy czym na głównej przekątnej występują admitancje własne węzłów, np. element y_{ii} macierzy Y_n równa się sumie admitancji gałęziowych połączonych w węźle i-tym. Poza przekątną główną znajdują się tzw. admitancje wzajemne węzłów, tzn. element y_{ij} macierzy Y_n równy jest admitancji gałęziowej ze znakiem minus, łączącej węzeł i-ty z węzłem j-tym. W konsekwencji, gdy między parą węzłów (i, j) nie ma połączenia, to element y_{ij} oraz y_{ji} jest równy zeru. Zakłada się przy tym brak sprzężeń magnetycznych w sieci.

Zgodnie z określeniem, sieci początkowej S_o ciągu $\{S_n\}$, odpowiadająca jej macierz admitancyjna Y_o spełnia po prostu ogólne warunki dla macierzy admitancyjnej sieci biernej. Jeśli S_o jest siecią zupełną, to macierz Y_o zawiera wyłącznie elementy różne od zera. Macierrz Y_o ma wymiary $(m_o \times m_o)$, gdzie m_o — liczba węzłów niezależnych macierzy admitancyjnej sieci początkowej S_o.

Struktura macierzy admitancyjnej Y_1 sieci S_1 podkreśla istotne cechy topologiczne tej sieci. W macierzy Y_1 można wyróżnić następujące macierze blokowe

$$Y_1 = \left[\begin{array}{c|c} A_o & D_1 \\ \hline D_1^T & H_1 \end{array}\right]. \tag{3}$$

Blok A_o ma wymiary $(m_o \times m_o)$, takie jak macierz admitancyjna Y_o sieci S_o. Różnice między elementami macierzy A_o i Y_o wynikają z faktu zmian co najmniej niektórych admitancji własnych i wzajemnych węzłów sieci S_o wskutek przyłączenia do niej warstwy ΔS_1. A zatem blok A_o składa się z admitancji własnych i wzajemnych węzłów sieci S_o po przyłączeniu do niej warstwy ΔS_1.

Bloki D_1 o wymiarach $(m_o \times r_1)$ i D_1^T o wymiarach $(r_1 \times m_o)$ zawierają admitancje wzajemne między np. węzłem i-tym, należącym do sieci S_o a węzłem j-tym, należącym do warstwy ΔS_1. Blok H_1 o wymiarach $(r_1 \times r_1)$ zawiera admitancje warstwy ΔS_1.

W drugim kroku rekurencyjnym powstaje sieć S_2 po dołączeniu do sieci S_1 warstwy ΔS_2 z zachowaniem warunku omówionego w p. 2. Macierz admitancyjną sieci S_2 można przedstawić w następującej postaci blokowej

$$Y_2 = \left[\begin{array}{c|c|c} A_o & D_1 & O_2 \\ \hline D_1^T & A_1 & D_2 \\ \hline O_2^T & D_2^T & H_2 \end{array}\right]. \tag{4}$$

Przyjmuje się, że węzły warstwy ΔS_2 są ponumerowane od (m_1+1) do (m_1+r_2), gdzie r_2 oznacza liczbę węzłów sieci ΔS_2.

Macierz Y_2 ma wymiary $(m_2 \times m_2)$, gdzie

$$m_2 = m_1 + r_2 = m_o + r_1 + r_2 \tag{5}$$

jest liczbą węzłów niezależnych sieci S_2.

Bloki A_o, D_1, D_1^T macierzy Y_2 są identyczne z odpowiednimi blokami macierzy Y_1.

Blok A_1 o wymiarach $(r_1 \times r_1)$ zawiera admitancje własne i wzajemne węzłów warstwy ΔS_1, zmodyfikowane po dołączeniu warstwy ΔS_2. Jest to zatem modyfikacja bloku H_1 macierzy Y_1.

Bloki D_2 o wymiarach $(r_1 \times r_2)$ i D_2^T o wymiarach $(r_2 \times r_1)$ to macierze admitancji wzajemnych między węzłami sieci ΔS_1 a węzłami sieci ΔS_2. Opisują one sposób połączenia obu kolejnych warstw.

W bloku H_2 o wymiarach $(r_2 \times r_2)$ znajdują się admitancje własne i wzajemne węzłów warstwy ΔS_2 po dołączeniu do sieci S_1.

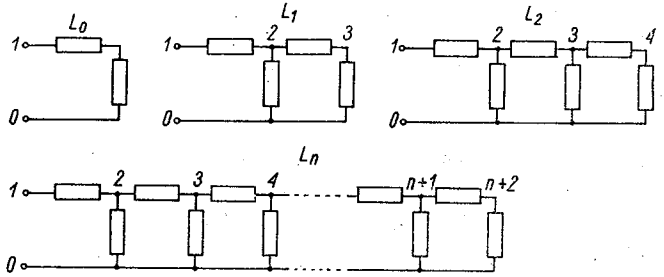
Istotnym faktem w strukturze macierzy Y_2 jest występowanie bloków zerowych O_2 o wymiarach $(m_o \times r_2)$ i odpowiednio O_2^T o wymiarach $(r_2 \times m_o)$, które są konsekwencją przyjętej w p. 2 zasady tworzenia ciągu sieci $\{S_n\}$ w sposób rekurencyjny, tj. założenia, że dla każdego $n \geqslant 2$ połączenia występują wyłącznie między następującymi po sobie warstwami.

W przypadku sieci S_2 bloki zerowe O_2, O_2^T macierzy admitancyjnej Y_2 oznaczają brak połączeń między warstwą ΔS_2 a tą częścią sieci S_1, która tworzyła sieć początkową S_o.

Strukturę macierzy admitancyjnej Y_{n+1} sieci S_{n+2} dla dowolnego $n \geqslant 2$ można otrzymać następująco. Przyjmuje się, że n-ty element ciągu sieci $\{S_n\}$ powstaje po dołączeniu do pewnej sieci S_{n-1} warstwy ΔS_n. Sieć S_n ma macierz admitancyjną Y_n o wymiarach $(m_n \times m_n)$, gdzie m_n — liczba węzłów niezależnych sieci S_n, o postaci blokowej

$$Y_n = \left[\begin{array}{c|c} U_{n-1} & D_n \\ \hline D_n^T & H_n \end{array}\right]. \tag{6}$$

Blok U_{n-1} o wymiarach $(m_{n-1} \times m_{n-1})$ jest macierzą admitancyjną sieci S_{n-1} po dołączeniu warstwy ΔS_n w sposób zilustrowany na rys. 2, przy czym m_{n-1} jest liczbą węzłów niezależnych sieci S_{n-1}.



Rys. 2. Ciąg sieci warstwowych $\{L_n\}$

Blok D_n o wymiarach $(m_{n-1} \times r_n)$ i blok D_n^T o wymiarach $(r_n \times m_{n-1})$ to bloki admitancji wzajemnych między węzłami sieci S_{n-1} a węzłami warstwy ΔS_n, która zawiera r_n węzłów ponumerowanych od $(m_{n-1}+1)$ do $(m_{n-1}+r_n)$.

Blok H_n o wymiarach $(r_n \times r_n)$ to macierz admitancyjna warstwy ΔS_n po dołączeniu do sieci S_{n-1}.

Zgodnie z zasadą tworzenia sieci warstwowych sieć S_{n+1} powstaje przez dołączenie do sieci S_n warstwy ΔS_{n+1} w taki sposób, aby połączenia wystąpiły tylko między węzłami warstw ΔS_n i ΔS_{n+1}, tzn. aby nie istniały połączenia między węzłami sieci S_{n-1} i ΔS_{n+1}.

Macierz admitancyjną Y_{n+1} można otrzymać przez rozbudowę oraz modyfikację macierzy Y_n. Otrzymuje się następującą postać blokową macierzy Y_{n+1}

$$Y_{n+1} = \left[\begin{array}{c|c|c} U_{n-1} & D_n & O_{n+1} \\ \hline D_n^T & A_n & D_{n+1} \\ \hline O_{n+1}^T & D_{n+1}^T & H_{n+1} \end{array}\right]. \tag{7}$$

Bloki U_{n-1}, D_n, D_n^T są określone jak dla sieci S_n. A_n o wymiarach $(r_n \times r_n)$ jest zmodyfikowanym blokiem H_n macierzy Y_n i składa się z admitancji węzłów warstwy ΔS_n po dołączeniu warstwy ΔS_{n+1}.

Bloki D_{n+1} o wymiarach $(r_n \times r_{n+1})$ i D_{n+1}^T o wymiarach $(r_{n+1} \times r_n)$ to bloki admitancji wzajemnych między węzłami warstw ΔS_n i ΔS_{n+1}, przy czym węzły sieci ΔS_{n+1} są ponumerowane od (m_n+1) do (m_n+r_{n+1}).

Macierz H_{n+1} o wymiarach $(r_{n+1} \times r_{n+1})$ to macierz admitancyjna warstwy ΔS_{n+1} po dołączeniu do sieci S_n.

Zgodnie ze wspomnianą zasadą rekurencyjnego określenia struktury sieci warstwowych, przedstawioną na rys. 2, wszystkie admitancje wzajemne między węzłami sieci S_{n-1} a węzłami sieci ΔS_{n+1} są równe zeru. W macierzy admitancyjnej Y_{n+1} występują zatem bloki zerowe O_{n+1} o wymiarach $(m_{n-1} \times r_{n+1})$ i O_{n+1}^T o wymiarach $(r_{n+1} \times m_{n-1})$.

W sposób indukcyjny można otrzymać inną postać macierzy Y_{n+1}, wyraźniej wskazującą warstwę struktury ciągu sieci $\{S_n\}$

$$Y_{n+1} = \left[\begin{array}{ccc|c|c} A_0 & D_1 & O_2 & & \\ D_1^T & A_1 & D_2 & & \\ O_2^T & D_2^T & A_2 & O_n & \\ & & & & O_{n+1} \\ & & A_{n-1} & D_n & \\ \hline & O_n^T & D_n^T & A_n & D_{n+1} \\ \hline & O_{n+1}^T & & D_{n+1}^T & H_{n+1} \end{array}\right] \tag{8}$$

Na głównej przekątnej znajdują się bloki $A_0, A_1, \ldots, A_n$ oraz H_{n+1}, które zawierają admitancje własne i wzajemne węzłów kolejnych warstw sieci S_{n+1}. Po obu stronach głównej przekątnej występują bloki: $D_1, D_2, \ldots, D_{n+1}$ oraz dopowiednio $D_1^T, D_2^T, \ldots, D_{n+1}^T$, opisujące połączenia między kolejnymi warstwami.

Szczególnie istotne jest istnienie w strukturze macierzy Y_{n+1} bloków zerowych $O_2, O_3, \ldots, O_n, O_{n+1}$ i odpowiednio $O_2^T, O_3^T, \ldots, O_n^T, O_{n+1}^T$, co jest konsekwencją ogólnego warunku tworzenia rekurencyjnego ciągu sieci warstwowych $\{S_n\}$. W związku z tym macierz Y_{n+1} staje się macierzą rzadką [6] dla większych n.

Powyższe rozważania można podsumować w postaci następującego twierdzenia: ciągu macierzy admitancyjnych $\{Y_n\}$, których strukturę blokową opisują wzory (3), (4) oraz (6), (7), istnieje wtedy i tylko wtedy, gdy odpowiadający ciągowi $\{Y_n\}$ ciąg elektrycznych liniowych sieci pasywnych jest ciągiem sieci warstwowych $\{S_n\}$, zdefiniowanym w punkcie 2.

4. ALGORYTM REKURENCYJNY OBLICZANIA ADMITANCJI MIĘDZYWĘZŁOWYCH

Jeżeli znana jest macierz admitancyjna Y pewnej elektrycznej sieci liniowej S, to admitancja między węzłami odniesienia R a drugim węzłem wyróżnionym P równa się [3]

$$Y_{zp} = \frac{\det Y}{D_{pp}Y}, \tag{9}$$

gdzie $\det Y$ oznacza wyznacznik macierzy Y a $D_{pp}Y$ jest minorem głównym macierzy Y_n po wykreśleniu p-tego wiersza i p-tej kolumny.

A zatem w celu obliczenia admitancji międzywęzłowej należy obliczyć wartość wyznacznika i minorów głównych macierzy admitancyjnej.

Jedną z najbardziej rozpowszechnionych metod numerycznych obliczania wyznaczników jest metoda Shipleya [4, 5], polegająca na sprowadzeniu wyznacznika stopnia n do wyznacznika stopnia $(n-1)$.

Wyznacznik pewnej nieosobliwej macierzy kwadratowej A o wymiarach $(n \times n)$

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \tag{10}$$

można przedstawić w postaci iloczynu elementów na głównej przekątnej

$$\det A = a_{11} \cdot a_{22}^{(1)} \cdot a_{33}^{(2)} \dots a_{n-1,n-1}^{(n-2)} \cdot a_{nn}^{(n-1)}. \tag{11}$$

Indeks górny we wzorze (11) oznacza numer kolejnego przekształcenia wszystkich elementów macierzy A według wzoru Gaussa [4, 5]

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - \frac{a_{ik}^{(k-1)} \cdot a_{kj}^{(k-1)}}{a_{kk}^{(k-1)}}, \quad a_{kk}^{(k-1)} \neq 0, \quad k = 2, 3, \dots, n, \quad i, j = k, k+1, \dots, n. \tag{12}$$

Warunek, aby mianownik wzoru (12) był różny od zera jest w sposób oczywisty spełniony w przypadku macierzy admitancyjnej.

Algorytm Shipleya ma właściwości istotne z punktu widzenia metody rekurencyjnej:

a) jeżeli macierz A jest macierzą symetryczną, to wystarcza przekształcić i zapamiętywać tylko połowę macierzy,

b) jeżeli macierz A ma następującą strukturę blokową

$$A = \begin{bmatrix} \overbrace{A_{11}}^{m} & A_{12} & \overbrace{O_{13}}^{k} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{matrix} \}m \\ \\ \}k \end{matrix}, \tag{13}$$

gdzie O_{13} jest blokiem zawierającym wyłącznie elementy równe zeru, to rozwinięcie wyznacznika wzdłuż przekątnej bloku A_{11} nie zmienia elementów bloków A_{23} i A_{33}, ponieważ, zgodnie z wzorem Gaussa (12), licznik poprawki równa się zeru; a zatem stosując (12) do macierzy (13) otrzymuje się

$$\det A = a_{11} a_{22}^{(1)} \dots a_{mm}^{(m-1)} \cdot \det \begin{bmatrix} A_{22}^{(m-1)} & A_{23} \\ A_{32}^{(m-1)} & A_{33} \end{bmatrix}. \tag{14}$$

Opisany powyżej algorytm obliczania wyznacznika jest w szczególności odpowiedni dla rekurencyjnego obliczania wyznacznika macierzy admitancyjnych $\{Y_n\}$ o postaci blokowej, wynikającej z warstwowej struktury sieci $\{S_n\}$.

Po zastosowaniu algorytmu Shipleya do obliczania wyznacznika macierzy admitancyjnej Y_1 o postaci (3), odpowiadającej sieci S_1, otrzymuje się następujące wyrażenia po wykonaniu $(m_o - 1)$ modyfikacji

$$\det Y_1 = \det \begin{bmatrix} A_o & D_1 \\ D_1^T & H_1 \end{bmatrix} = a_{11} \cdot a_{22}^{(1)} \dots a_{m_o, m_o}^{(m_o-1)} \cdot \det[H_1^{(m_o-1)}], \tag{15}$$

gdzie:

$a_{ii}^{(j)}$ — element na głównej przekątnej bloku A_o macierzy Y_1, po wykonaniu j-tej modyfikacji według wzoru (12);

$H_1^{(m_o-1)}$ — blok H_1 macierzy Y_1 o wymiarach $(r_1 \times r_1)$ po $(m_o - 1)$ modyfikacjach.

Macierz Y_2 sieci S_2 ma postać (4). Istnienie bloków zerowych O_2 i O_2^T umożliwia wykorzystanie do obliczenia wyznacznika tej macierzy części obliczeń wykonanych dla sieci S_1 na podstawie właściwości (14).

Jak wiadomo, macierz Y_2 otrzymuje się przez dodanie do macierzy Y_1 bloków D_2, D_2^T i O_2, O_2^T oraz H_2, ponadto dokonuje się przekształcenia bloku H_1 na blok A_1. Przekształcenie to jest związane ze zmianami co najmniej niektórych admitancji własnych i wzajemnych węzłów warstwy ΔS_1 po dołączeniu warstwy ΔS_2. Przekształcenie to jest addytywne i można je opisać jako

$$A_1 = H_1 + \Delta A_1, \tag{16}$$

gdzie ΔA_1 — macierz zmian admitancji węzłów warstwy ΔS_1 po dołączeniu warstwy ΔS_2.

Na podstawie zależności (14) można napisać

$$A_1^{(m_o-1)} = H_1^{(m_o-1)} + \Delta A_1, \tag{17}$$

a więc dołączenie warstwy ΔS_2 nie powoduje konieczności powtórzenia obliczeń wykonanych w pierwszych (m_o-1) modyfikacjach. Ą zatem oznaczając iloczyn

$$a_{11} \cdot a_{22}^{(1)} \dots a_{m_o,m_o}^{(m_o-1)} = \det A_o, \tag{18}$$

otrzymuje się wyrażenie na wyznacznik macierzy Y_2

$$\det Y_2 = \det A_o \cdot \det \left[\begin{array}{c|c} A_1^{(m_o-1)} & D_2 \\ \hline D_2^T & H_2 \end{array} \right]. \tag{19}$$

W kroku tym pozostaje zatem do obliczenia wyznacznik stopnia (r_1+r_2).

Rozwijając drugi czynnik iloczynu według wzoru Gaussa (12) otrzymuje się

$$\det Y_2 = \det A_o \cdot \det[A_1^{(m_o-1)}] \cdot \det[H_2^{(m_1-1)}]. \tag{20}$$

W trakcie obliczeń dla n-tej sieci ciągu $\{S_n\}$ wyrażenie na wyznacznik macierzy Y_n o postaci blokowej (6) można przedstawić następująco

$$\det Y_n = \det \left[\begin{array}{c|c} U_{n-1} & D_n \\ \hline D_n^T & H_n \end{array} \right] = \det U_{n-1} \cdot \det[H_n^{(m_{n-1}-1)}]. \tag{21}$$

W celu obliczenia wyznacznika następnej macierzy Y_{n+1}, odpowiadającej sieci warstwowej S_{n+1}, należy najpierw utworzyć tę macierz na podstawie macierzy Y_n. Blok U_{n-1} pozostaje bez zmian. Blok zmodyfikowany $H_n^{(m_{n-1}-1)}$ zostaje przekształcony odpowiednio do zmian admitancji węzłów warstwy ΔS_n po dołączeniu warstwy ΔS_{n+1}, co można ogólnie zapisać jako tworzenie zmodyfikowanego bloku $A_n^{(m_{n-1}-1)}$

$$A_n^{(m_{n-1}-1)} = H^{(m_{n-1}-1)} + \Delta A_n, \tag{22}$$

gdzie ΔA_n — macierz zmian admitancji węzłów warstwy ΔS_n po dołączeniu warstwy ΔS_{n+1}.

Po utworzeniu bloków H_{n+1}, D_{n+1}, D_{n+1}^T, związanych z warstwą ΔS_{n+1}, wyrażenie na wyznacznik macierzy Y_{n+1} jest następujące

$$\det Y_{n+1} = \det U_{n-1} \cdot \det \left[\begin{array}{c|c} A_n^{(m_{n-1}-1)} & D_{n+1} \\ \hline D_{n+1}^T & H_{n+1} \end{array} \right]. \tag{23}$$

Rozwijając iloczyn wzdłuż przekątnej bloku $A_n^{(m_{n-1}-1)}$ otrzymuje się

$$\det Y_{n+1} = \det U_n \cdot \det[H_{n+1}^{(m_n-1)}], \tag{24}$$

a więc wyrażenie o postaci (21) po zwiększeniu indeksów oznaczających numer sieci o 1. Oznacza to, że obliczenie wyznaczników macierzy admitancyjnych sieci warstwowych oraz odpowiednich minorów głównych może być dokonane w sposób rekurencyjny, a także na podstawie (9) możliwe jest rekurencyjne obliczenie admitancji międzywęzłowych dla sieci tej klasy.

Ze wzoru (24) wynika, że wyznacznik macierzy Y_{n+1} może być obliczany wyłącznie na podstawie znajomości bloku H_{n+1}, określającego admitancje warstwy ΔS_{n+1} oraz przekształconego bloku A_n, związanego z warstwą ΔS_n. Zapamiętywanie bloku U_{n-1}, a więc również sieci S_{n-1}, jest zbędne. Wynika stąd główna zaleta metody rekurencyjnej. W $(n+1)$-tym kroku zamiast całej sieci S_{n+1} bierze się pod uwagę podczas obliczeń tylko węzły i gałęzie należące do warstw ΔS_n i ΔS_{n+1} i w konsekwencji można zredukować zapamiętywanie macierzy Y_{n+1} o wymiarach równych liczbie węzłów niezależnych sieci S_{n+1} do wymiarów podmacierzy złożonej z bloków A_n, D_n, H_{n+1} o wymiarach równych liczbie węzłów 2 ostatnich warstw. Umożliwia to zastosowanie metody w przypadku sieci złożonych z wielu tysięcy węzłów i gałęzi, które powstają wskutek agregacji kolejnych warstw.

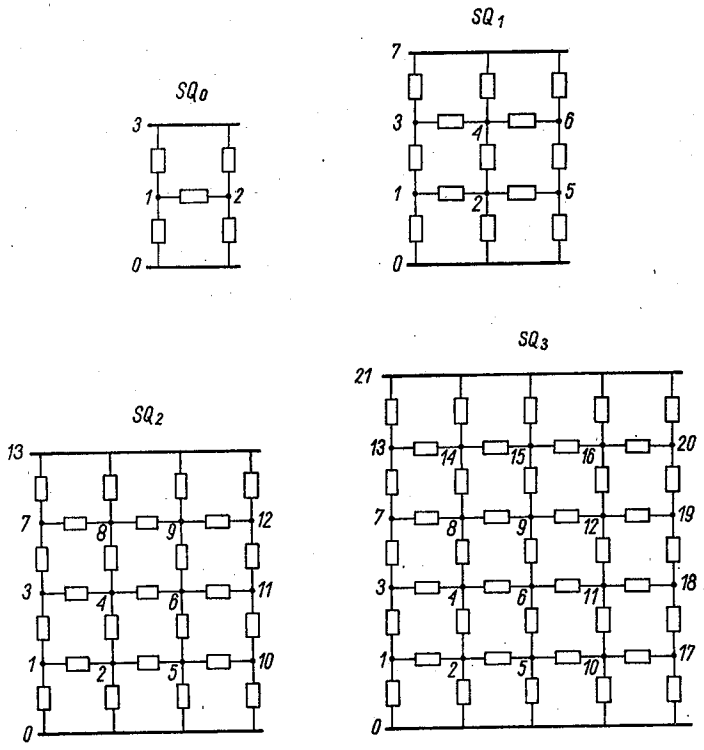
Metoda obliczeń rekurencyjnych admitancji międzywęzłowych, wykorzystująca wzór Gaussa, należy do tzw. dokładnych metod numerycznych, tzn. błędy powstają tylko na skutek zaokrągleń. Z tego punktu widzenia korzystny jest przypadek, gdy w sieciach warstwowych ciągu $\{S_n\}$ występują elementy reaktancyjne jednego typu. Wówczas elementy na głównej przekątnej macierzy admitancyjnych $\{Y_n\}$, równe sumie admitancji incydentnych z danym węzłem, są największe co do modułu.

5. PRZYKŁADY CIĄGÓW SIECI WARSTWOWYCH

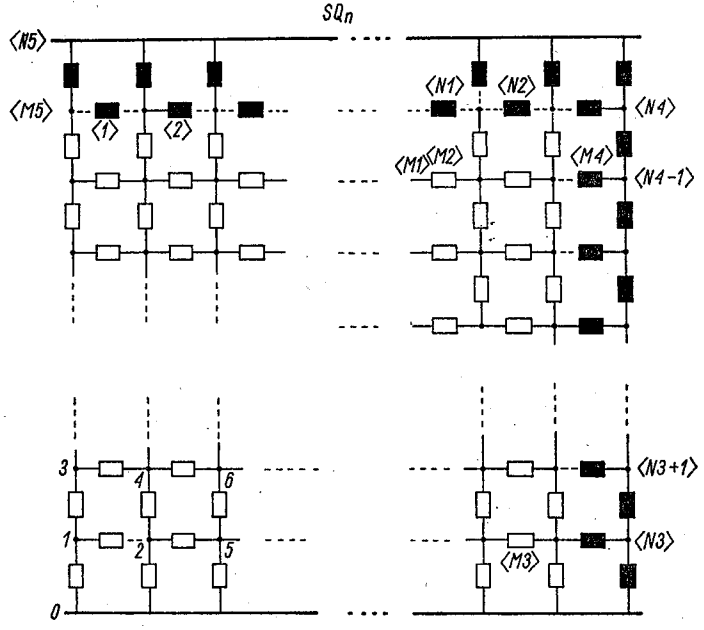
W celu zilustrowania wprowadzonych pojęć oraz możliwości zastosowań metody przytoczono proste przykłady ciągów sieci warstwowych.

A) Na rys. 2 przedstawiono pierwsze elementy ciągu sieci $\{L_n\}$. Ciąg sieci $\{L_n\}$ prowadzi do określenia linii łańcuchowej dla dużych n. Wspólnym węzłem odniesienia jest węzeł 0. Warstwy mają postać ogólną typu Γ i składają się z 1 węzła i 2 gałęzi. Z tego powodu bloki macierzy admitancyjnej redukują się do elementów i macierz admitancyjna linii łańcuchowej jest macierzą trójdiagonalną. Możliwe jest zatem rekurencyjne obliczanie admitancji wejściowej i wyjściowej dla dowolnie dużej liczby ogniw. Szczególne postaci rozwiązań [2] wynikają z przyjęcia określonych, funkcyjnych rozkładów admitancji gałęziowych w celu modelowania realizacji technicznych linii transmisyjnych.

B) Interesującym przykładem ciągu sieci warstwowych jest ciąg sieci o regularnej topologii kwadratowej, nazywany $\{SQ_n\}$. Pierwsze elementy ciągu $\{SQ_n\}$ przedstawiono na rys. 3. Zadanie polega na wyznaczeniu admitancji zastępczej każdej sieci SQ_n między dolnym brzegiem, będącym węzłem odniesienia, a brzegiem górnym. Sieci ciągu $\{SQ_n\}$ mają następującą właściwość: jeśli wszystkie admitancje gałęziowe sieci są identyczne i równe Y_o, to prąd nie płynie przez gałęzie poziome i admitancja zastępcza każdej sieci ciągu jest równa admitancji gałęziowej Y_o. Właściwość ta może być wykorzystana w celu sprawdzenia poprawności algorytmu. W sposób ogólny n-ty element ciągu — sieć SQ_n przedstawiono na rys. 4. Zaczernione są admitancje warstwy ΔSQ_n, dołączonej w n-tym



Rys. 3. Pierwsze elementy ciągu sieci kwadratowych $\{SQ_n\}$

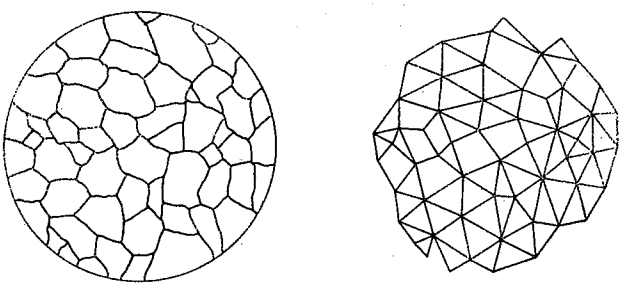


Rys. 4. Ogólna postać sieci kwadratowej SQ_n

kroku rekurencyjnym. Regularność topologiczna sieci $\{SQ_n\}$ umożliwia automatyczne tworzenie macierzy admitancyjnej y_{n+1} sieci SQ_{n+1} na podstawie znajomości sieci SQ_n oraz reguły dołączania warstwy ΔSQ_{n+1}. Warstwa ta składa się z grupy węzłów i gałęzi dołączonych od strony prawej sieci oraz w wyniku rozwarcia górnego brzegu i dołączenia oznaczonych na czarno węzłów i gałęzi. Ciąg określonych w powyższy sposób sieci kwadratowych $\{SQ_n\}$ wykorzystano w celu wykazania zarówno efektywnego działania metody rekurencyjnej, jak też wskazania, że może ona być użytecznym narzędziem poznawczym w materiałoznawstwie. Umożliwia ona badanie jakościowo nowych elektrycznych modeli zastępczych, których idea zostanie skrótowo przytoczona poniżej.

Z punktu widzenia zastosowań przedstawionej metody, bardzo ważnym jest fakt, że strukturę warstwową mają grafy sieci zastępczych, odpowiadających materiałom ziarnistym, takim jak ciała polikrystaliczne, granulaty, proszki itp. Również sieci, których grafy odpowiadają strukturom krystalograficznym, należą do klasy sieci warstwowych. W materiałach ziarnistych połączenia elektryczne występują tylko między stykającymi się ziarnami — w sposób naturalny można zatem zdekomponować je na kolejne „warstwy".

Na rys. 5 pokazano strukturę ziarnistą ferrytu polikrystalicznego w przekroju płaskim oraz odpowiadający jej graf sieci zastępczej, utworzony następująco: wewnątrz każdego ziarna jest umieszczony węzeł, a gałęzie między węzłami występują tylko wtedy, gdy od-

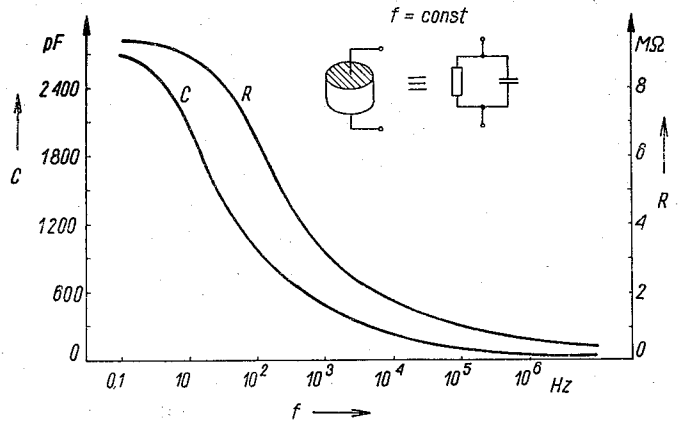


Rys. 5. Struktura polikrystaliczna ferrytu i graf jej sieci zastępczej

powiednie ziarna stykają się ze sobą. Sieć, której graf ukazano na rysunku 5b, jest siecią warstwową i może ona być analizowana metodą rekurencyjną, jednak ze względu na losowy charakter grafu, dane dotyczące łączenia kolejnych warstw trzeba by wprowadzać do maszyny cyfrowej w każdym kroku. Określony ciąg sieci $\{SQ_n\}$ można uważać za idealizację, polegającą na założeniu regularnej topologii sieci zastępczej. Obliczenie admitancji zastępczej między przeciwległymi brzegami odpowiada sytuacji przy pomiarach właściwości materiału umieszczonego między elektrodami. Zadanie modelowania polega na określeniu związków między właściwościami materiału w mikroobszarach oraz jego właściwościami efektywnymi.

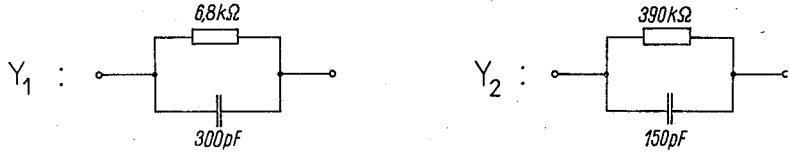
Losowy charakter przewodzenia we wnętrzach oraz na stykach ziaren przedstawiono w rozważaniach modelowych za pomocą losowego rozkładu wartości admitancji gałęziowych, zbudowanych z elementów *RC*. Rdzenie ferrytowe traktowane są podczas pomiarów jako równoległe dwójniki *RC*. Przebiegi rezystancji R_z i pojemności zastępczej C_z wykazują znaczne zmiany w szerokim paśmie częstotliwości, jak na rys. 6. Zjawisko

to jest nazywane dyspersją dielektryczną. Osobliwą cechą ferrytów jest bardzo duży wzrost pojemności zastępczej dla małych częstotliwości w stosunku do pojemności dla dużych częstotliwości [7].



Rys. 6. Właściwości elektryczne ferrytów polikrystalicznych

Sieci ciągu $\{SQ_n\}$, w których admitancje gałęziowe typu RC przyjmują wartości losowe, stanowią pewien wyidealizowany zbiór modeli przewodzenia prądu w niejednorodnym materiale ziarnistym. Dla przeprowadzenia obliczeń na maszynie CYBER 70 przyjęto, że admitancje gałęziowe sieci $\{SQ_n\}$ są równoległymi dwójnikami RC. Rodzaje dwójników elementarnych i wartości ich elementów przedstawiono na rys. 7. Stosunek rezy-



Rys. 7. Dwójniki elementarne stosowane w sieciach losowych $\{SQ_n\}$

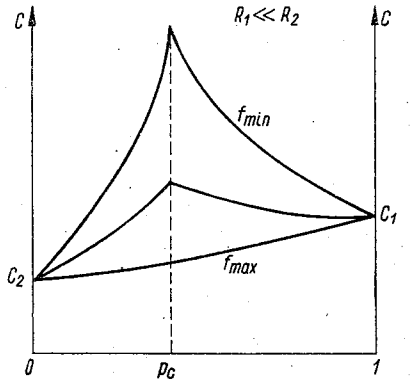
stancji R_z/R_1 przekracza 50, a stosunek pojemności C_z/C_1 równa się 0,5. Przyjęto, że dwójnik Y_1 występuje z prawdopodobieństwem p, a dwójnik Y_2 z prawdopodobieństwem $(1-p)$. Ponadto przyjęto, że poszczególne admitancje gałęziowe są statystycznie niezależne od wzajemnego usytuowania w sieci.

Po przeprowadzeniu serii obliczeń można sformułować następujące wnioski:

1. częstotliwościowe pasmo dyspresji jest określone przez wartości stałych czasu dwójników elementarnych i nie zależy od wartości parametru p rozkładu binarnego,
2. zmiany rezystancji R_z i pojemności C_z sieci zależą od wartości p, od admitancji gałęziowych Y_1 i Y_2 oraz od liczby węzłów sieci.

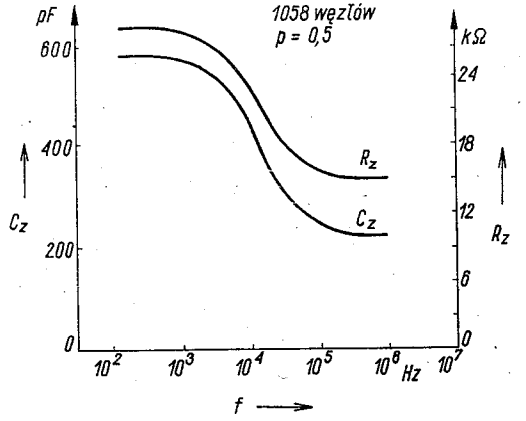
Najważniejszym faktem zaobserwowanym w wyniku symulacji numerycznej było występowanie maksymalnego wzrostu pojemności zastępczej C_z dla małych częstotliwości dla wartości parametru rozkładu $p = 0{,}5$. Symulacje numeryczne przeprowadzone dla sieci o innych topologiach (np. trójkątnej i sześciennej prostej [8]) wykazały, że zjawisko

takie zachodzi zawsze dla pewnej wartości charakterystycznej p_c rozkładu binarnego, zależnej tylko od topologii sieci. Maksimum C_z rośnie ze wzrostem stosunku R_2/R_1 przy stałym stosunku C_2/C_1, jak pokazano na rys. 8. Na tej podstawie można postawić hipotezę, że duża wartość pojemności zastępczej ferrytów dla małych częstotliwości wynika ze znacznego rozrzutu przewodności w mikroobszarach materiału.



Rys. 8. Wzrost pojemności zastępczej w binarnych sieciach *RC*

Badania numeryczne sieci losowych $\{SQ_n\}$ wskazują również, że „zbieżność" obliczeń ze wzrostem liczby węzłów sieci jest dość powolna i za reprezentatywne można uważać obliczenia dla sieci zawierających ponad 1000 węzłów. Przykładowe wyniki obliczeń dla sieci SQ_{31} o liczbie węzłów 1058 pokazano na rys. 9.



Rys. 9. Rezystancja i pojemność zastępcza sieci SQ_{31}

Przykładowa realizacja maszynowa algorytmu rekurencyjnego w systemie Cyber wykazała, że możliwe jest obliczanie admitancji zastępczych sieci ciągu $\{SQ_n\}$, zawierających ponad 3300 węzłów. Jest to wielkość znacznie przekraczająca możliwości znanych metod obliczeniowych analizy sieci.

Czas obliczeń jest proporcjonalny do $m_n^3/6$, gdzie m_n — liczba węzłów sieci i wynosił około 900 s dla 1 częstotliwości dla $m_n \approx 3300$. Rozpatrywanie tak dużych sieci losowych *RC* jest uzasadnione faktem bardzo wolnej „zbieżności" wyników obliczeń ze wzrostem liczby elementów sieci. Jest to właściwość sieci o losowych rozkładach admitancji. Należy zaznaczyć, że właściwości takich sieci są dotychczas mało poznane. Omówienie ich przekracza zamierzenia tego artykułu.

6. PODSUMOWANIE

W artykule opisano podstawowe zasady rekurencyjnej metody obliczania admitancji międzywęzłowych dla pewnej klasy liniowych pasywnych sieci elektrycznych. Obliczanie admitancji międzywęzłowych w takich sieciach przebiega na ogół w sposób wymagający odwracania macierzy rzędu równego liczbie węzłów niezależnych sieci. Liczba działań arytmetycznych, które należy wykonać, rośnie bardzo szybko, co prowadzi do powiększenia czasu obliczeń. Druga bariera wiąże się ze skończoną i ograniczoną pamięcią operacyjną maszyn cyfrowych. Z tych powodów efektywność obliczeń przy użyciu maszyn cyfrowych maleje ze wzrostem liczby węzłów sieci tak, że nawet przy użyciu dużych maszyn nie można obliczyć admitancji międzywęzłowych, gdy liczba węzłów niezależnych przekracza kilkaset, a w szczególnych przypadkach około tysiąc węzłów.

Idea przedstawionej metody obliczeń wywodzi się z teorii niejednorodnej linii łańcuchowej, której uogólnieniem jest teoria łańcucha $2n$-wrotników niejednorodnych.

Jak wykazano, postępowanie rekurencyjne w celu obliczenia admitancji międzywęzłowych może być stosowane w przypadku szerokiej klasy sieci elektrycznych, nazwanych sieciami warstwowymi.

Analogia między ciągiem sieci warstwowych i łańcuchem $2n$-wrotników polega na tym, że dana sieć ΔS_n ma połączenia z warstwą ΔS_{n-1}, nie ma natomiast połączeń z warstwami ΔS_{n-2}, ΔS_{n-3}, ..., zaś w łańcuchu $2n$-wrotników dane ogniwo ma połączenia z poprzedzającym, ale nie ma ich z dalszymi ogniwami.

Struktura topologiczna ciągu sieci warstwowych $\{S_n\}$ implikuje szczególną postać blokową ciągu macierzy admitancyjnych $\{Y_n\}$, które zostały opisane w p. 3. Istotnym faktem z punktu widzenia metody rekurencyjnej jest istnienie w strukturze macierzy Y_n bloków zerowych $O_2, \ldots, O_n$ oraz $O_2^T, \ldots, O_n^T$.

Aby wyznaczyć admitancje międzywęzłowe według (9) należy obliczyć wyznacznik i odpowiednie minory główne macierzy admitancyjnej. W pracy zaproponowano zastosowanie w tym celu tzw. algorytmu Shipleya [3], opartego na zależności Gaussa (12). Algorytm ten należy do tzw. dokładnych metod numerycznych, tzn. błędy są skutkiem zaokrąglenia wyników działań cząstkowych. W efekcie obliczania admitancji międzywęzłowych dla ciągu sieci warstwowych można prowadzić rekurencyjnie, począwszy od sieci o niewielkiej liczbie węzłów i gałęzi aż do sieci złożonych z wielu tysięcy węzłów i gałęzi. Jest to możliwe dzięki temu, że w każdym kroku rekurencyjnym w pamięci maszyny cyfrowej przechowuje się wyłącznie dane dotyczące 2 ostatnich warstw, nie trzeba natomiast zapamiętywać całej macierzy admitancyjnej.

Klasa sieci warstwowych obejmuje zarówno sieci o topologii regularnej, jak i nieregularnej, sieci planarne oraz nieplanarne. Z punktu widzenia zastosowań metody szcze-

gólnie ważny jest fakt, że strukturę warstwową mają grafy sieci zastępczych, które odpowiadają materiałom ziarnistym oraz strukturom krystalograficznym.

Ogólne założenia zakresu stosowalności metody rekurencyjnej nie obejmują ograniczeń dotyczących wartości admitancji gałęziowych, które mogą przyjmować dowolne wartości zespolone. Z tego powodu metoda rekurencyjna jest stosowana w numerycznych badaniach właściwości wielkich sieci elektrycznych, w których admitancje gałęziowe są określone przez zmienną losową. Tego rodzaju sieci wydają się być właściwymi modelami elektrycznymi niejednorodnych ośrodków ziarnistych. Badanie takich modeli może dopomóc w rozstrzyganiu wielu hipotez dotyczących związków między właściwościami mikrostrukturalnymi ośrodka a jego właściwościami makroskopowymi; np. do takich zagadnień należy wyjaśnienie zjawiska małoczęstotliwościowej dyspersji dielektrycznej materiałów polikrystalicznych.

Temat artykułu jest ściśle związany z badaniami prowadzonymi w Instytucie Podstaw Elektroniki Politechniki Warszawskiej na zlecenie Zakładu Materiałów Magnetycznych „Polfer" w ramach Programu Rządowego PR-3.

LITERATURA

1. I. Koźniewska, *Równania rekurencyjne*, PWN, Warszawa 1972.
2. T. Cholewicki, *Elektryczne linie długie i układy drabinkowe niejednorodne*, PWN, Warszawa 1974.
3. S. Seshu, M.B. Reed, *Linear Graphs and Electric Networks*, Addison-Wesley, Reading. Mass., 1961.
4. I.M. Gelfand, *Wykłady z algebry liniowej*, PWN, Warszawa 1971.
5. H.E. Brown, *Solution of Large Networks by Matrix Methods*, Wiley, New York 1975.
6. J. Dobrowolski, *Algorytmy i struktury zapamiętywania danych techniki macierzy rzadkich do analizy obwodów mikrofalowych*, Rozpr. Elektrotechn., Tom XXIII, z. 4, 1977.
7. S. Krupička, *Fizika ferritów i rodstviennych im magnitnych okisłow*, Moskwa 1976.
8. V.K.S. Shante, S. Kirkpatrick, *An Introduction to Percolation Theory*, Adv. in Phys., 20, 325—357, 1971.

S. JANKOWSKI

PRINCIPLES OF RECURRENCE METHOD OF INTERNODAL ADMITTANCE COMPUTATION

Summary

The paper presents the principles of recurrence method of internodal admittance computation for a sequence of linear passive electric networks. It was proved that this method can be applied for so called „stratified networks". These networks are obtained by recurrence aggregation of so called „layer networks" provided that bond connections appear between neighbouring layers only, i.e. between n-th and $(n+1)$th layers. Stratified networks are either topologically regular or irregular and either planar or non-planar as well. Block structure of stratified network admittance matrices enables to reduce the admittance matrix of n-th network to the admittance matrix of last two layers for the purpose of computation. Thus, recurrence algorithms of internodal admittance computation are efficient for networks composed of thousands of nodes and bonds. Examples of stratified networks sequences are also presented in the paper. Recurrence method can be applied to the analysis of large networks with randomly distributed bond admittances for the purpose of modelling electric properties of nonhomogeneous granular media.

S. JANKOWSKI

PRINCIPES DE LA MÉTHODE RÉCURRENTE DE CALCUL DES ADMITTANCES INTERNODALES

Résumé

Dans cet article on a présenté les principes de la méthode récurrente de calcul des admittances internodales pour la séquence des réseaux linéaires passifs électriques. On a prouvé que la méthode peut être appliquée uniquement pour les réseaux appelés dans ce travail „réseaux laminés". Ils sont créés par l'agrégation des réseaux appelés couches, de telle façon que les branches des connections paraissent seulement entre les couches successives, par exemple entre la couche n et $(n+1)$. Les réseaux laminés peuvent être topologiquement réguliers ou irréguliers, plans ou non plans. La structure de bloc des matrices des admittances des réseaux laminés permet de réduire la matrice admittance du réseau n-e à la matrice admittance de deux dernières couches. Pour cette raison les algorithmes récurrents de calcul des admittances internodales sont efficaces pour les réseaux composés de milliers de noeuds et branches.

Dans l'article on a présenté aussi des exemples des séquences des réseaux laminés. La méthode récurrente peut être appliquée également à l'étude des grands réseaux contenant les admittances des branches de valeur aléatoire pour modeler des propriétés électriques des milieux granulaires hétérogènes.

S. JANKOWSKI

GRUNDLAGEN DER REKURRENZMETHODE FÜR DIE BERECHNUNG DER ZWISCHENKNOTENADMITTANZ

Zusammenfassung

Im obigen Artikel wurden die Grundlagen der Rekurrenzmethode für die Berechnung der Zwischenknotenadmittanz für eine Folge linearer passiver Elektronetze dargestellt. Es wurde bewiesen, daß diese Methode lediglich im Falle derjenigen Netze angewandt werden kann, die wir hier Schichtnetze nennen. Diese Netze entstehen zufolge einer Aggregation in einer aufeinander folgenden Reihe von Rekurrenzschritten mancher Netze, Schichten benannt, in derartiger Weise, damit Zweigverbindungen nur zwischen den aufeinander folgenden Schichten auftreten, z.B. zwischen der n-Schicht und der $(n+1)$-Schicht. Die Schichtnetze können topologisch regelmäßig oder unregelmäßig sein, planar oder auch unplanar. Die Blockstruktur der Admittanzmatrix der Schichtnetze gestattet es, daß beim Berechnen die Admittanzmatrix des n-Netzes auf die Admittanzmatrix zweier letzten Schichten reduziert wird. Somit sind die Rekurrenzalgorithmen der Berechnung der Zwischenknotenadmittanzen soger für Netze mit vielen Tausenden Knoten und Zweigen effektiv. Im weiteren wurden ebenfalls Beispiele für Schichtnetzfolgen angeführt. Die Rekurrenzmethode kann zum Prüfen großer Netze mit stochastischer Verteilung der Zweigadmittanz angewandt werden zwecks Modellierung elektrischer Eigenschaften heterogener körniger Zentren.

С. ЯНКОВСКИ

ОСНОВЫ РЕКУРРЕНТНОГО МЕТОДА ВЫЧИСЛЕНИЯ МЕЖДУУЗЛОВОГО АДМИТАНСА

Резюме

В статье представлены основы рекуррентного метода вычисления междуузлового адмитанса для последовательности линейных пассивных электрических сетей. Показано, что метод можно применять единственно в случае сетей называемых в статье слоевыми. Эти сети образуются путем сложения в постепенных рекуррентных шагах определенных сетей называемых слоями, таким образом чтобы ветвенные связи выступали только между постепенными слоями, на пример n-тым и $(n+1)$-ым. Слоевые сети могут быть топологически регулярными или нерегулярными, плос-

кими или непланарными. Блочная структура матриц адмитансов слоевых сетей позваляет уменьшать во время вычислений матрицу адмитансов n-той сети к матрице адмитансов двух последних слоев. Таким образом рекуррентные алгоритмы вычисления междуузловых адмитансов явлвются эффективными для сетей содержащих даже много тысяч узлов и ветвей. В статье предсавлены также примеры последовательностей слоевых сетей. Рекуррентный метод может применяться для исследования больших сетей со случайно распределенными ветвенными адмитансами с целью моделирования электрических свойств неоднородных зернистых сред.

ROZPRAWY ELEKTROTECHNICZNE 1982, 28, z. 3-4, ss. 343—353

621.396.677.83

Zastosowanie metody równań całkowych do analizy elipsoidalnego elektromagnetycznego ekranu

PIOTR KRASOŃ (SZCZECIN)

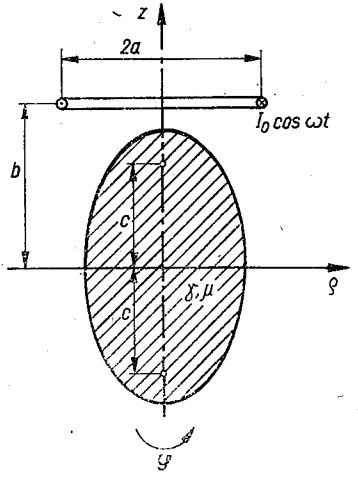
Ośrodek Elektronicznej Techniki Obliczeniowej Politechniki Szczecińskiej

Otrzymano 14.10.1980

W pracy dokonano analizy rozkładu wektora gęstości prądu indukowanego wewnątrz elektromagnetycznego elipsoidalnego ekranu metodą równania całkowego. Podano także sposób wyznaczania współczynnika ekranowania. Analiza prowadzona była poprzez sprowadzenie opisu zjawisk elektromagnetycznych do równania całkowego Fredholma II rodzaju. Wyprowadzono wzory umożliwiające szybkie napisanie algorytmu obliczeniowego na EMC, a tym samym wykazano szczególną przydatność metody równania całkowego w osiowosymetrycznych zagadnieniach ekranowania.

1. WSTĘP

Analiza rozkładu gęstości prądu w przekrojach podłużnych brył obrotowych w układach o osiowej symetrii ma znaczenie przy projektowaniu ekranów elektromagnetycznych. Tradycyjną metodą wykorzystywaną przy obliczeniach jest zwykle metoda rozdzielenia zmiennych [5, 6]. Przykładowo w [2] analizowany jest rozkład gęstości prądu wewnątrz elipsoidy obrotowej znajdującej się w polu pojedynczego zwoju kołowego (rys. 1).



Rys. 1. Elipsoida obrotowa znajdująca się w polu zwoju kołowego

Po wprowadzeniu eliptycznego układu współrzędnych i zastosowaniu metody rozdzielenia zmiennych znalezienie potencjału wektorowego w obszarze przewodzącym sprowadza się do rozwiązania równania o postaci

$$(1-z^2)\frac{d^2w}{dz^2}-2z\frac{dw}{dz}+\left(\lambda-c^2z^2-\frac{\mu}{1-z^2}\right)w=0, \tag{1}$$

w którym c^2 jest stałą urojoną, natomiast w obszarze nieprzewodzącym równania Legendre'a

$$(1-z^2)\frac{d^2w}{dz^2}-2z\frac{dw}{dz}+\left[\nu(\nu+1)-\frac{\mu^2}{1-z^2}\right]w=0. \tag{2}$$

O ile rozwiązanie ogólne równania (2) dla odpowiedniej wartości parametru ν łatwo jest znaleźć, to rozwiązanie ogólne (1) nie jest opracowane w literaturze.

Autor [2] korzysta z rozwinięcia rozwiązania równania (1) względem specjalnego systemu funkcji ortogonalnych i przez skorzystanie z warunków ciągłości otrzymuje ostatecznie nieskończony układ równań pozwalający określić nieznane współczynniki rozwinięcia. Współczynniki układu równań są iloczynami wartości funkcji specjalnych Legendre'a I i II rodzaju oraz funkcji kulistych. Komplikacja obliczeń numerycznych jest więc duża.

W niniejszej pracy wykorzystano do obliczeń metodę równań całkowych. Sformułowanie problemu w postaci równania całkowego Fredholma II rodzaju jako modelu matematycznego posiada zalety w stosunku do innych metod. Podstawowymi zaletami są:
— możliwość automatycznego spełnienia warunku brzegowego na powierzchni przewodnika,
— prostota obliczeń numerycznych,
— możliwość obliczania bardziej skomplikowanych konfiguracji geometrycznych niż w przypadku metody rozdzielenia zmiennych.

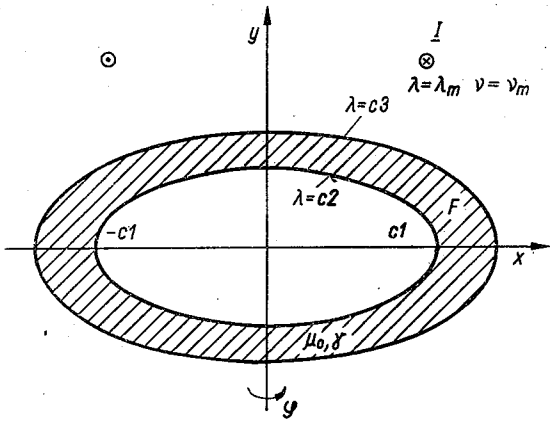
2. OPIS BADANEGO UKŁADU

Na rys. 2 przedstawiono ekran w kształcie elipsoidy obrotowej wydrążonej, znajdujący się w polu zewnętrznym wywołanym obecnością zwoju kołowego wiodącego prąd $i=\sqrt{2}\,I\sin(\omega t)$. Zakłada się, że wydrążenie ma także kształt elipsoidy obrotowej. Badanie układu z rys. 2 dogodnie jest prowadzić w eliptycznym układzie współrzędnych. Jako siatkę współrzędnych (rys. 3) przyjęto rodzinę współogniskowych i współosiowych stożkowych

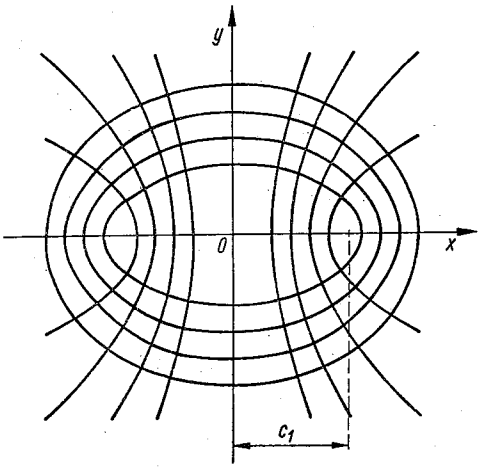
$$\frac{x^2}{l^2}+\frac{y^2}{l^2-c_1^2}=1, \tag{3}$$

hiperbol dla $0<l<c_1$, elips dla $c_1<l$.

Po oznaczeniu symbolem λ wartości parametru l dla elipsy przechodzącej przez punkt (x, y) oraz symbolem ν parametru l dla hiperboli przechodzącej przez punkt (x, y) wza-



Rys. 2. Ekran elektromagnetyczny elipsoidalny



Rys. 3. Eliptyczny układ współrzędnych

jemna zależność współrzędnych wyraża się wzorami:

$$x = \pm \frac{\lambda \nu}{c_1}, \quad y = \pm \frac{\sqrt{(\lambda^2 - c_1^2)(c_1^2 - \nu^2)}}{c_1}. \tag{4}$$

Jakobian przekształcenia współrzędnych dany jest wzorem

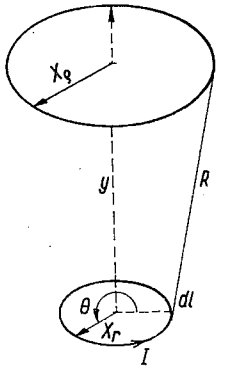
$$\frac{D(x, y)}{D(\lambda, \nu)} = \frac{\nu^2 - \lambda^2}{\sqrt{(\lambda^2 - c_1^2)(c_1^2 - \nu^2)}}. \tag{5}$$

3. MATEMATYCZNE SFORMUŁOWANIA PROBLEMU W POSTACI RÓWNANIA CAŁKOWEGO FREDHOLMA II RODZAJU

W celu wyznaczenia równania całkowego opisującego rozkład gęstości prądu $\vec{J} = J\vec{1}_\varphi$ w przekroju podłużnym ekran można podzielić obszar F (rys. 2) na małe kawałki o polu

o powierzchni dF i traktować tak powstałe „nitki" prądowe jako nieskończenie cienkie zwoje kołowe. Potencjał wektorowy nieskończenie cienkiego zwoju kołowego z prądem $\underline{J}$ wyraża się wzorem [5] (rys. 4)

$$\underline{A} = \frac{\mu_o x_r \underline{J}}{4\pi} \int_0^{2\pi} \frac{\cos\Theta \, d\Theta}{\sqrt{x_r^2 + x_\varrho^2 - 2x_r x_\varrho \cos\Theta + y^2}}. \tag{6}$$



Rys. 4. Przewód kołowy z prądem I

Po przekształceniu (6) można przedstawić w postaci kombinacji całek eliptycznych

$$\underline{A} = \frac{\mu_o \underline{J}}{2\pi} \sqrt{\frac{x_r}{x_\varrho}} \left[\left(\frac{2}{k} - k \right) K(k) - \frac{2}{k} E(k) \right], \tag{7}$$

gdzie

$$k = \frac{2\sqrt{x_r x_\varrho}}{\sqrt{(x_r + x_\varrho)^2 + y^2}},$$

$$K(k) = \int_0^{\frac{\pi}{2}} \frac{d\beta}{\sqrt{1 - k^2 \sin^2\beta}} \quad \text{— całka eliptyczna zupełna I rodzaju,} \tag{8}$$

$$E(k) = \int_0^{\frac{\pi}{2}} \sqrt{1 - k^2 \sin^2\beta} \, d\rho \quad \text{— całka eliptyczna zupełna II rodzaju.}$$

Potencjał wektorowy w dowolnym punkcie P przestrzeni wytworzony przez elementarny zwój kołowy dany jest zgodnie z (7) wzorem

$$d\underline{A} = \frac{\mu_o}{2\pi} \sqrt{\frac{x_r}{x_\varrho}} \underline{J}(x_r, y_r) \left[\left(\frac{2}{k} - k \right) K(k) - \frac{2}{k} E(k) \right] dF,$$

gdzie $\underline{J}(x_r, y_r)$ gęstość prądu w punkcie (x_r, y_r).

Całkując otrzymuje się wzór na potencjał wektorowy w dowolnym punkcie P przestrzeni

$$\underline{A} = \frac{\mu_o}{2\pi} \int_F \underline{J}(x_r, y_r) \sqrt{\frac{x_r}{x_\varrho}} F_1(k) \, dF + \underline{A}_o, \tag{9}$$

gdzie

$$F_1(k) = \left(\frac{2}{k} - k\right)K(k) - \frac{2}{k}E(k), \tag{9a}$$

$\underline{A}_o$ — potencjal wektorowy zewnętrznego pola elektromagnetycznego wytworzonego przez zwój kołowy.

Korzystając ze związku $\underline{J} = -j\omega\gamma\underline{A}$ otrzymuje się ostatecznie równanie całkowe Fredholma II rodzaju, w którym niewiadomą funkcją jest gęstość prądu

$$\underline{J}(x_\varrho, y_\varrho) = j\omega\gamma\underline{A}_o(x_\varrho, y_\varrho) + \frac{j\omega\mu_o\gamma}{2\pi}\int_F \sqrt{\frac{x_r}{x_\varrho}}\,\underline{J}(x_r, y_r)F_1(k)\,dx_r\,dy_r, \tag{10}$$

Przechodząc z kartezjańskiego układu współrzędnych do układu współrzędnych eliptycznych równanie (10) transformuje się do postaci

$$\underline{J}(\lambda_\varrho, \nu_\varrho) = j\omega\gamma\underline{A}_o(\lambda_\varrho, \nu_\varrho) + \frac{j\omega\mu_o\gamma}{2\pi}\int_{C_2}^{C_3}\int_0^{C_1}\sqrt{\frac{x(\lambda,\nu)}{x(\lambda_\varrho,\nu_\varrho)}}\,\underline{J}(\lambda,\nu)F_1(k)\left|\frac{D(x,y)}{D(\lambda,\nu)}\right|d\nu\,d\lambda, \tag{11}$$

przy czym zależność $x(\lambda, \nu)$ i $y(\lambda, \nu)$ dane są wzorami (4).

4. ROZWIĄZANIE RÓWNANIA FREDHOLMA II RODZAJU

Jedną z najbardziej efektywnych metod rozwiązywania równań całkowych o postaci

$$J(x) = f(x) + \alpha\int_F K(x, y)J(y)\,dy \tag{12}$$

jest metoda Fredholma (z wykorzystaniem wzoru na kwadraturę). Ideą tej metody jest sprowadzenie równania całkowego do układu równań algebraicznych liniowych przez zastąpienie całki sumą całkową. W myśl metody Fredholma obszar F dzieli się na elementarne podobszary o objętości dV, w każdym z podobszarów wybiera się punkt pośredni x_k, zastępuje się całkę sumą całkową otrzymując równanie

$$J(x) = f(x) + \alpha dV\sum_{j=1}^{n} K(x, x_j)J(x_j)$$

i żąda się aby spełnione było ono w punktach $x_1, \ldots, x_n$. Po uporządkowaniu otrzymuje się układ równań:

$$[a_{ij}]\begin{bmatrix} J_1 \\ \vdots \\ J_n \end{bmatrix} = \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}, \tag{13}$$

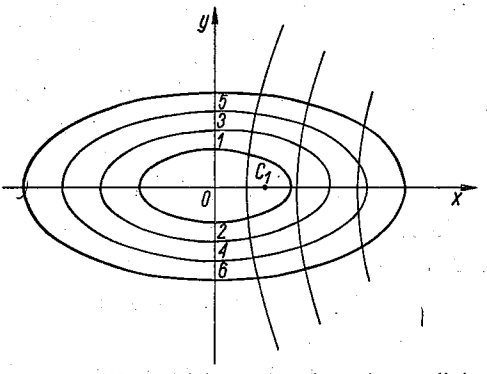
którego współczynniki wyrażają się wzorami:

$$a_{ij} = \begin{cases} 1 - \alpha dVK_{ii}, & i = j, \\ -\alpha dVK_{ij}, & i \neq j, \end{cases} \tag{13a}$$

$$d_j = f(x_j) \quad \text{gdzie} \quad K_{ij} = K(x_i, x_j), \quad J_j = J(x_j).$$

Okazuje się, że rozwiązania $J_1, \ldots, J_n$ dążą do rozwiązań równania (12) gdy $n \to \infty$, $dV \to 0$, przy założeniu całkowalności z kwadratem jądra $K(x, y)$.

W przypadku gdy K_{ii} jest nieokreślone a jądro równania całkowego jest całkowalne z kwadratem, można zastosować metodę Fredholma do jego rozwiązania przyjmując $K_{ii} = 0$ [4]. Wykazanie, że jądro równania (11) jest całkowalne z kwadratem nie przed-



Rys. 5. Sposób podziału na kwadraty krzywoliniowe

stawia trudności. Przedział zmienności współrzędnej λ został podzielony na m_1 części $c_2 < \lambda_1 < \lambda_2 < \ldots < \lambda_{m_1} < c_3$, podczas gdy przedział zmienności ν na m_2 części $0 < \nu_1 < \nu_2 < \ldots < \nu_m < c_1$. Wprowadzono numerację zmiennych zgodną z numeracją kwadratów krzywoliniowych przedstawioną na rys. 5, tzn. J_k oznacza gęstość prądu w k-tym kwadracie. Po obliczeniu współczynniki (13a) wyrażają się wzorami

$$a_{ij} = \begin{cases} 1, & i = j, \\ -\dfrac{j\omega\mu_o\gamma}{2\pi} \sqrt{\dfrac{x_u}{x_t}}\, F_1(k_{u,t}) \left| \dfrac{\nu_u^2 - \lambda_u^2}{\sqrt{(\lambda_u^2 - c_1^2)(c_1^2 - \nu_u^2)}} \right| \Delta\lambda\Delta\nu, \end{cases} \tag{14}$$

gdzie:

$$k_{u,t} = \frac{2\sqrt{x_u x_t}}{\sqrt{(x_u + x_t)^2 + (y_u - y_t)^2}},$$

$$\Delta\lambda = \frac{c_3 - c_2}{m_1}, \quad \Delta\nu = \frac{c_1}{m_2},$$

$$\lambda_u = c_2 + l_1\Delta\lambda,$$

$$\nu_u = l_2\Delta\nu,$$

$$\lambda_t = c_2 + l_3\Delta\lambda,$$

$$\nu_t = l_1\Delta\nu,$$

$$l_2 = \begin{cases} \text{entier } (j/2m_1) + 1 & \text{dla } 2m_1 \nmid j, \\ \dfrac{j}{2m_1} & \text{dla } 2m_1 \mid j, \end{cases} \tag{14a}$$

$$l_1 = \begin{cases} \text{entier } [(j - 2m_1 l_2)/2] + 1 & \text{dla } j - 2m_1 l_2 \text{ nieparzystego}, \\ (j - 2m_1 l_2)/2 & \text{dla } j - 2m_1 l_2 \text{ parzystego}, \end{cases}$$

(l_3, l_4 wyrażają się analogicznie jak l_1, l_2, należy tylko we wzorach (14) j zastąpić przez i

$$x_u = \frac{\lambda_u v_u}{c_1}, \tag{15}$$

$$y_u = \begin{cases} \sqrt{(\lambda_u^2 - c_1^2)(c_1^2 - v_u^2)}/c_1 & \text{dla } j \text{ nieprzystego,} \\ -\sqrt{(\lambda_u^2 - c_1^2)(c_1^2 - v_u^2)}/c_1 & \text{dla } j \text{ parzystego.} \end{cases}$$

Zastępując we wzorach (15) λ_u, v_u, j odpowiednio przez λ_t, v_t, i, otrzymuje się wzory na x_t i y_t

$$d_j = \frac{j\omega\mu_0\gamma J}{2\pi} \sqrt{\frac{x_m}{x_t}} F_1(k_{m,t}),$$

gdzie:

$$x_m = \frac{\lambda_m v_m}{c_1}, \qquad y_m = \frac{\sqrt{(\lambda_m^2 - c_1^2)(c_1^2 - v_m^2)}}{c_1},$$

$$k_{m,t} = \frac{2\sqrt{x_m x_t}}{\sqrt{(x_m + x_t)^2 + (y_m - y_t)^2}}.$$

5. OBLICZENIA NUMERYCZNE

Układ równań (13) został zaprogramowany na podstawie wzorów (14), przy czym zastosowano następujące aproksymacje całek eliptycznych podane w [1]:

$$E(m) = (1 + a_1 m_1 + a_2 m_1^2 + a_3 m_1^3 + a_4 m_1^4) + (b_1 m_1 + b_2 m_1^2 + b_3 m_1^3 + b_4 m_1^4)\ln(1/m_1) + \varepsilon(m),$$

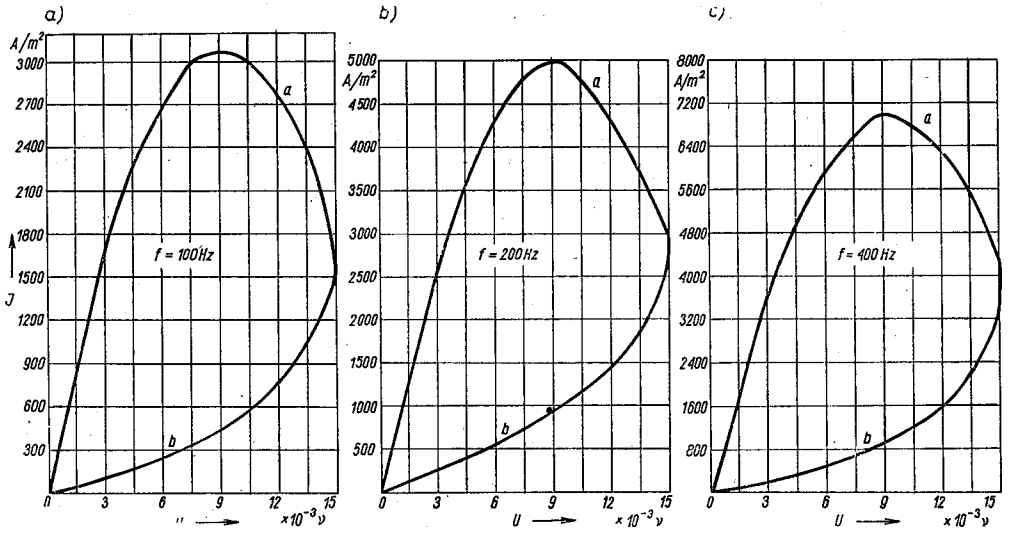
$$m = k^2, \quad m_1 = 1 - m, \quad |\varepsilon(m)| < 2 \cdot 10^{-8},$$

$a_1 = 0{,}443215141463,$ $b_1 = 0{,}2499836810,$
$a_2 = 0{,}06260601220,$ $b_2 = 0{,}09200180037,$
$a_3 = 0{,}4757383546,$ $b_3 = 0{,}04069697526,$
$a_4 = 0{,}01736506451,$ $b_4 = 0{,}00526449639;$

$$K(m) = (a_0 + a_1 m_1 + a_2 m_1^2 + a_3 m_1^3 + a_4 m_1^4) + (b_0 + b_1 m_1 + b_2 m_1^2 + b_3 m_1^3 + b_4 m_1^4)\ln(1/m_1) + \varepsilon(m),$$

$$m = k^2, \quad m_1 = 1 - m, \quad |\varepsilon(m)| < 2 \cdot 10^{-8},$$

$a_0 = 1{,}38629436112,$ $b_0 = 0{,}5,$
$a_1 = 0{,}09666344259,$ $b_1 = 0{,}12498593597,$
$a_2 = 0{,}03590092383,$ $b_2 = 0{,}06880248576,$
$a_3 = 0{,}03742563713,$ $b_3 = 0{,}03328355346,$
$a_4 = 0{,}01451196212,$ $b_4 = 0{,}00441787012.$



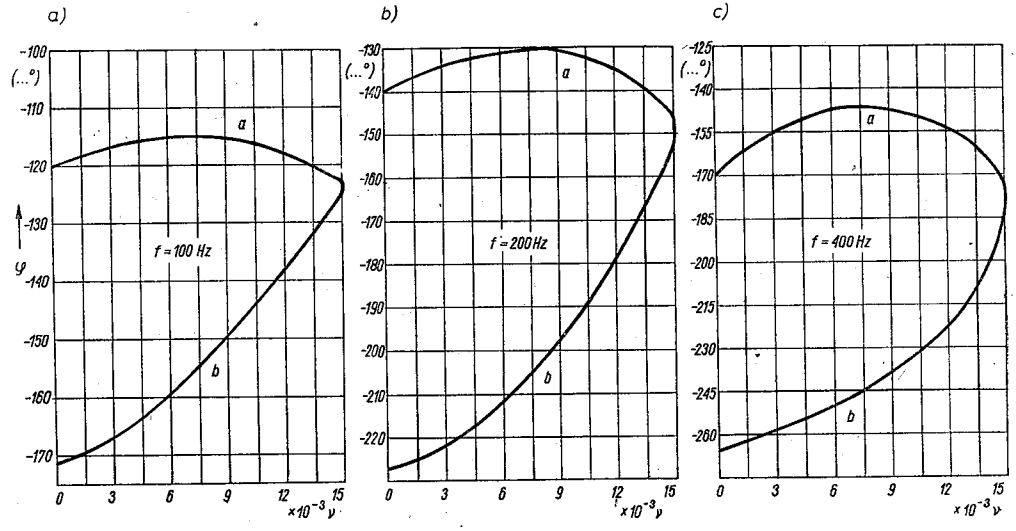
Rys. 6. Wykres modułu gęstości prądu na linii $\lambda = (c_2 + c_3)/2$
a) gałąź krzywej odpowiadająca $y > 0$, b) gałąź krzywej odpowiadająca $y < 0$

Obliczenia przeprowadzono dla następujących danych:

$$\gamma = 5{,}7 \cdot 10^7 \Omega^{-1} m^{-1}, \quad \lambda_m = 3 \cdot 10^{-2}, \quad \mu = 4\pi \cdot 10^{-7} \text{ H/m},$$

$$v_m = 8 \cdot 10^{-3}, \quad c_1 = 15 \cdot 10^{-3} \text{ m}, \quad c_2 = 18 \cdot 10^{-3} \text{ m}, \quad c_3 = 20{,}5 \cdot 10^{-3} \text{ m}.$$

Rozkłady modułu gęstości prądu oraz fazy przedstawione są na rys. 6 i 7.



Rys. 7. Wykres kąta przesunięcia fazowego wektora gęstości prądu na linii $\lambda = (c_2 + c_3)/2$
a) gałąź krzywej odpowiadająca $y > 0$, b) gałąź krzywej odpowiadająca $y < 0$

6. WYZNACZENIE WSPÓŁCZYNNIKA EKRANOWANIA

Współczynnik ekranowania zdefiniowano jako

$$k = \frac{\underline{B}}{\underline{B}_o},$$

gdzie:

$\underline{B}$ — wartości indukcji w początku układu współrzędnych (rys. 2) w przypadku obecności ekranu,

$\underline{B}_o$ — wartość indukcji w początku układu współrzędnych wytworzonej przez zwój kołowy przy nieobecności ekranu.

Wielkość B można wyznaczyć stosując superpozycję pól wytworzonych przez elementarne zwoje (rys. 5) traktowane jako nieskończenie cienkie zwoje kołowe

$$\underline{B} = \sum_{F} \frac{\mu_o \underline{J}_i \, dF_i}{4\pi(x_{dF_i}^2 + y_{dF_i}^2)} + \underline{B}_0,$$

gdzie:

$$\underline{B}_o = \frac{\mu_o \underline{J}}{4\pi(x_m^2 + y_m^2)}$$

x_{dF_i}, y_{dF_i} — współrzędne kartezjańskie środka i-tego zwoju elementarnego.

Zmiany współczynnika ekranowania dla danych z p. 5 w zależności od częstotliwości ilustruje poniższa tabelka

f	k
100 Hz	0,85
200 Hz	0,64
400 Hz	0,40

7. ZAKOŃCZENIE

W pracy zastosowano metodę równania całkowego do analizy ekranu elektromagnetycznego o kształcie elipsoidy obrotowej. Z matematycznego punktu widzenia sprowadzono problem do rozwiązania równania całkowego Fredholma II rodzaju (równanie (11)). Podano sposób algorytmizacji obliczeń przy wykorzystaniu metody Fredholma jako metody przybliżonego rozwiązania równania (11). Należy podkreślić, że metoda równania całkowego prowadzi w badanym przypadku do prostszego algorytmu obliczeniowego, aniżeli inne powszechnie stosowane metody.

LITERATURA

1. M, Abramowitz, J.A. Stegun, *Handbook of mathematical functions*, Dover Publication Inc., 1965.
2. A. Kost, *Wirbelstrom Verteilung im gestreckten Rotationsellipsoid*, Archiv für Elektrotechnik, 60, 1978.

3. M. Krakowski, *Elektrotechnika teoretyczna*, t. 2, PWN, W-wa 1978.
4. W. Pogorzelski, *Równania całkowe i ich zastosowania*, PWN, 1953.
5. R. Sikora, *Teorie pola elektromagnetycznego*, WNT, 1977.
6. J. Turowski, *Elektrodynamika techniczna*, WNT, W-wa 1968.
7. S.G. Michlin, C.L. Smolicki, *Metody przybliżone rozwiązywania równań różniczkowych i całkowych*, PWN, W-wa 1972.

P. KRASOŃ

APPLICATION OF INTEGRAL EQUATIONS METHOD TO ANALYSIS OF ELIPSOIDAL ELECTROMAGNETIC SCREEN

Summary

The analysis of decomposition of current density vector inside an elipsoidal electromagnetic screen using integral equation method is carried out in this paper. Screening coefficient calculation method is given also. Analysis was carried on by reducing electromagnetic effects description to Fredholm II integral equation of the 2nd kind. Formulas allowing fast construction of numerical algorithm for computer purposes are calculated in the paper, and that is why it shows its particular usefulness of the integral equation method for axial-symmetrical screening problems.

P. KRASOŃ

APPLICATION DE LA MÉTHODE DES ÉQUATIONS INTÉGRALES À L'ANALYSE DE L'ÉCRAN ÉLECTROMAGNÉTIQUE ÈLLIPSOÏDAL

Résumé

Dans l'étude on a analysé la répartition du vecteur de densité de courant intérieur électromagnétique et ellipsoïdal par la méthode de l'équation intégrale. On a présenté aussi le mode de détermination du coefficient de blindage. L'analyse a été faite en amenant la description des phénomènes électromagnétiques à la forme de l'équation intégrale de Fredholm II-e genre. On a introduit des formules rendant possible une rapide détermination de l'algorithme de calcul sur la EMC en démontrant, en même temps, une utilité toute particulière de la méthode de l'équation intégrale aux problèmes de blindage axialement symétrique.

P. KRASOŃ

ANWENDUNG DER METHODE DER INTEGRALGLEICHUNGEN FÜR DIE ANALYSE DER ELEKTROMAGNETISCHEN ELLIPSOIDEN ABSCHIRMUNG

Zusammenfassung

In der Arbeit wurde eine Analyse der Verteilung des Stromdichtevektors innerhalb der elektromagnetischen ellipsoiden Abschirmung mit Hilfe der Integralgleichung durchgeführt. Es wurde auch die Methode für die Bestimmung des Abschirmungskoeffizienten entwickelt. Bei der Analyse wurde die Beschreibung der elektromagnetischen Erscheinungen auf die Lösung der Fredholmschen Integralgleichung zweiter Art reduziert. Die angegebenen Formeln ermöglichen eine schnelle numerische Berechnung des Algorithmus und sind sehr günstig beim Lösen axialsymmetrischer Aufgaben der elektromagnetischen Abschirmung.

П. КРАСОНЬ

ПРИМЕНЕНИЕ МЕТОДА ИНТЕГРАЛЬНЫХ УРАВНЕНИЙ К АНАЛИЗУ ЭЛЛИПСОИДООБРАЗНОГО ЭЛЕКТРОМАГНИТНОГО ЭКРАНА

Р е з ю м е

Проведен анализ распределения вектора плотности тока внутри электромагнитного эллипсоидообразного экрана методом интегрального уравнения. Приведен также способ вычисления коэффицента экранирования. Анализ проведен сведением описания электромагнитных явлений к интегральному уравнению Фредгольма второго рода. Выведены формулы дающие возможность быстро написать алгоритм для вычисления на ЦВМ и тем самым доказана особенная пригодность метода интегрального уравнения для решения осевосимметрических вопросов экранирования.

ROZPRAWY ELEKTROTECHNICZNE 1982, 28, z. 3-4, ss. 355—365

621.3.014.4:621.3.014.12

Zjawisko naskórkowości i prądy wirowe w cienkiej taśmie metalowej przy wymuszeniu sinusoidalnym

HANNA MORAWSKA (ŁÓDŹ)

Instytut Podstaw Elektrotechniki Politechniki Łódzkiej

MACIEJ KRAKOWSKI (ŁÓDŹ)

Instytut Informatyki Politechniki Łódzkiej

Otrzymano 30.7.1980

Przedstawiono metodę obliczania gęstości prądu w cienkiej taśmie metalowej, w której płynie prąd sinusoidalny. Wyprowadzono wzory na rezystancję i indukcyjność wewnętrzną cienkiej taśmy oraz podano ich charakterystyki częstotliwościowe. Przedstawiono metodę obliczania gęstości prądu i strat wiroprądowych w cienkiej taśmie umieszczonej w poprzecznym polu magnetycznym. Omówiono taśmę z prądem umieszczoną w polu magnetycznym i wyprowadzono wzór na straty mocy w tej taśmie.

1. WSTĘP

Analiza pola elektromagnetycznego w cienkiej taśmie metalowej przy przepływie prądu sinusoidalnego była od szeregu lat przedmiotem wielu badań naukowych. W pracy [2] sprowadzono zagadnienie do równań całkowych, lecz nie wykonano obliczeń liczbowych. W pracy [5] wyznaczono gęstość prądu w taśmie, rozwiązując równanie całkowe na drodze rozwinięcia jądra tego równania w szereg potęgowy, w [1, 10] zastąpiono taśmę przez nieograniczoną płytę i rozpatrzono jedynie zmianę gęstości prądu wzdłuż jej grubości. W pracy [6] przedstawiono metodę przybliżoną, w której zastosowano metodę kolejnych reakcji prądów wirowych, a w pracy [7] zastosowano metodę równań całkowych. Analizę prądów wirowych w cienkiej taśmie umieszczonej w poprzecznym polu magnetycznym przeprowadzono w [4, 8, 9].

Niniejsza praca zawiera rozwiązanie analityczne podane w pracy [8]. W rozważaniach pominięto prądy przesunięcia i rozpatrzono nieskończenie długą taśmę o grubości niewielkiej w porównaniu z głębokością wnikania fali elektromagnetycznej. Założono, że przenikalność magnetyczna taśmy metalowej jest równa μ_0.

Wszystkie wielkości charakteryzujące pole elektromagnetyczne zależą od czasu jak $\exp(j\omega t)$ i są przedstawione w postaci zespolonej. Wektory zespolone są elementami

trójwymiarowej zespolonej przestrzeni euklidesowej, a ich normę przedstawia wzór $||\mathbf{W}|| = \sqrt{\mathbf{W} \cdot \mathbf{W}^*}$, przy czym $\mathbf{W}^*$ jest wektorem sprzężonym z $\mathbf{W}$. Przyjęto, że norma wektorów zespolonych jest równa wartości skutecznej.

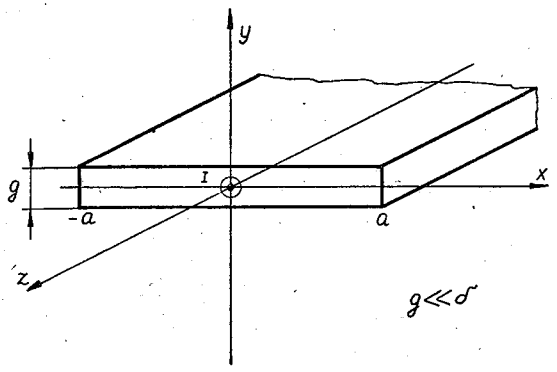
Wykaz najważniejszych oznaczeń

$\mathbf{B}$ — indukcja magnetyczna
$\mathbf{E}$ — natężenie pola elektrycznego
g — grubość taśmy
I — wartość zespolona prądu
$\mathbf{J}$ — gęstość prądu
k^2 — $j\omega\mu_o\gamma$
P — straty mocy, straty wiroprądowe
γ — konduktywność taśmy
ω — pulsacja

2. ZJAWISKO NASKÓRKOWOŚCI W CIENKIEJ TAŚMIE METALOWEJ

2.1. Zależności podstawowe

Cienka taśma metalowa wiodąca prąd I jest umieszczona w płaszczyźnie x, z prostokątnego układu współrzędnych (rys. 1). Przyjmuje się, że taśma metalowa jest bardzo



Rys. 1. Cienka taśma metalowa wiodąca prąd

cienka, wobec czego pole w jej wnętrzu nie zależy od y. Gęstość prądu i natężenie pola elektrycznego mają tylko składowe $J_z(x)$ i $E_z(x)$. Na podstawie równań Maxwella w postaci zespolonej otrzymuje się równanie Helmholtza

$$\frac{d^2E_z}{dx^2} - k^2E_z(x) = 0. \tag{1}$$

Ze względu na symetrię układu, rozwiązanie tego równania przyjmuje się w postaci

$$E_z(x) = C\mathrm{ch}kx,$$

gdzie $k = \sqrt{\omega\mu_0\gamma}\ \exp\left(j\frac{\pi}{4}\right)$, zaś stałą C wyznacza się na podstawie warunku

$$g\int_{-a}^{a}\gamma E_z(x)\,dx = I,$$

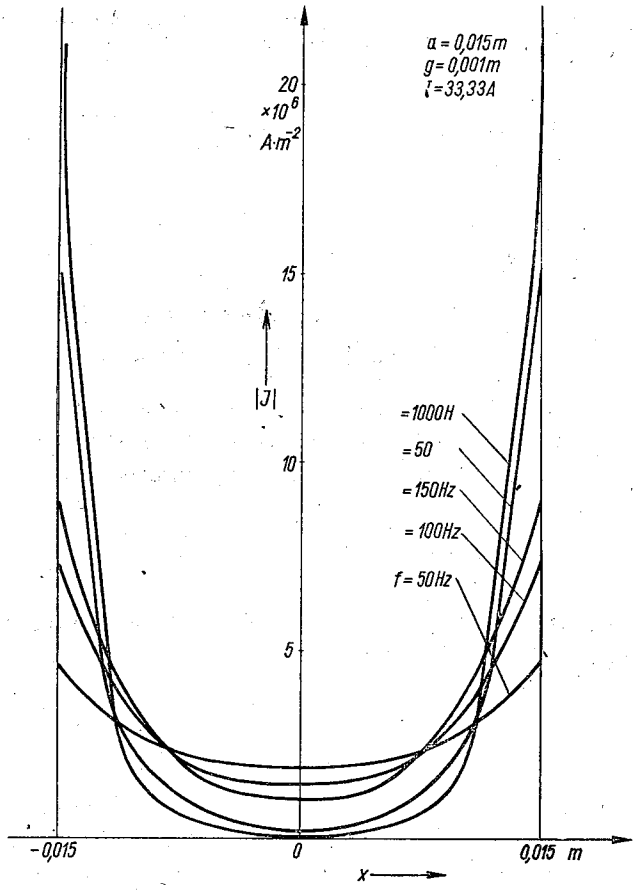
który oznacza, że prąd całkowity w taśmie jest równy I. W wyniku otrzymuje się

$$E_z(x) = \frac{kI}{2g\gamma}\frac{\operatorname{ch}kx}{\operatorname{sh}ka}, \qquad x \in (-a, a), \tag{2}$$

a stąd

$$J_z(x) = \gamma E_z(x) = \frac{kI}{2g}\frac{\operatorname{ch}kx}{\operatorname{sh}ka}, \qquad x \in (-a, a). \tag{3}$$

Wzór (3) pozwala obliczyć rozkład prądu w taśmie. Na rys. 2 przedstawiono rozkład prądu w taśmie przy różnych częstotliwościach.



Rys. 2. Rozkład prądu w cienkiej taśmie przy różnych częstotliwościach

2.2. Rezystancja i indukcyjność wewnętrzna

Moc czynna pobierana przez taśmę przy przepływie prądu (na jednostkę długości)

$$P = \frac{1}{\gamma}\int_S ||\mathbf{J}||^2 dS = \frac{1}{\gamma}\int_{-a}^{a} \frac{kI}{2g}\frac{\operatorname{ch}kx}{\operatorname{sh}ka}\frac{k^*I^*}{2g}\frac{\operatorname{ch}k^*x}{\operatorname{sh}k^*a}g\,dx \tag{4}$$

przy uwzględnieniu wzoru (3), gdzie S — powierzchnia przekroju poprzecznego taśmy. Stąd po wykonaniu całkowania znajduje się na podstawie zależności $R = P/|I|^2$ wzór na rezystancję taśmy przy prądzie sinusoidalnym (na jednostkę długości)

$$R = \frac{\omega\mu_o}{2\sqrt{2}\,g|k|}\,\frac{\operatorname{sh}(|k|a\sqrt{2})+\sin(|k|a\sqrt{2})}{\operatorname{ch}(|k|a\sqrt{2})-\cos(|k|a\sqrt{2})}. \tag{5}$$

Na podstawie II prawa Maxwella otrzymuje się

$$H_y(x) = \frac{1}{j\omega\mu_o}\frac{dE_z}{dx} = \frac{I}{2g}\frac{\operatorname{sh}kx}{\operatorname{sh}ka}, \qquad x\in(-a,a).$$

Moc bierną związaną z polem magnetycznym we wnętrzu taśmy (na jednostkę długości) oblicza się z wzoru

$$Q = 2\omega\int_S \frac{1}{2}\mu_o||\mathbf{H}||^2 dS = \frac{\omega\mu_o|I|^2}{4g^2}\int_{-a}^{a}\frac{\operatorname{sh}kx}{\operatorname{sh}ka}\frac{\operatorname{sh}k^*x}{\operatorname{sh}k^*a}\,g\,dx.$$

Po wykonaniu całkowania i uwzględnieniu wyrażenia $L_w = Q/(\omega|I|^2)$ otrzymuje się wzór na indukcyjność wewnętrzną taśmy metalowej na jednostkę długości

$$L_w = \frac{\mu_o}{2\sqrt{2}\,|k|g}\,\frac{\operatorname{sh}(|k|a\sqrt{2})-\sin(|k|a\sqrt{2})}{\operatorname{ch}(|k|a\sqrt{2})-\cos(|k|a\sqrt{2})}. \tag{6}$$

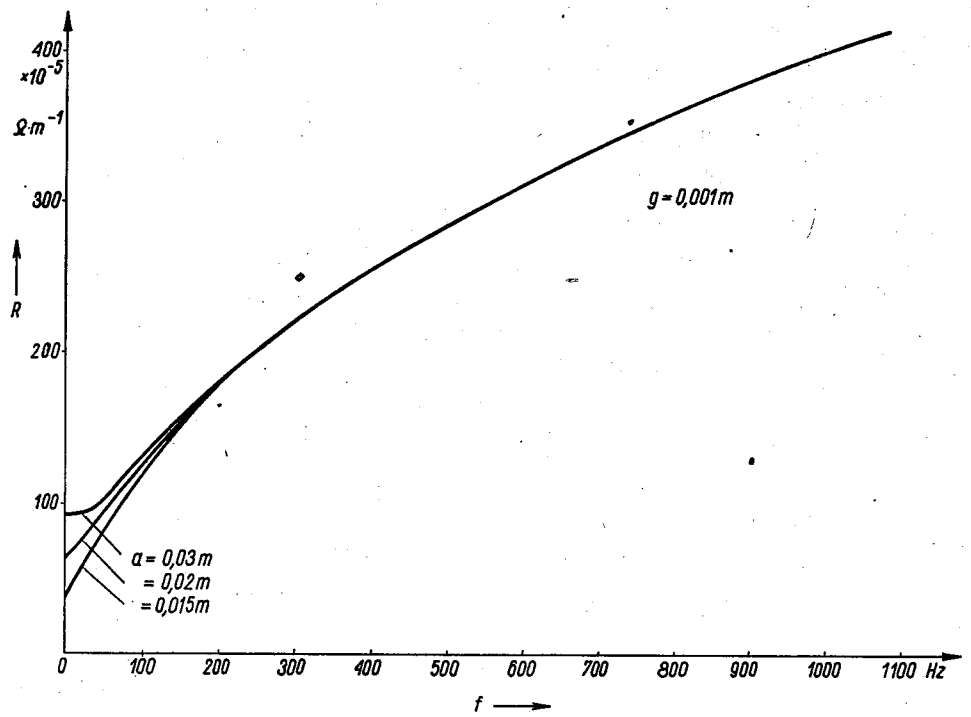
Zatrzymując dwa pierwsze wyrazy w rozwinięciach w szeregi potęgowe funkcji trygonometrycznych i hiperbolicznych, a następnie podstawiając je do (5) i (6), znajduje się wzory na rezystancję i indukcyjność wewnętrzną taśmy przy prądzie stałym

$$R_o = \frac{1}{2ag\gamma}, \qquad L_{wo} = \frac{\mu_o a}{6g}. \tag{7}$$

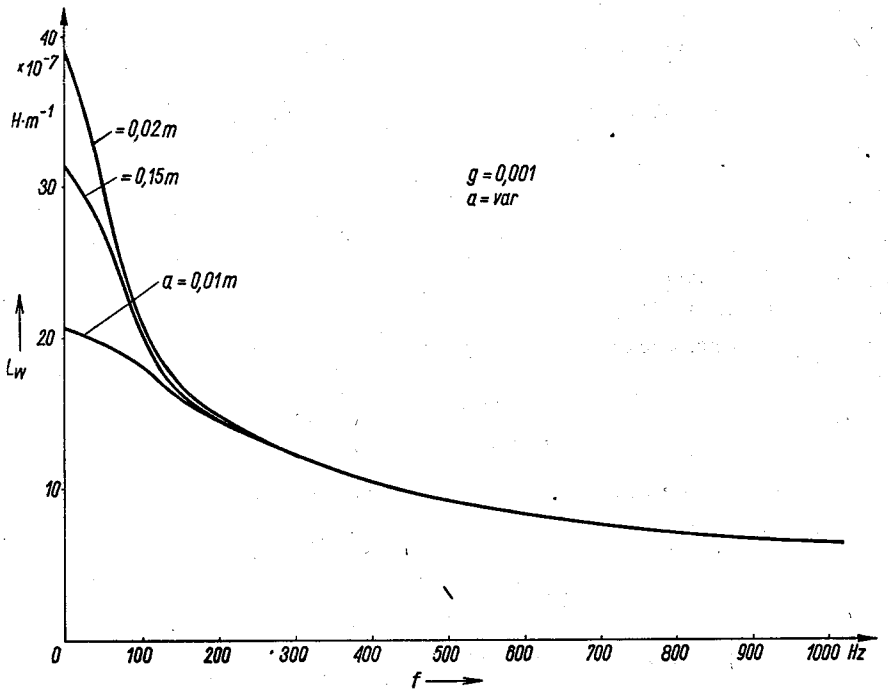
Przy $|k|a\sqrt{2} > 5$, ułamek zawierający funkcje hiperboliczne w (5) i (6) jest w przybliżeniu równy 1, wobec czego

$$R \approx \frac{\omega\mu_o}{2\sqrt{2}\,|k|g}, \qquad L_w \approx \frac{\mu_o}{2\sqrt{2}\,|k|g}. \tag{8}$$

Oznacza to, że przy dostatecznie dużych częstotliwościach rezystancja i indukcyjność wewnętrzna taśmy nie zależą od jej szerokości. Jest to oczywiste z fizycznego punktu widzenia, gdyż w tych warunkach prąd płynie jedynie przy brzegach taśmy, nie wykorzystując całej powierzchni przekroju poprzecznego. Jak wynika z wzorów (8), przy uwzględ-



Rys. 3. Charakterystyki częstotliwościowe rezystancji taśmy



Rys. 4. Charakterystyki częstotliwościowe indukcyjności wewnętrznej taśmy

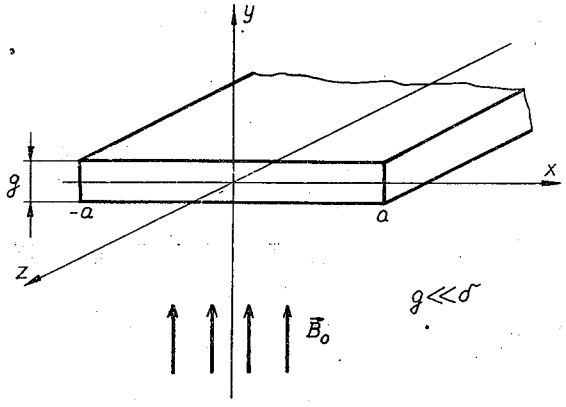
nieniu zależności $|k| = \sqrt{\omega\mu_0\gamma}$, rezystancja taśmy jest proporcjonalna, a jej indukcyjność odwrotnie proporcjonalna do $\sqrt{\omega}$.

Charakterystyki częstotliwościowe rezystancji i indukcyjności wewnętrznej taśm metalowych o jednakowej grubości i różnych szerokościach są przedstawione na rys. 3 i 4.

3. TAŚMA W POPRZECZNYM POLU MAGNETYCZNYM

3.1. Zależności podstawowe

Metalowa taśma jest umieszczona w zewnętrznym równomiernym polu magnetycznym o indukcji magnetycznej $\mathbf{B}_o = 1_y B_o$ (rys. 5). Równania opisujące zewnętrzne pole



Rys. 5. Cienka taśma metalowa w równomiernym polu magnetycznym

elektromagnetyczne w obszarze zajmowanym przez taśmę, jeśli jest ona usunięta, są następujące:

$$\nabla \times \mathbf{H}_o = 0, \tag{9a}$$

$$\nabla \times \mathbf{E}_o = -j\omega \mathbf{B}_o. \tag{9b}$$

Indukowane w taśmie prądy wirowe wytwarzają wtórne pole magnetyczne, opisane za pomocą wektorów **E, H, B, J**. Pole wypadkowe jest superpozycją pola zewnętrznego i pola wtórnego, wobec czego równania tego pola są następujące:

$$\nabla \times (\mathbf{H}_o + \mathbf{H}) = \mathbf{J}, \tag{10a}$$

$$\nabla \times (\mathbf{E}_o + \mathbf{E}) = -j\omega(\mathbf{B}_o + \mathbf{B}). \tag{10b}$$

Po uwzględnieniu zależności $\mathbf{J} = \gamma(\mathbf{E}_o + \mathbf{E})$ oraz (9a), równania opisujące pole wtórne przybierają postać

$$\nabla \times \mathbf{H} = \mathbf{J}, \tag{11a}$$

$$\nabla \times \mathbf{J} = -j\omega\gamma(\mathbf{B} + \mathbf{B}_o) \tag{11b}$$

przy założeniu, że $\gamma = \text{const}$. Na podstawie powyższych równań otrzymuje się

$$\nabla \times \nabla \times \mathbf{H} = -k^2 \mathbf{H} - j\omega\gamma \mathbf{B}_o,$$

przy czym

$$k^2 = j\omega\mu_o\gamma. \tag{12}$$

Uwzględniając tożsamość wektorową $\nabla\times\nabla\times\mathbf{H} = \nabla(\nabla\cdot\mathbf{H}) - \nabla^2\mathbf{H}$ oraz wzór $\nabla\cdot\mathbf{H} = 0$ prawdziwy przy μ = const, otrzymuje się w wyniku

$$\nabla^2\mathbf{H} - k^2\mathbf{H} = j\omega\gamma\mathbf{B}_o. \tag{13}$$

3.2. Gęstość prądu i straty wiroprądowe

Gęstość prądu **J** w taśmie jest równoległa do osi Oz, a składowa H_y we wnętrzu bardzo cienkiej taśmy nie zależy od y i jest funkcją jednej zmiennej x. Równanie (13) dla składowej H_y przybiera postać

$$\nabla^2 H_y - k^2 H_y = j\omega\gamma B_o,$$

czyli

$$\frac{d^2H_y}{dx^2} - k^2 H_y = j\omega\gamma B_o. \tag{14}$$

Całkowity prąd w taśmie jest równy zeru, a warunek brzegowy [4]

$$H_y(-a) = H_y(a) = 0. \tag{15}$$

Rozwiązaniem omawianego zagadnienia brzegowego jest

$$H_y(x) = \frac{B_o}{\mu_o}\left(\frac{\operatorname{ch} kx}{\operatorname{ch} ka} - 1\right), \qquad x \in (-a, a), \tag{16}$$

przy czym $k = \sqrt{\omega\mu_o\gamma}\exp(j\pi/4)$.

Gęstość prądu wyznacza się z (11a), otrzymując

$$J_z(x) = \frac{dH_y}{dx} = \frac{kB_o}{\mu_o}\frac{\operatorname{sh} kx}{\operatorname{ch} ka}, \qquad x \in (-a, a). \tag{17}$$

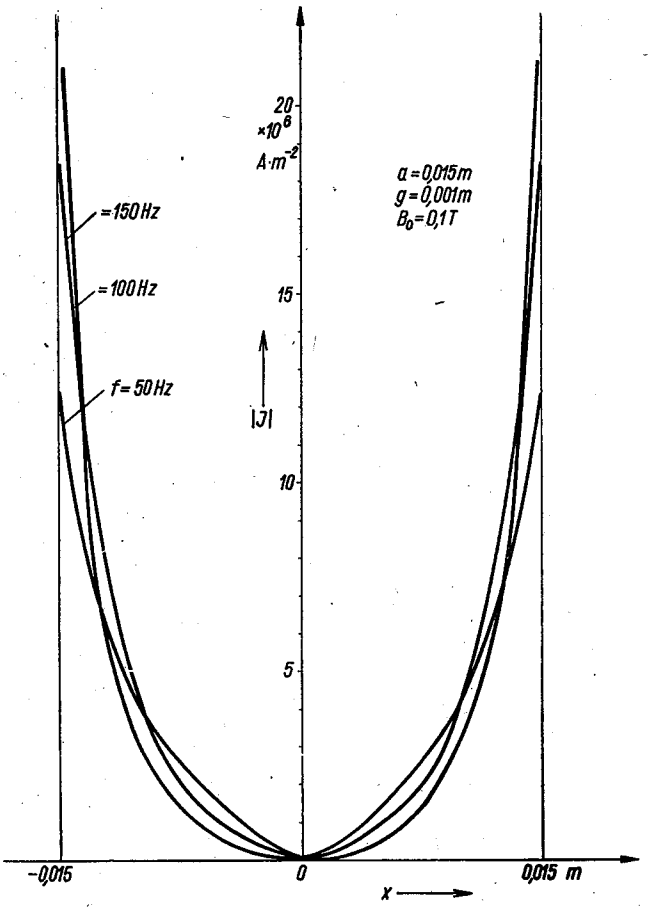
Rozkład prądu w taśmie jest przedstawiony na rys. 6. Straty wiroprądowe w taśmie przedstawia wzór

$$P = \frac{g}{\gamma}\int_{-a}^{a} |J_z(x)|^2 dx, \tag{18}$$

przy czym $|J_z(x)|^2 = J_z(x)\,J_z^*(x)$.

Podstawiając (17) do (18) otrzymuje się w wyniku wzór na straty wiroprądowe w taśmie umieszczonej w poprzecznym polu magnetycznym

$$P_e = \frac{\sqrt{2}|k|g|B_o|^2}{\mu_o^2\gamma}\,\frac{\operatorname{sh}(|k|a\sqrt{2}) - \sin(|k|a\sqrt{2})}{\operatorname{ch}(|k|a\sqrt{2}) + \cos(|k|a\sqrt{2})}. \tag{19}$$



Rys. 6. Rozkład prądu w cienkiej taśmie przy $B_o = 0{,}1$ T

4. TAŚMA Z PRĄDEM W POPRZECZNYM POLU MAGNETYCZNYM

Przyjmuje się, że w taśmie z rys. 5 płynie prąd I w stronę osi Oz. Pole elektromagnetyczne w taśmie jest superpozycją pola wywołanego przez prąd I płynący w taśmie przy braku zewnętrznego pola magnetycznego (p. 2) i pola wywołanego przez zewnętrzne pole magnetyczne przy braku prądu w przewodzie (p. 3), wobec czego

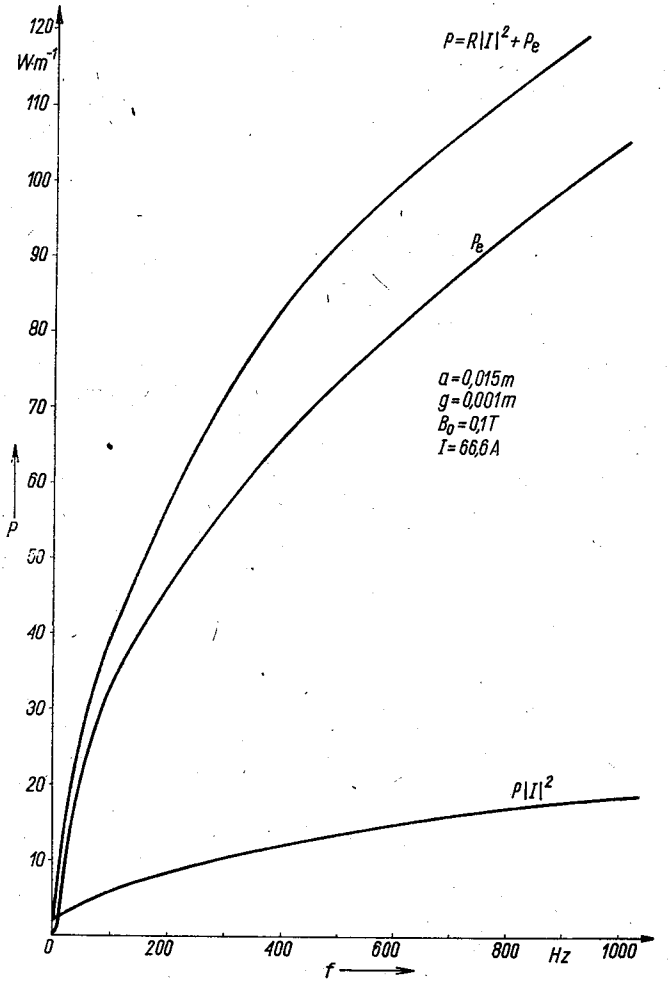
$$J_z(x) = \frac{kI}{2g}\frac{\operatorname{ch} kx}{\operatorname{sh} ka} + \frac{kB_o}{\mu_o}\frac{\operatorname{sh} kx}{\operatorname{ch} ka}, \qquad x \in (-a, a). \tag{20}$$

Straty mocy w taśmie otrzymuje się, podstawiając (20) do (18); po wykonaniu prostych obliczeń otrzymuje się

$$P = R|I|^2 + P_e, \tag{21}$$

gdzie R oznacza rezystancję przewodu taśmowego (wzór 5), natomiast P_e przedstawia wzór (19).

Straty mocy w taśmie z prądem umieszczonej w zewnętrznym polu magnetycznym są równe sumie mocy $P|I|^2$ wywołanej przez prąd w przewodzie taśmowym oraz strat wiroprądowych wywołanych przez zewnętrzne pole magnetyczne. Straty mocy w taśmie z prądem, umieszczonej w polu magnetycznym w funkcji częstotliwości są przedstawione na rys. 7.



Rys. 7. Straty mocy w taśmie z prądem umieszczonej w polu magnetycznym w funkcji częstotliwości

LITERATURA

1. W. A. Goworkow, *Pola elektryczne i magnetyczne*, WNT, Warszawa 1962.
2. H. G. Gross, *Die Berechnung der Stromverteilung in zylindrischen Leitern mit rechteckigen und elliptischen Querschnitt*, Archiv für Elektrotechnik, t. 34, z. 5, 1940, s. 241—268.
3. M. Krakowski, *Elektrotechnika teoretyczna*, t. 2, PWN, Warszawa 1979.
4. M. Krakowski, *Eddy currents in thin plates induced by time-harmonic magnetic field*, Bulletin de l'Academie des Sciences, Serie des sciences techniques, t. 27, z. 12, 1979, s. 949—963.

5. A. Krawczyk, *Uproszczone obliczanie gęstości prądu i temperatury w przewodzie taśmowym*, Zeszyty Naukowe PŁ „Elektryka", z. 56, 1976, s. 151—161.
6. W. Lipiński, *Analiza rozkładu gęstości prądu w wydłużonej szynie prostokątnej*, Archiwum Elektrotechniki, t. XXIII, z. 1, 1974, s. 49—57.
7. W. Lipiński, M. Gramz, P. Krasoń, *Ein Beitrag zur Berechnung der Stromverdrängung in einem Bandleiter*, Archiv für Elektrotechnik, t. 62, 1980, s. 51—55.
8. H. Morawska-Koraszewska, *Prądy wirowe w szynie metalowej umieszczonej w równomiernym polu magnetycznym*, Praca doktorska, Politechnika Łódzka, 1979.
9. J. Sonntag, *Abschirmung einer Doppelleitung durch einen parallelen Bandleiter*, Archiv für Elektrotechnik, t. 61, 1979, s. 117—121.
10. R. L. Stoll, *The analysis of eddy currents*, Clarendon Press, Oxford 1974.

H. MORAWSKA, M. KRAKOWSKI

SKIN EFFECT AND EDDY CURRENTS IN A THIN METALLIC TAPE DUE TO SINUSOIDAL EXCITATION

Summary

A method of calculation of current density in a metallic tape carrying a.c. is presented. Relationships for the resistance and the internal inductance of the tape are derived. A method of calculation of current density and eddy-current losses in a thin metallic tape situated in a transverse magnetic field is presented. A tape carrying a.c. in a magnetic field is considered. The power losses in this tape are derived.

H. MORAWSKA, M. KRAKOWSKI

SKIN-EFFECT ET LES COURANTS DE FOUCAULT DANS UNE BANDE FINE DE MÉTAL AVEC L'EXCITATION SINUSOÏDALE

Résumé

On a décrit la méthode de calcul de la densité du courant dans une bande fine de métal parcourue par le courant sinusoïdal. On a établi les formules qui déterminent la résistance et l'inductance intérieure de la bande fine et on a donné leurs réponses en fréquence. On a présenté la méthode de caclul de la densité du courant et des pertes de Foucault dans une bande fine placée dans le champ magnétique transversal. On a décrit aussi la bande avec le courant placée dans le champ magnétique et on a établi la formule qui détermine des pertes de puissance dans cette bande.

H. MORAWSKA, M. KRAKOWSKI

SKINEFFEKT UND WIRBELSTRÖME IN EINEM DÜNNEN BANDLEITER MIT SINUSFÖRMIGER ERREGUNG

Zusammenfassung

Es wurde die Methode der Berechnung der Stromdichte in einem dünnen, von dem sinusförmigen Strom durchflossenen metallischen Bandleiter dargestellt. Die Formeln für den Wechselstromwiderstand und die innere Induktivität dieses Leiters, sowie ihre Frequenzkennlinien wurden angegeben. Es wurde die Methode der Berechnung der Stromdichte und der Wirbelstromverluste in einem im magnetischen Feld untergebrachten Bandleiter dargestellt. Der stromführende, im magnetischen Feld untergebrachte Bandleiter wurde betrachtet und die Formel für die Wirbelstromverluste wurde angegeben.

Х. МОРАВСКА, М. КРАКОВСКИ

ПОВЕРХНОСТНЫЙ ЭФФЕКТ И ВИХРЕВЫЕ ТОКИ В ТОНКОЙ МЕТАЛЛИЧЕСКОЙ ЛЕНТЕ ПРИ СИНУСОИДАЛЬНОМ ВОЗМУЩЕНИИ

Р е з ю м е

Представлен метод вычисления плотности тока в тонкой металлической ленте ведущей синусоидальный ток. Определены формулы для активного сопротивления и внутренней индуктивности тонкой ленты. Представлен метод вычисления плотности тока в тонкой металлической ленте расположенной в поперечном магнитом поле. Рассмотрена лента ведущая ток, расположенная в магнитном поле. Определены формулы для потерь активной мощности в этой ленте.

ROZPRAWY ELEKTROTECHNICZNE 1982, 28, z. 3-4, ss. 367—383

621.317.322.029.4

Metoda kompensacyjna precyzyjnego pomiaru małych wartości skutecznych odkształconego napięcia przemiennego małej częstotliwości

MICHAŁ LISOWSKI (WROCŁAW)

Instytut Metrologii Elektrycznej Politechniki Wrocławskiej

Otrzymano 10.8.1980

W pracy opisano metodę kompensacji odkształconego napięcia przemiennego przez odwzorowanie w napięciu kompensacyjnym — za pomocą wzmacniacza — kształtu mierzonego napięcia. Metoda ta, analogiczna do metody kompensacji odkształconego prądu przemiennego opisanej przez autora w pracy [15], umożliwia precyzyjne pomiary wartości skutecznej napięcia bez poboru energii z mierzonego źródła. Przeprowadzono analizę metrologiczną dwóch układów kompensacyjnych: o przetwarzaniu napięcie — prąd i napięcie — napięcie. Przeanalizowano wpływ zmian wartości i częstotliwości mierzonego napięcia na równowagę układu oraz czułość i dokładność pomiaru. Podano parametry i wyniki badań wykonanego kompensatora.

Wykonany kompensator przy współpracy z komparatorem prądu AC/DC w układzie o przetwarzaniu napięcie — prąd lub komparatorem napięcia AC/DC w układzie o przetwarzaniu napięcie — napięcie, umożliwia precyzyjne pomiary napięcia od 0,25 mV do 250 mV, w zakresie częstotliwości 40÷1000 Hz. Współczynnik przetwarzania układów został określony z niedokładnością 0,01%. Opracowany kompensator umożliwia tysiąckrotne rozszerzenie w dół zakresu napięciowego komparatora AC/DC lub pomiary małych napięć przy pracy komparatora AC/DC na zakresach prądowych.

1. WSTĘP

Najdokładniejszymi przyrządami umożliwiającymi pomiary wartości skutecznych napięć odkształconych są komparatory AC/DC, porównujące prąd przemienny z prądem stałym [1, 5, 6, 7, 8, 18, 20, 22, 23]. Komparatory te umożliwiają pomiary napięć odkształconych z niedokładnością ok. 0,01%, lecz nie nadają się do pomiarów małych napięć, gdyż ich dolne zakresy pomiarowe wynoszą ok. 0,25 V. Ponadto urządzenia te obciążają mierzone źródło.

Obniżenie dolnego zakresu pomiarowego komparatorów AC/DC częściowo umożliwiają precyzyjne wzmacniacze pomiarowe [3, 4, 10, 11]. Jednak ze wzrostem wzmocnienia następuje spadek dokładności pomiaru. Przy wzmocnieniu około 100 błąd ten wynosi około 0,1%.

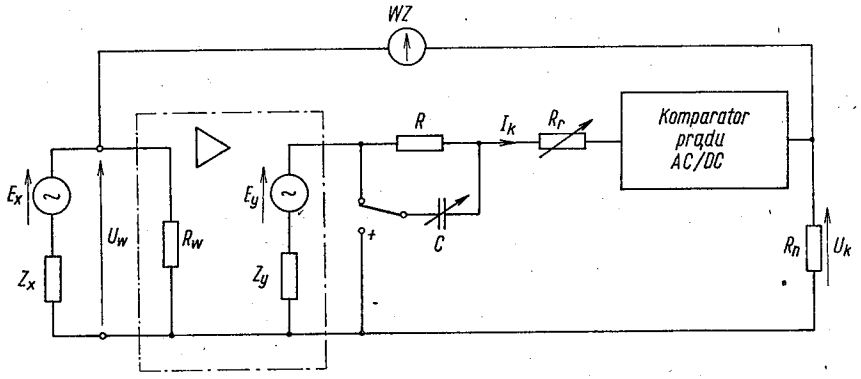
Małe wartości napięć prądu przemiennego można mierzyć również kompensatorami [5, 13, 14, 17, 19, 21]. Umożliwiają one pomiary bez poboru energii z mierzonego źródła. Kompensatory te przy prądzie przemiennym mają stosunkowo małą dokładność. Zasad-

niczą przyczyną błędów są różne kształty przebiegów czasowych wielkości mierzonej i kompensującej. Zgodność przebiegów czasowych można uzyskać zastępując niezależne źródło prądu pomocniczego źródłem sterowanym sygnałem mierzonym. Metoda ta była stosowana wcześniej przez autora w kompensatorze prądowym prądu przemiennego [15] oraz w kompensacyjnych woltomierzach cyfrowych wartości skutecznej [12]. Zasadę działania takiego woltomierza w literaturze polskiej przedstawiono w [9]. Jest to automatyczny kompensator o stałym prądzie pomocniczym i zmiennej rezystancji kompensacyjnej.

W układach kompensacyjnych do rozszerzenia zakresu pomiarowego komparatorów znacznie wygodniejsze jest zastosowanie stałej rezystancji kompensacyjnej i regulowanego prądu pomocniczego. Układ taki może przejąć zadanie wzmacniacza pomiarowego i pełnić funkcję przetwornika, przetwarzając mierzone napięcie na prąd lub na napięcie o takiej wartości, którą można zmierzyć komparatorem AC/DC [16].

2. KOMPENSATOR O PRZETWARZANIU NAPIĘCIE — PRĄD

Schemat ideowy kompensatora z odtwarzaniem kształtu mierzonego napięcia w napięciu kompensującym przedstawiono na rys. 1. Zródło mierzonego napięcia E_x steruje



Rys. 1. Schemat kompensatora z odtwarzaniem kształtu przebiegu czasowego mierzonego napięcia w napięciu kompensacyjnym o przetwarzaniu napięcie — prąd

WZ — wskaźnik zera; ▷ — wzmacniacz; E_x — mierzone napięcie; U_w — napięcie wejściowe wzmacniacza; E_y — napięcie źródła na wyjściu wzmacniacza; I_k — prąd w obwodzie komparatora; R_n — rezystor kompensacyjny; Z_x — impedancja mierzonego źródła; R_w — rezystancja wejściowa wzmacniacza; Z_y — impedancja wyjściowa wzmacniacza; R_r — rezystor regulacyjny; R, C — rezystor i kondensator do korekcji przesunięć fazowych; *AC* — prąd przemienny; *DC* — prąd stały

wejściem wzmacniacza o bardzo dużej rezystancji wejściowej $R_w \approx 10^{12}\ \Omega$. Wyjście tego wzmacniacza jest źródłem prądu I_k o kształcie praktycznie zgodnym z kształtem mierzonego napięcia E_x. Prąd I_k przepływając przez rezystor R_n wytwarza na nim napięcie kompensujące U_k. Układ równoważy się ręcznie zmieniając moduł prądu I_k rezystorem R_r, a fazę kondensatorem C. W stanie równowagi wskaźnik zera *WZ* wskaże brak sygnału. Wówczas napięcie U_k na rezystorze R_n będzie równe napięciu U_w, czyli

$$\underline{U}_w = \underline{I}_k R_n. \tag{1}$$

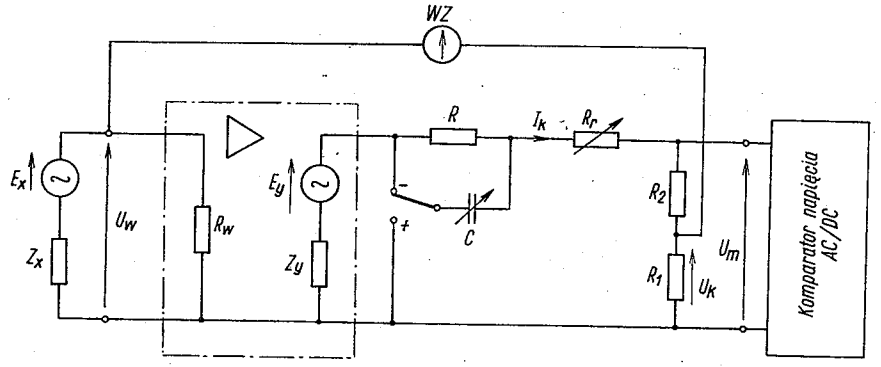
Jeżeli rezystancja wejściowa wzmacniacza R_w jest dużo większa od impedancji mierzonego źródła Z_x, to $\underline{U}_w = \underline{E}_x$, zatem wzór (1) przyjmie postać

$$\underline{E}_x = \underline{I}_k R_n. \tag{2}$$

Wartość rezystora R_n dobiera się tak, aby prąd I_k miał wartość większą od dolnego zakresu pomiarowego komparatora AC/DC.

3. KOMPENSATOR O PRZETWARZANIU NAPIĘCIE — NAPIĘCIE

W niektórych komparatorach AC/DC, np. firmy Fluke model 540 B [23], pomiar prądu jest obarczony większym błędem od błędu pomiaru napięcia. Pomiar jest dokładniejszy, jeżeli mierzone napięcie o małej wartości przetworzy się na napięcie o wartości większej niż dolny zakres pomiarowy komparatora AC/DC. Schemat ideowy kompensatora spełniającego powyższe wymagania przedstawiono na rys. 2. Prąd I_k przepływający



Rys. 2. Schemat kompensatora o przetwarzaniu napięcie — napięcie
R_1, R_2 — rezystory tworzące dzielnik napięcia; U_m — napięcie mierzone komparatorem; oznaczenia pozostałych symboli jak na rys. 1

przez układ regulacji modułu i fazy (elementy R_r, R i C) rozgałęzia się na dwa prądy: I_n — przepływający przez dzielnik napięcia złożony z rezystorów R_1 i R_2 oraz prąd I_m — pobierany przez komparator AC/DC, mierzący napięcie U_m. Napięcie kompensujące U_k określa się na podstawie wskazań komparatora AC/DC z zależności

$$\underline{U}_k = \underline{U}_m \frac{R_1}{R_1+R_2}. \tag{3}$$

W stanie równowagi układu

$$\underline{U}_w = \underline{U}_k = \underline{U}_m \frac{R_1}{R_1+R_2}. \tag{4}$$

Jeżeli rezystancja wejściowa wzmacniacza R_w jest dużo większa od impedancji mierzonego źródła Z_x, to $\underline{U}_w \approx \underline{E}_x$, zatem wzór (4) przyjmie postać

$$\underline{E}_x = \underline{U}_m \frac{R_1}{R_1+R_2}. \tag{5}$$

Przedstawiony kompensator można rozpatrywać jako wzorcowy „przekładnik podwyższający", na którego wyjściu przyłączony jest komparator. Jego przekładnię, zdefiniowaną jako stosunek napięcia E_x do napięcia na zaciskach komparatora U_m, określa wzór

$$m = \frac{\underline{E}_x}{\underline{U}_m} = \frac{R_1}{R_1+R_2}. \tag{6}$$

Do określenia przekładni kompensatora wystarcza znajomość stosunku rezystancji dzielnika. Wartość m nie zależy praktycznie od parametrów wzmacniacza ani wartości elementów regulacyjnych.

Przedstawione układy kompensacyjne są strukturalnie zbliżone do układów wzmacniaczy z ujemnym sprzężeniem zwrotnym o przetwarzaniu napięcie — prąd (rys. 1) i napięcie — napięcie (rys. 2). Zasadniczą różnicę stanowi miejsce przyłączenia wejścia wzmacniacza. W układach ze sprzężeniem zwrotnym wzmacniacz sterowany jest sygnałem niezrównoważenia, czyli — porównując z przedstawionym układem — z punktów przyłączenia wskaźnika zera.

Dla wzmacniaczy o dużych wzmocnieniach, wartość wzmocnienia określona jest nie tylko przez rezystancję pętli sprzężenia zwrotnego, ale także przez wzmocnienie wzmacniacza z otwartą pętlą sprzężenia zwrotnego oraz przez współczynniki uwzględniające wpływ impedancji mierzonego źródła i impedancji obciążenia [2]. Natomiast współczynniki przetwarzania przedstawionych układów określone są w układzie o przetwarzaniu napięcie — prąd przez rezystancję R_n, a w układzie o przetwarzaniu napięcie — napięcie przez stosunek rezystancji R_1 i R_2.

W przedstawionych układach występuje również błąd niezrównoważenia, zależny od czułości kompensatora, jednakże — jak wykazały badania — jego wartość jest tak mała, że można ją pominąć.

Składowe reaktancyjne impedancji komparatora, wzmocnienia wzmacniacza i jego impedancji wyjściowej oraz reaktancje resztkowe rezystorów powodują występowanie przesunięcia fazowego między harmonicznymi podstawowymi napięć: mierzonego i kompensacyjnego. Jeżeli w układzie stosuje się komparator AC/DC termoelektryczny, to udział składowych reaktancyjnych jest nieznaczny; jednak powinny one być skompensowane. W układzie stosuje się więc korekcję fazy dodatnią lub ujemną, złożoną z elementów RC (rys. 1 i 2).

4. ANALIZA METROLOGICZNA

4.1. Zachowanie się opisanych układów przy zmianie wartości i częstotliwości mierzonego napięcia oraz wzmocnienia wzmacniacza

Napięcie źródłowe E_y na wyjściu wzmacniacza (rys. 1 i 2) zależy od wzmocnienia napięciowego k i napięcia wejściowego U_w. Napięcie U_w zależy z kolei od wartości napięcia mierzonego E_x, wartości rezystancji wejściowej wzmacniacza R_w oraz impedancji mierzonego źródła Z_x i wynosi

$$\underline{U}_w = \underline{E}_x \frac{R_w}{R_w + \underline{Z}_x}. \tag{7}$$

Źródłowe napięcie wyjściowe wzmacniacza określa zależność

$$\underline{E}_y = \underline{U}_w \underline{k} = \underline{E}_x \underline{k} \frac{R_w}{R_w + \underline{Z}_x}. \tag{8}$$

Dla układu o przetwarzaniu napięcie — prąd przy ujemnej korekcji fazy (na rys. 1 przełącznik P w pozycji „−") słuszne jest równanie

$$\underline{E}_y = \underline{I}_k \left(\underline{Z}_y + \frac{R}{1+j\omega CR} + R_r + \underline{Z}_k + R_n \right), \tag{9}$$

w którym Z_k — impedancja wejściowa komparatora AC/DC przy pomiarze prądu.

Po przyrównaniu stronami wzorów (8) i (9) i przekształceniu otrzymuje się

$$\underline{I}_k = \frac{\underline{E}_x \underline{k}}{\left(\underline{Z}_y + \frac{R}{1+j\omega CR} + R_r + \underline{Z}_k + R_n \right)} \frac{R_w}{R_w + \underline{Z}_x}. \tag{10}$$

Podstawiając wyrażenia (7) i (10) do wzoru (1) otrzymuje się równanie stanu równowagi układu o przetwarzaniu napięcie — prąd przy ujemnej korekcji fazy

$$\underline{Z}_y + \frac{R}{1+j\omega C} + R_r + \underline{Z}_k + R_n = \underline{k} R_n. \tag{11}$$

Dla obwodu wyjściowego wzmacniacza przy dodatniej korekcji fazy (na rys. 1 przełącznik P w pozycji „+")

$$\underline{I}_k = \frac{\underline{E}_y \frac{1}{j\omega C}}{\left(\underline{Z}_y + R + \frac{1}{j\omega C} \right)(R_r + \underline{Z}_k + R_n) + \frac{1}{j\omega C}(\underline{Z}_y + R)}. \tag{12}$$

Podstawiając wyrażenia (7), (8) i (12) do wzoru (1) otrzymuje się równanie stanu równowagi układu o przetwarzaniu napięcie — prąd przy dodatniej korekcji fazy

$$\left(\underline{Z}_y + R + \frac{1}{j\omega C} \right)(R_r + \underline{Z}_k + R_n) + \frac{1}{j\omega C}(\underline{Z}_y + R) = \underline{k} \frac{R_n}{j\omega C}. \tag{13}$$

Dla układu o przetwarzaniu napięcie — napięcie przy ujemnej korekcji fazy (na rys. 2 przełącznik P w pozycji „−") równanie obwodu wyjściowego wzmacniacza ma postać

$$\underline{E}_y = \underline{I}_k \left(\underline{Z}_y + \frac{R}{1+j\omega CR} + R_r \right) + \underline{U}_m. \tag{14}$$

Prąd

$$\underline{I}_k = \frac{\underline{U}_m}{\underline{Z}_m} + \frac{\underline{U}_m}{R_1 + R_2}, \tag{15}$$

gdzie Z_m — impedancja wejściowa komparatora AC/DC przy pomiarze napięcia.

Po podstawieniu wyrażenia (15) do wzoru (14) i po przekształceniach otrzymuje się

$$\underline{U}_m = \frac{\underline{E}_y \underline{Z}_m (R_1 + R_2)}{(R_1 + R_2 + \underline{Z}_m) \left(R_r + \underline{Z}_y + \frac{R}{1+j\omega CR} \right) + \underline{Z}_m (R_1 + R_2)}. \tag{16}$$

Podstawiając wyrażenia (7), (8) i (16) do wzoru (4) otrzymuje się równanie stanu równowagi układu o przetwarzaniu napięcie — napięcie przy ujemnej korekcji fazy

$$(R_1 + R_2 + \underline{Z}_m) \left(R_r + \underline{Z}_y + \frac{R}{1+j\omega CR} \right) + \underline{Z}_m (R_1 + R_2) = \underline{k} \underline{Z}_m (R_1 + R_2). \tag{17}$$

Dla obwodu wyjściowego wzmacniacza przy dodatniej korekcji fazy (na rys. 2 przełącznik P w pozycji „+") wzór na napięcie U_m wyprowadzony metodą potencjałów węzłowych ma postać

$$\underline{U}_m = \frac{\underline{E}_y}{R_r(\underline{Z}_y+R)} \frac{1}{\left(\frac{1}{\underline{Z}_y+R}+\frac{1}{R_r}+j\omega C\right)\left(\frac{1}{R_r}+\frac{1}{R_1+R_2}+\frac{1}{\underline{Z}_m}\right)-\frac{1}{R_r^2}}. \tag{18}$$

Podstawiając wyrażenia (7), (8) i (18) do wzoru (4) otrzymuje się równanie stanu równowagi układu o przetwarzaniu napięcie — napięcie przy dodatniej korekcji fazy

$$R_r(\underline{Z}_y+R)\left[\left(\frac{1}{\underline{Z}_y+R}+\frac{1}{R_r}+j\omega C\right)\left(\frac{1}{R_r}+\frac{1}{R_1+R_2}+\frac{1}{\underline{Z}_m}\right)-\frac{1}{R_r^2}\right] = \underline{k}\frac{R_1}{R_1+R_2}. \tag{19}$$

W równaniach stanu równowagi układu (11), (13), (17), (19) nie występuje E_x. Wynika stąd, że równowaga układu nie zależy od wartości napięcia mierzonego E_x, jeżeli wzmocnienie wzmacniacza k jest stałe, a elementy wchodzące w skład obwodu kompensacyjnego są liniowe. Praktycznie tylko wzmocnienie k zależy od napięcia wejściowego. Jednakże budowane obecnie wzmacniacze mają tak dużą liniowość, że wpływ zmian wartości mierzonego napięcia na wzmocnienie jest praktycznie pomijalnie mały. Układ zachowuje więc równowagę pomimo wahań mierzonego napięcia.

W równaniach stanu równowagi występują wyrażenia zależne od częstotliwości. Jeżeli jednak udział składowej reaktancyjnej w obwodzie wyjściowym wzmacniacza jest bardzo mały, to wahania częstotliwości praktycznie nie mają wpływu na równowagę układu.

Analizując równania (11), (13), (17), (19) można stwierdzić, że zmiana wzmocnienia wzmacniacza powoduje rozrównoważenie układu. Jednakże do wykonania pomiaru wystarcza krótkotrwała stałość wzmocnienia na czas pomiaru, a zmiany długoterminowe są korygowane rezystorem dekadowym R_r. Natomiast wartość wzmocnienia nie ma wpływu na wynik pomiaru.

4.2. Czułość układu

4.2.1. Analiza układu o przetwarzaniu napięcie — prąd

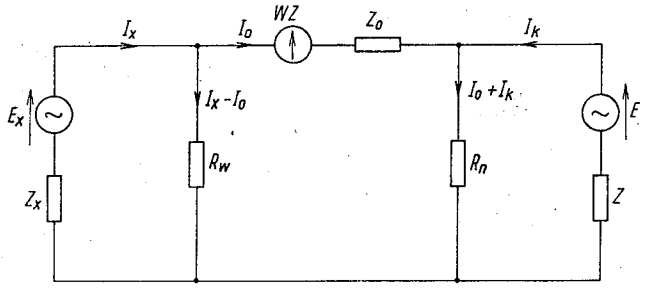
Schemat zastępczy kompensatora o przetwarzaniu napięcie — prąd przedstawiono na rys. 3. Dla układu z ujemną korekcją fazy $\underline{E} = \underline{E}_k$, a impedancja

$$\underline{Z} = \underline{Z}_y+\frac{R}{1+j\omega CR}+R_r+\underline{Z}_k. \tag{20}$$

Natomiast dla układu z dodatnią korekcją fazy

$$\underline{E} = \frac{\underline{E}_y}{1+j\omega C(R+\underline{Z}_y)}, \tag{21}$$

$$\underline{Z} = \frac{R+\underline{Z}_y}{1+j\omega C(R+\underline{Z}_y)}+R_r+\underline{Z}_k. \tag{22}$$



Rys. 3. Schemat zastępczy kompensatora o przetwarzaniu napięcie — prąd
I_o — prąd w gałęzi wskaźnika zera; Z_o — impedancja wskaźnika zera; E — zastępcze napięcie źródłowe wyjścia wzmacniacza; Z — zastępcza impedancja źródła wyjściowego napięcia wzmacniacza; pozostałe oznaczenia jak na rys. 1

Z rozwiązania równań oczkowych układu przedstawionego na rys. 1 otrzymuje się zależność na prąd w gałęzi wskaźnika zera

$$\underline{I}_o = \frac{\underline{E}_x R_w(\underline{Z}+R_n) - \underline{E} R_n(\underline{Z}_x+R_w)}{(\underline{Z}_x+R_w)(\underline{Z}_o+R_n+R_w)(\underline{Z}+R_n) - R_n^2(R_w+\underline{Z}_x) - R_w^2(\underline{Z}+R_n)} . \tag{23}$$

Po zróżniczkowaniu wyrażenia (23) i przekształceniach otrzymuje się wzór na względną czułość prądową układu kompensacyjnego

$$S_{uI} = \frac{dI_o}{\frac{d\underline{E}_x}{\underline{E}_x}} = \frac{\underline{E}_x(R_n+\underline{Z})R_w}{R_n[(\underline{Z}_x+R_w)(\underline{Z}_o+\underline{Z})+\underline{Z}_x R_w]+\underline{Z}[\underline{Z}_o\underline{Z}_x+R_w(\underline{Z}_x+\underline{Z}_o)]} . \tag{24}$$

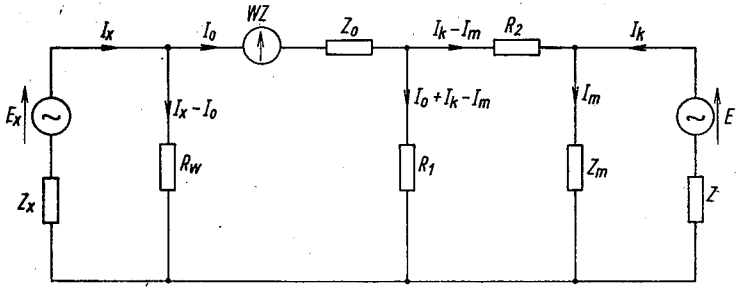
Jeżeli rezystancja wejściowa wzmacniacza jest bardzo duża w porównaniu z pozostałymi impedancjami, co ma miejsce w praktyce, to wyrażenie (24) przyjmie postać

$$S_{uI} = \frac{\underline{E}_x(R_n+\underline{Z})}{R_n(\underline{Z}_o+\underline{Z}_x+\underline{Z})+\underline{Z}(\underline{Z}_o+\underline{Z}_x)} . \tag{25}$$

Z analizy wyrażenia (25) widać, że ze wzrostem R_n czułość układu maleje. Największą czułość uzyskuje się przy małych wartościach R_n.

4.2.2. Analiza układu o przetwarzaniu napięcie — napięcie

Schemat zastępczy kompensatora o przetwarzaniu napięcie — napięcie przedstawiono na rys. 4. Dla układu przedstawionego na rys. 4 prąd w gałęzi wskaźnika zera jest określony



Rys. 4. Schemat zastępczy kompensatora o przetwarzaniu napięcie — napięcie
oznaczenia jak na rys. 2 i 3

zależnością

$$\underline{I}_o = \frac{\underline{E}_x R_w[(R_1+R_2)(\underline{Z}+\underline{Z}_m)+\underline{Z}\underline{Z}_m]-\underline{E}\underline{Z}_m R_1(\underline{Z}_x+R_w)}{[(R_1+R_2)(\underline{Z}+\underline{Z}_m)+\underline{Z}\underline{Z}_m][(\underline{Z}_x+R_w)(\underline{Z}_o+R_1)+\underline{Z}_x R_w]-R_1^2(\underline{Z}+\underline{Z}_m)(\underline{Z}_x+R_w)}. \tag{26}$$

Po zróżniczkowaniu wyrażenia (26) i przekształceniach otrzymuje się wzór na względną czułość prądową układu kompensacyjnego

$$S_{uI} = \frac{d\underline{I}_o}{\frac{d\underline{E}_x}{\underline{E}_x}} = \frac{\underline{E}_x R_w[(R_1+R_2)(\underline{Z}+\underline{Z}_m)+\underline{Z}\underline{Z}_m]}{[(R_1+R_2)(\underline{Z}+\underline{Z}_m)+\underline{Z}\underline{Z}_m][\underline{Z}_o(\underline{Z}_x+R_w)+\underline{Z}_x R_w]+ \\ +R_1(Z_x+R_w)[R_2(\underline{Z}+\underline{Z}_m)+\underline{Z}\underline{Z}_m]}. \tag{27}$$

Dla $R_w \to \infty$

$$S_{uI} = \frac{\underline{E}_x[(R_1+R_2)(\underline{Z}+Z_m)+\underline{Z}\underline{Z}_m]}{[(R_1+R_2)(\underline{Z}+\underline{Z}_m)+\underline{Z}\underline{Z}_m](\underline{Z}_o+\underline{Z}_x)+R_1[R_2(\underline{Z}+\underline{Z}_m)+\underline{Z}\underline{Z}_m]}. \tag{28}$$

Z analizy zależności (28) wynika, że układ ma największą czułość przy małych wartościach rezystancji R_1 i R_2.

4.3. Dokładność pomiaru

Dokładność pomiaru napięcia proponowaną metodą zależy od dokładności przetwarzania sygnału w układzie kompensacyjnym oraz od dokładności pomiaru I_k lub napięcia U_m komparatorem AC/DC.

Dla układu o przetwarzaniu napięcie — prąd błąd pomiaru napięcia E_x określa wyrażenie

$$\delta E_x = \delta I_k + \delta G_I, \tag{29}$$

w którym:

δI_k — błąd pomiaru prądu I_k komparatorem AC/DC,

δG_I — błąd przetworzenia układu o przetwarzaniu napięcie — prąd.

Błąd δG_I określony jest zależnością

$$\delta G_I = \delta R_n + \delta ks + \delta R_w + \delta p + \delta d, \tag{30}$$

w której:

δR_n — niedokładność rezystora kompensacyjnego,

δks — błąd spowodowany niewiernym odtwarzaniem przez wzmacniacz kształtu napięcia mierzonego w napięciu kompensacyjnym,

δR_w — błąd związany ze skończoną wartością rezystancji wejściowej wzmacniacza,

δp — błąd pobudliwości,

δd — błędy dodatkowe.

Dla układu o przetwarzaniu napięcie — napięcie

$$\delta E_x = \delta U_m + \delta G_U, \tag{31}$$

gdzie:

δU_m — błąd pomiaru napięcia U_m komparatorem AC/DC,

δG_U — błąd przetworzenia układu o przetwarzaniu napięcie — napięcie.

Błąd δG_U określa zależność

$$\delta G_U = \delta m + \delta ks + \delta R_w + \delta p + \delta d, \tag{32}$$

w której δm — błąd przekładni dzielnika rezystancyjnego złożonego z rezystorów R_1 i R_2.

4.3.1. Niedokładność rezystora R_n i przekładni m

Dokładność wywzorcowania rezystora R_n w układzie na rys. 1 i stosunku rezystorów R_1/R_2 w układzie na rys. 2 ma bezpośredni wpływ na błąd przetwarzania δG_I i δG_U. Również zmiany rezystancji pod wpływem zmian temperatury otoczenia, samonagrzewania się rezystorów od przepływającego przez nie prądu, starzenia się rezystorów, częstotliwości mogą wpływać na dokładność rezystora R_n i przekładni m.

Stosując odpowiednie materiały rezystancyjne oraz właściwą technologię można wykonać rezystory na poziomie niedokładności 10^{-5} i co najmniej z taką samą dokładnością wykonać dzielnik rezystancyjny. Istnieje również możliwość zastąpienia dzielnika rezystancyjnego dzielnikiem indukcyjnym.

4.3.2. Niewierne odtworzenie przez wzmacniacz kształtu napięcia mierzonego w napięciu kompensacyjnym

Mierzone napięcie można przedstawić w postaci szeregu Fouriera

$$e_x(t) = \sum_{n=1}^{n=p} U_{mn} \cos(n\omega_1 t + \varphi_n), \tag{33}$$

w którym:

U_{mn} — amplituda n-tej harmonicznej,
ω_1 — pulsacja pierwszej harmonicznej,
t — czas,
φ_n — przesunięcie fazowe n-tej harmonicznej,
n — liczba naturalna.

Napięcie kompensacyjne ma kształt odwzorowany kształtem napięcia mierzonego, jednak to odwzorowanie obarczone jest pewnymi błędami amplitudowymi ΔU_{mn} i fazowymi $\Delta\varphi_n$. Jego przebieg czasowy określa wyrażenie

$$u_k(t) = \sum_{n=1}^{n=p} (U_{mn} + \Delta U_{mn}) \cos(n\omega_1 t + \varphi_n + \Delta\varphi_n). \tag{34}$$

Wartości skuteczne napięcia mierzonego i kompensacyjnego wynoszą

$$E_x = \sqrt{\sum_{n=1}^{n=p} U_n^2} = U_1 \sqrt{1 + \sum_{n=2}^{n=p} \left(\frac{U_n}{U_1}\right)^2}, \tag{35}$$

$$U_k = \sqrt{\sum_{n=1}^{n=p} (U_n + \Delta U_n)^2} = (U_1 + \Delta U_1) \sqrt{1 + \sum_{u=2}^{n=p} \frac{U_n + \Delta U_n}{U_1 + \Delta U_1}}, \tag{36}$$

przy czym:

U_n — wartość skuteczna n-tej harmonicznej,

ΔU_n — bezwzględny błąd odtworzenia wartości skutecznej n-tej harmonicznej,

U_1 — wartość skuteczna pierwszej harmonicznej,

ΔU_1 — bezwzględny błąd odtworzenia wartości skutecznej pierwszej harmonicznej.

Przy zastosowaniu selektywnego wskaźnika zera czułego na pierwszą harmoniczną błąd odtworzenia tej harmonicznej zależy jedynie od czułości wskaźnika zera i może być pominięty. Wówczas

$$U_k = U_1 \sqrt{1+\sum_{n=2}^{n=p}\left(\frac{U_n+\Delta U_n}{U_1}\right)^2}. \tag{37}$$

Pozostałe harmoniczne nie są w pełni skompensowane, gdyż mierzone napięcie i napięcie kompensacyjne różnią się nieznacznie kształtami. Ta różnica kształtów spowodowana jest dodatkowymi zniekształceniami napięcia kompensacyjnego wprowadzonymi przez wzmacniacz i elementy reaktancyjne w obwodzie wyjściowym wzmacniacza.

Błąd pomiaru napięcia E_x, powstały na skutek niewiernego odtworzenia kształtu krzywej napięcia E_x, definiuje wzór

$$\delta ks = \frac{U_k - E_x}{E_x} = \frac{U_k}{E_x} - 1, \tag{38}$$

Po podstawieniu wyrażeń (35) i (37) do wzoru (38) otrzymuje się

$$\delta ks = \sqrt{\frac{1+\sum_{n=2}^{n=p}\left(\frac{U_n+\Delta U_n}{U_1}\right)^2}{1+\sum_{n=2}^{n=p}\left(\frac{U_n}{U_1}\right)^2}} - 1. \tag{39}$$

Po wprowadzeniu oznaczeń

$$\alpha = \sum_{n=2}^{n=p}\left(\frac{U_n+\Delta U_n}{U_1}\right)^2,$$

$$\beta = \sum_{n=2}^{n=p}\left(\frac{U_n}{U_1}\right)^2,$$

wzór (39) przyjmie postać

$$\delta ks = \sqrt{\frac{1+\alpha}{1+\beta}} - 1. \tag{40}$$

Jeżeli $\alpha \ll 1$ i $\beta \ll 1$, wzór (40) można zapisać w postaci

$$\delta ks = \frac{\alpha - \beta}{2}. \tag{41}$$

Podstawiając do wzoru (41) wyrażenia za α i β otrzymuje się

$$\delta ks = \frac{1}{U_1^2} \sum_{n=2}^{n=p} U_n^2 \frac{\Delta U_n}{U_n} + \frac{1}{2} \frac{1}{U_1^2} \sum_{n=2}^{n=p} \Delta U_n^2. \tag{42}$$

We wzorze (42) $\frac{\Delta U_m}{U_n}$ oznacza względny błąd modułowy odtworzenia n-tej harmonicznej, gdy występuje kompensacja pierwszej harmonicznej.

Wzmacniacz można rozpatrywać jako źródło harmonicznych dodających się do harmonicznych mierzonego napięcia. Gdyby przesunięcia fazowe były równe zeru, odpowiednie harmoniczne dodawałyby się algebraicznie. Wówczas błąd bezwzględny odtworzenia n-tej harmonicznej byłby równy n-tej harmonicznej wnoszonej przez wzmacniacz i elementy reaktancyjne w obwodzie wyjściowym wzmacniacza. Stąd

$$\frac{\sum_{n=2}^{n=p} \Delta U_n^2}{U_1^2} \leqslant w^2, \tag{43}$$

przy czym w — współczynnik zniekształceń nieliniowych, wnoszonych przez wzmacniacz i elementy reaktancyjne w jego obwodzie wyjściowym.

Nachylenie wzmocnienia prądowego wzmacniacza w funkcji pulsacji określa wyrażenie

$$k_i(\omega) = \frac{I_k(\omega)}{U_w(\omega)}, \tag{44}$$

w którym:

$I_k(\omega)$ — prąd wyjściowy wzmacniacza w funkcji pulsacji,

$U_w(\omega)$ — napięcie wejściowe wzmacniacza w funkcji pulsacji.

Bezwzględną nieliniowość nachylenia wzmocnienia prądowego określa wzór

$$\Delta k_i(\omega) = \frac{\Delta I_k(\omega)}{U_w(\omega)}, \tag{45}$$

a względną nieliniowość wyraża zależność

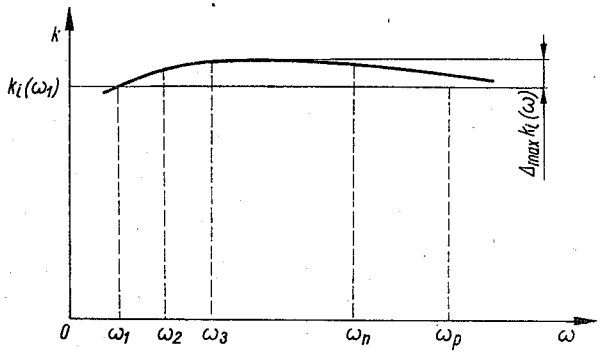
$$\delta k_i(\omega) = \frac{\Delta k_i(\omega)}{k_i(\omega)} = \frac{\Delta I_k(\omega)}{I_k(\omega)}. \tag{46}$$

Zależność nachylenia wzmocnienia prądowego wzmacniacza w funkcji pulsacji przedstawia wykres na rys. 5.

Prąd wyjściowy wzmacniacza jest proporcjonalny do napięcia kompensacyjnego U_k, zatem

$$\delta k_i(\omega) = \frac{\Delta U_k(\omega)}{U_k(\omega)} = \frac{\Delta U_n}{U_n}. \tag{47}$$

Każda z harmonicznych jest wzmacniana z innym błędem. Jednak dla uproszczenia założono graniczny przypadek, że wszystkie harmoniczne są przenoszone z jednakowym



Rys. 5. Zależność nachylenia wzmocnienia prądowego wzmacniacza w funkcji pulsacji

błędem maksymalnym

$$\delta_{max} k_i(\omega) = \frac{\Delta_{max} k_i(\omega)}{k_i(\omega)} = \frac{\Delta_{max} U_n}{U_n}. \tag{48}$$

Podstawiając we wzorze (42)

$$\frac{1}{U_1^2} \sum_{n=2}^{n=p} U_n^2 = h^2, \tag{49}$$

gdzie h — współczynnik zawartości wyższych harmonicznych w mierzonym napięciu, oraz uwzględniając zależności (48) i (49) otrzymuje się ostatecznie zależność

$$\delta ks \leqslant h^2 \delta_{max} k_i(\omega) + \frac{1}{2} w^2. \tag{50}$$

Wzór (50) określa wpływ niewierności odtworzenia przez wzmacniacz kształtu napięcia mierzonego w napięciu kompensacyjnym na błąd pomiaru wartości skutecznej napięcia E_x. Wzór ten jest taki sam jak wzór wyprowadzony przez autora w pracy [15], określający wpływ niewierności odtworzenia przez wzmacniacz kształtu mierzonego prądu w prądzie kompensacyjnym w kompensatorze prądowym.

4.3.3. Wpływ rezystancji wejściowej wzmacniacza

Rezystancja wejściowa wzmacniacza wraz z impedancją mierzonego źródła tworzy dzielnik napięcia, na skutek czego napięcie wejściowe wzmacniacza jest zgodnie z zależnością (7) pomniejszone o spadek napięcia na impedancji mierzonego źródła. Ponieważ kompensowane jest napięcie U_w a nie E_x, wynik pomiaru obarczony jest błędem

$$\delta R_w = \frac{R_w}{R_w + Z_x} - 1. \tag{51}$$

Ze wzoru (51) wynika, że błąd δR_w ma pomijalnie małą wartość, jeżeli $R_w \gg Z_x$, a odgrywa rolę przy porównywalnych wartościach R_w i Z_x. Dlatego w układzie należy stosować wzmacniacz o bardzo dużej rezystancji wejściowej. Stosując w obwodzie wej-

ściowym wzmacniacza układ scalony z wejściem „FET-owym", którego rezystancja wejściowa jest rzędu $10^{12}\Omega$, błąd δR_w jest pomijalnie mały, na poziomie 10^{-5} w zakresie impedancji mierzonego źródła Z_x do $10^7\Omega$.

4.3.4. Składowa stała w prądzie wyjściowym wzmacniacza

Komparatorem AC/DC reagującym na wartość skuteczną mierzy się prąd lub napięcie łącznie z jego składową stałą. Występowanie składowej stałej w prądzie wyjściowym wzmacniacza, gdy nie występuje składowa stała w napięciu mierzonym, powoduje błąd pomiaru.

W dobrze skonstruowanym wzmacniaczu składowa stała może mieć pomijalnie mały wpływ na dokładność pomiaru.

4.3.5. Inne źródła błędów

Źródłami dodatkowych błędów są sprzężenia pasożytnicze, zakłócenia, szumy i prądy upływnościowe. W układzie należy starannie ekranować poszczególne elementy układu i stosować odpowiednie materiały izolacyjne oraz w obwodzie wejściowym wzmacniacza — elementy niskoszumowe. Szumy i zakłócenia ograniczają dolny zakres pomiarowy. Dla napięć większych od 0,25 mV, przy dobrze skonstruowanym układzie, błędy dodatkowe mogą mieć pomijalnie małą wartość.

5. PARAMETRY I WYNIKI BADAŃ MODELU

W Instytucie Metrologii Elektrycznej skonstruowano pod kierunkiem autora model kompensatora stanowiącego połączenie układów pomiarowych z rys. 1 i rys. 2. Przeprowadzono badania wszystkich czynników wpływających na dokładność pomiaru. Model ten ma następujące ważniejsze parametry techniczne:

— zakres mierzonego napięcia 0,25 ÷ 250 mV,
w trzech podzakresach 2,5 mV, 25 mV, 250 mV,
— zakres częstotliwości 40 ÷ 1000 Hz,
— wartość prądu mierzonego komparatorem AC/DC w układzie o przetwarzaniu napięcie — prąd 10 ÷ 100 mA,
— wartość napięcia mierzonego komparatorem AC/DC w układzie o przetwarzaniu napięcie — napięcie 0,25 ÷ 2,5 V,
— niedokładność przetwarzania układu 0,01%,
— zasilanie układu 220 V; 50 Hz.

Rezystory kompensacyjne i dzielnik napięcia wykonano z zeraninu o współczynniku temperaturowym zmian rezystancji $3 \cdot 10^{-6}$/K.

Wzmacniacz zbudowano z trzech stopni. Pierwszym stopniem wejściowym jest wzmacniacz operacyjny typu 3521L z wejściem „FET-owym" firmy Burr-Brown. Jako drugi stopień wzmacniający zastosowano niskoszumowy wzmacniacz operacyjny μA 725.

Stopień wyjściowy stanowi wzmacniacz o wzmocnieniu napięciowym równym jedności, wykonany z elementów dyskretnych z tranzystorami wyjściowymi BD 254 i BD 255. Całkowite wzmocnienie wzmacniacza wynosi 2500 dla zakresu 2,5 mV, 250 dla 25 mV i 25 dla 250 mV. Zmiany wzmocnienia dokonuje się równocześnie ze zmianą zakresu pomiarowego za pomocą dzielnika włączonego między pierwszym a drugim stopniem wzmocnienia.

Do równoważenia układu stosowano jako wskaźnik zera selektywny nanowoltomierz typu 233 firmy Unipan. Ponieważ wskaźnik ten ma jeden potencjał na obudowie, włączono go do układu za pośrednictwem przedwzmacniacza o symetrycznym wejściu. Przedwzmacniacz ten o wzmocnieniu 10 zbudowano z dwóch wejściowych niskoszumowych wzmacniaczy operacyjnych μA 725 i wyjściowego μA 710.

Tablica 1

Wyniki badań wykonanego kompensatora

Lp.	Przyczyna błędu	Wartość błędu	Uwagi
1	Niedokładność rezystora R_n	$\leqslant 3 \cdot 10^{-5}$	wyznaczono przy $f = 120$ Hz
2	Niedokładność przekładni dzielnika m	$\leqslant 2 \cdot 10^{-5}$	wyznaczono przy $f = 120$ Hz
3	Stała czasowa rezystorów	$\leqslant 3 \cdot 10^{-6}$	do $f = 1000$ Hz
4	Niewierne odtworzenie w napięciu kompensacyjnym za pomocą wzmacniacza kształtu napięcia mierzonego	$\leqslant 3 \cdot 10^{-6}$	przy $h = 0{,}3\%$ i $f = 1000$ Hz
		$\leqslant 4 \cdot 10^{-5}$	przy $h = 10\%$ i $f = 1000$ Hz
5	Składowa stała wnoszona przez wzmacniacz	$\leqslant 4 \cdot 10^{-6}$	przy $E_x = 0{,}25$ mV
6	Szumy i zakłócenia	$\leqslant 6 \cdot 10^{-6}$	przy $E_x = 0{,}25$ mV
7	Nieczułość kompensatora	$\leqslant 1 \cdot 10^{-5}$	przy wskaźniku zera o zakresie 1 μV na pełne wychylenie

Ważniejsze końcowe wyniki badań zestawiono w tablicy 1. Z tablicy tej wynika, że zasadniczy wpływ na błąd przetwarzania układu ma błąd rezystora kompensacyjnego w układzie o przetwarzaniu napięcie — prąd i niedokładność przekładni dzielnika w układzie o przetwarzaniu napięcie — napięcie. Ponadto, przy wyższych częstotliwościach i znacznych odkształceniach mierzonego napięcia, uwydatnia się znacznie błąd spowodowany dodatkowymi odkształceniami napięcia kompensacyjnego, wprowadzonymi przez wzmacniacz i elementy reaktancyjne w obwodzie wyjściowym wzmacniacza. Błąd ten rośnie ze wzrostem współczynnika zawartości wyższych harmonicznych w mierzonym napięciu.

W najbardziej niekorzystnych warunkach, tj. dla $E_x = 0{,}25$ mV przy częstotliwości harmonicznej podstawowej 1000 Hz i współczynniku zawartości wyższych harmonicznych do 10%, błąd przetwarzania układu nie przekracza 0,01% dla układu o przetwarzaniu napięcie — prąd i 0,009% dla układu o przetwarzaniu napięcie — napięcie.

6. WNIOSKI

Przedstawione układy pomiarowe pełnią rolę bardzo dokładnych przetworników wartości lub wielkości i mogą służyć do pomiarów małych napięć przy współpracy z zewnętrznym komparatorem AC/DC. Odwzorowanie w napięciu kompensacyjnym za pomocą wzmacniacza, kształtu napięcia mierzonego umożliwia kompensację praktycznie wszystkich harmonicznych. Metoda ta pozwala na precyzyjne pomiary napięcia odkształconego, praktycznie bez poboru energii z mierzonego źródła, w zakresie częstotliwości uwarunkowanym głównie pasmem przenoszenia wzmacniacza i stałą czasową rezystorów kompensacyjnych.

Dokładność pomiaru napięcia zaproponowanymi układami kompensacyjnymi zależy od dokładności przetwarzania układu i dokładności komparatora AC/DC. W zakresie małych częstotliwości błąd przetwarzania może być mniejszy od błędu komparatora AC/DC lub z nim porównywalny. Błąd przetwarzania wykonanego modelu kompensatora nie przekracza 0,01% dla prądu o współczynniku zawartości wyższych harmonicznych poniżej 10%.

LITERATURA

1. R.F. Aknaer, O.P. Gałachowa, T.B. Roždestvenskaja, *Metody i sredstva obesnečenija edinstva izmerenij napraženija peremennogo toka*, Izmeritelnaja technika, 1976, nr 4, s. 66—71.
2. F. Arnold, *Gemeinsame Signalflusstheorie des Übertragungs- und Fehlerverhaltens der Messverstärker—Grundtypen*, ATM, Z 631—8/9, 1968, nr 1/2.
3. G. Frizte, H. Klausche, *Wechselstromverstärker für Präzisionsmessungen*, Siemens — Zeitschrift, 1955, nr 9, z. 10.
4. R. Gotszalk, Z. Nawrocki, *Przekładniki elektroniczne i urządzenia do ich wzorcowania*, Pomiary Automatyka Kontrola, 1974, nr 12.
5. H. Helke, *Wechselstrom-Gleichstrom-Konverter und Kompensationsverfahren für Wechselstrom*, ATM, J. 942 — 2/3, 1967, nr 10/11, J 942—4, 1968, nr 1.
6. F.L. Hermach, *Thermal Converters as AC — DC Transfer for Current and Voltage Measurements at Audio Frequencies*, National Bureau of Standars, Vol. 48, 1952, nr 2.
7. F.L. Hermach, E.S. Williams, *Thermal Converters for Audio-Frequency Voltage Measurements of High Accuracy*, IEEE Trans. Instrum. Meas., Vol. IM-15, December 1966, nr 4, s. 260—268.
8. S. Iwamato, H. Hirayoma, *AC/DC Thermal Converters of the ETL*, IEEE Trans. Instrum. Meas., Vol. IM-23, December 1974, nr 4, s. 326—329.
9. A. Jellonek, Z. Karkowski, *Miernictwo radiotechniczne*, WNT, Warszawa 1972, s. 339—344.
10. W. Kaiser, *Elektrische Präzisions-Wechselstrom-Messverstärker*, Elektronik, 1957, nr 7, z. 11.
11. H. Kalusche, *Präzisions-Wechselstromverstärker*, ATM, Z 631—7, 1964, nr 3.
12. I.F. Klistorin, *Cyfrowe voltometry dejstvujuščich značenij*, Avtometrija, 1962, nr 2.
13. M. Lisowski, K. Morawski, *Precyzyjny kompensator prądu przemiennego o częstotliwości 50 Hz*, Prace Naukowe Inst. Metrol. Elektr. Politechniki Wrocławskiej, 1971, nr 3.
14. M. Lisowski, *Układ zasilania precyzyjnego kompensatora napięcia przemiennego 50 Hz*, Prace Naukowe Inst. Metrol. Elektr. Politechniki Wrocławskiej, 1976, nr 12.
15. M. Lisowski, *Metoda kompensacji prądowej do precyzyjnego pomiaru małych wartości skutecznych prądu przemiennego małej częstotliwości*, Rozprawy Elektrotechniczne, 1977, z. 4, s. 941—956.
16. M. Lisowski, *Metoda rozszerzania dolnego zakresu pomiarowego precyzyjnych komparatorów*, Materiały XIV Międzyuczelnianej Konferencji Metrologów, Częstochowa, 1979, s. 142—148.

17. K. Morawski, *Dokładność pomiaru kompensatorami prądu zmiennego w układzie o współrzędnych prostokątnych i biegunowych przy częstotliwości 50 Hz*, Zeszyty Naukowe Politechniki Wrocławskiej, Elektryka XXXIII, 1968, nr 185.
18. T.B. Roždestvenskaja, R.F. Aknaer, O.P. Gałachova, *Gosudarstvennyj special-nyj ětalon edinicy napražsnija peremennogo toka v diapazone častot 20 — 3·10⁷ Hz*, Izmeritelnaja technika, 1976, nr 3, s. 44—46.
19. G. Schlamp, *Der komplexe Wechselstromkompensator und seine Anwendung*, ETZ, A, 1960, nr 22, s. 784—789.
20. Cz. Szczepaniak, *Komparatory elektryczne pomiarowe*, WNT, Warszawa 1979.
21. Cz. Szczepaniak, *Kompensatory prądu przemiennego*, WNT, Warszawa 1976.
22. M.S. Veksler, M.V. Popov, *Točnye izmerenija veličin peremennogo toka s iskažennoj formoj krivoj v širokom diapazone častot*, Sb. nauč. tr. Ivanov, energ. in-ta 1972, vyp. 23, s. 74—77.
23. *540 B Thermal Transfer Standard*, Fluke Catalog 1980/81, s. 105—106.

M. LISOWSKI

COMPENSATING METHOD ALLOWING PRECISE MEASUREMENTS OF ROOT MEAN SQUARE L. F. DEFORMED TENSION VALUES

Summary

In this paper a method of compensation of the deformed a.c. tension is described. The identity of the shapes of the deformed — and the compensating tensions is obtained by image-conversion in a measured tension shape-forming amplifier. This method is analogous to the compensating method allowing a measurement of the deformed l.f. current values described by the author in [15]. A metrological analysis of two compensating circuits: converting tension into current and tension into tension is described. The analysis deals with the influence of value and frequency change of the measured tension on the measuring accuracy, sensitivity and the balance of the circuit. Metrological properties of such a potentiometer are given. If the potentiometer is coupled to an a.c./d.c. current comparator in a tension-current conversion circuit or to an a.c./d.c. tension comparator in a tension-tension conversion circuit it would allow precise measurements of a.c. tensions in the range between 0.25 mV÷250 mV at frequencies between 40÷1000 Hz. The conversion coeficient is established with an error of 0.01%. The potentiometer makes it possible to enlarge the tension range of the a.c./d.c. comparator one thousand times downward or to measure small tension values by means of an a.c./d.c. comparator working at current ranges.

M. LISOWSKI

MÉTHODE COMPENSATOIRE PERMETTANT DE MESURER AVEC PRÉCISION DE FAIBLES VALEURS EFFICACES DE LA TENSION ALTERNATIVE DÉFORMÉE À BASSE FRÉQUENCE

Résumé

L'article présente une méthode de compensation de la tension alternative déformée en faisant la proection, à l'aide d'un amplificateur, de la forme de tension mesurée dans la tension compensatoire. Cette méthode, analogiquement à la méthode de compensation du courant alternatif déformé, présentée par l'auteur dans article [15], permet de mesurer avec précision les valeurs efficaces, sans avoir recours à la prise de courant de la source mesurée. On a procédé à l'analyse métrologique de deux systèmes compensatoires: transformant la tension en courant et la tension en tension. On a étudié l'influence des variations de l'intensité et des fréquences de la tension mesurée sur l'équilibre du système et sur la sensibilité et la précision de mesure. On présente les résultats et les paramètres obtenus au cours de l'étude et concernant le potentiomètre construit.

A l'aide d'un comparateur du courant AC/DC, transformant la tension en courant, ou d'un compa-

rateur de la tension AC/DC, transformant la tension en tension, le potentiomètre construit permet de mesurer avec une grande précision la tension dans les limites de 0,25 mV à 250 mV, à fréquence de 40 à 1000 Hz. Le facteur de conversion des systèmes a été défini avec inexactitude de 0,01%. Le potentiomètre construit permet d'élargir mille fois vers le bas le domaine des mesures du comparateur AC/DC ou les mesures de basses tensions à l'aide de comparateur AC/DC, en utilisant le domaine du courant.

M. LISOWSKI

KOMPENSATIONSMETHODE FÜR FEINMESSUNG KLEINER EFFEKTIVWERTE DER VERUNSTALTETEN NIEDERFREQUENTEN WECHSELSPANNUNG

Zusammenfassung

Im Artikel wird eine Methode der Kompensation der verunstalteten Wechselspannung beschrieben. Die Form dieser Spannung wird bei der Kompensationsspannung mittels eines Verstärkers abgebildet. Diese Methode ist der vom Verfasser in [15] beschriebenen Methode der Wechselstromkompensation analog. Sie ermöglicht die Feinmessung des Effektivwertes der Wechselspannung ohne Energieentnahme aus der gemessenen Quelle.

Es wurde eine metrologische Analyse zweier Kompensationsschaltungen mit Umwandlung: Spannung-Strom und Spannung-Spannung durchgeführt. Der Artikel behandelt den Einfluß der Wert- und Frequenzänderungen der gemessenen Spannungen auf das Verhalten der Anordnung sowie auf ihre Empfindlichkeit und Meßgenauigkeit. Eigenschaften des ausgeführten Kompensators sowie Prüfergebnisse werden gebracht.

Dieser Kompensator ermöglicht in Zusammenarbeit mit einem AC/DC-Stromkomparator in der Umwandlungsschaltung Spannung-Strom bzw. mit einem AC/DC-Spannungskomparator in der Umwandlungsschaltung Spannung-Spannung sehr genaue Spannungsmessungen von 0,25 mV bis 250 mV im Frequenzbereich von 40 bis 1000 Hz. Der Umwandlungskoeffizient der Schaltungen ist mit einer Ungenauigkeit von 0,01% behaftet. Der erwogene Kompensator bietet die Möglichkeit, die untere Grenze des Spannungsbereiches des AC/DC-Komparators um das 1000-fache hinunterzusetzen, auch Messungen von Kleinspannungen beim Einsatz des AC/DC-Komparators in den Strombereich durchzuführen.

М. ЛИСОВСКИ

КОМПЕНСАЦИОННЫЙ МЕТОД ТОЧНОГО ИЗМЕРЕНИЯ МАЛЫХ ЭФФЕКТИВНЫХ ЗНАЧЕНИЙ ПЕРЕМЕННОГО НАПРЯЖЕНИЯ НИЗКОЙ ЧАСТОТЫ

Резюме

Приведен метод компенсации переменного напряжения путем воспроизведения в компенсационном напряжении — с помощью усилителя — формы кривой измеряемого напряжения. Этот метод — аналогичный методу компенсации переменного тока с искажённой формой кривой, описанному автором в труде [15] — обеспечивает точные измерения эффективного значения напряжения без потребления энергии от измеряемого источника. Приведен метрологический анализ двух компенсаторов: с преобразованием напряжение—ток и напряжение-напряжение. Рассмотрено влияние изменений значения и частоты измеряемого напряжения на равновесие компенсатора и влияние их на чувствительность и погрешности измерения. Приведены основные параметры и результаты испытания для построенного компенсатора. Построенный компенсатор работающий с компаратором тока AC/DC в схеме с преобразованием напряжение-ток или компаратором напряжения AC/DC в схеме с преобразованием напряжение-напряжение, обеспечивает точные измерения напряжения $0{,}25 \div 250$ мв частотой $40 \div 1000$ Гц. Коэффициент преобразования компенсатора определен с погрешностью 0,01%. Представленный компенсатор изменяет нижний предел применения компараторов AC/DC в 10^3 раз или измерения малых напряжений при работе компаратора AC/DC в токовом режиме.

ROZPRAWY ELEKTROTECHNICZNE 1982, 28, z. 3-4, ss. 385—400

621.37

Zastosowanie metody Kantorowicza do rozwiązywania równania przewodnictwa

JAN PURCZYŃSKI (SZCZECIN), ANDRZEJ BRYKALSKI (SZCZECIN)

Ośrodek Elektronicznej Techniki Obliczeniowej Politechniki Szczecińskiej

Otrzymano 15.4.1980

W pracy zastosowano metodę Kantorowicza do rozwiązania kilku problemów związanych z przewodzeniem ciepła w stanie nieustalonym. Wyznaczono rozkład temperatury w wiodącym prąd przewodzie o przekroju w kształcie prostokąta. Uwzględniono pierwsze i trzecie zagadnienie brzegowe, uzyskując proste wzory określające rozkład temperatury. Ponadto wyznaczony został rozkład pola magnetycznego w przewodzącym pierścieniu obejmującym przewód z prądem i rdzeniem umieszczonym wewnątrz cewki.

1. WPROWADZENIE

Analiza pól elektromagnetycznych oraz procesów przewodzenia i wymiany ciepła prowadzi do poszukiwania rozwiązania równania różniczkowego cząstkowego z odpowiednimi warunkami granicznymi. W przypadku analizy pól stacjonarnych są to równania Laplace'a lub Poissona, a więc równania typu eliptycznego. Procesy nieustalone opisane są równaniami typu parabolicznego lub eliptyczno-parabolicznego. Dokładne rozwiązanie tych równań jest możliwe do uzyskania w stosunkowo niewielu przypadkach, w takich, w których brzeg rozpatrywanego obszaru pozwala na zastosowanie jednego z układów współrzędnych.

Większe możliwości zapewniają metody bezpośrednie rachunku wariacyjnego. Dzięki doborowi odpowiednich funkcji przybliżających, uwzględnić można obszary o brzegu bardziej złożonym, przy czym obliczenia prowadzą do uzyskania prostych i łatwych do interpretacji zależności. Wykorzystanie tych metod do rozwiązywania równania przewodnictwa stało się w ostatnich latach problemem aktualnym, o czym świadczą ukazujące się prace [4, 5, 6, 7, 9, 10, 11, 12].

Zadanie poszukiwania funkcji $U(u, v, t)$ spełniającej równanie przewodnictwa

$$\frac{\partial}{\partial u}\left[\lambda(u, v, t)\frac{\partial U}{\partial u}\right]+\frac{\partial}{\partial v}\left[\lambda(u, v, t)\frac{\partial U}{\partial v}\right]-C(u, v)\frac{\partial U}{\partial t}=-Q(u, v, t) \tag{1.1}$$

oraz warunki na brzegu $\Gamma=\Gamma_1\cup\Gamma_2\cup\Gamma_3$ obszaru płaskiego D

$$[U-U_1]_{\Gamma_1}=0, \tag{1.2}$$

$$\left[\lambda \frac{\partial U}{\partial n}+\sigma\right]_{\Gamma_2}=0, \tag{1.3}$$

$$\left[\lambda \frac{\partial U}{\partial n}+\alpha(U-U_3)\right]_{\Gamma_3}=0 \tag{1.4}$$

i warunek początkowy

$$U(u, v, 0)=U_o(u, v) \tag{1.5}$$

zastępuje się zagadnieniem wariacyjnym polegającym na minimalizacji następującego funkcjonału

$$F[U]=\int_D\int\left[\lambda\left(\frac{\partial U}{\partial u}\right)^2+\lambda\left(\frac{\partial U}{\partial v}\right)^2+2\left(CU\frac{\partial U}{\partial t}-QU\right)\right]du\,dv+$$

$$+2\int_{\Gamma_2}\sigma U ds+2\int_{\Gamma_3}\alpha U\left(\frac{U}{2}-U_3\right)ds. \tag{1.6}$$

Równanie (1.1) wraz z podanymi warunkami granicznymi pojawia się przy opisie różnych pól niestacjonarnych. Występującym funkcjom C, α, λ, σ nie będziemy przypisywali żadnej interpretacji fizycznej, poprzestając na stwierdzeniu, że są to parametry (funkcje zależne od parametrów) charakteryzujące własności fizyczne obszaru, a funkcja Q jest wymuszeniem.

Sens fizyczny tych funkcji wyniknie z rozpatrywanych przykładów w dalszej części pracy. Zgodnie z zasadami przyjętymi w pracy [16] zmienna t jest traktowana jako parametr, a minimum funkcjonału (1.6) występuje dla każdej ustalonej chwili t.

Dla minimalizacji funkcjonału (1.6) w pracach [6, 7, 9, 10] użyto „zmodyfikowanej metody Kantorowicza". Metoda ta jest pewną odmianą metody Ritza, którą podobne zagadnienie rozwiązano w pracy [11]. Zmodyfikowana metoda Kantorowicza prowadzi do równań różniczkowych zwyczajnych ze względu na zależne od czasu współczynniki, podczas gdy w „klasycznej" metodzie Kantorowicza [2] funkcje wyznaczane z równań różniczkowych są funkcjami zmiennej, względem której w minimalizowanym funkcjonale odbywa się całkowanie. W pracy [4] równanie przewodnictwa rozwiązano minimalizując funkcjonał powstały w oparciu o metodę momentów, a w [12] zastosowana została metoda Bubnowa-Galerkina. Niektórzy autorzy [5, 13], przez zastosowanie do równania przewodnictwa przekształcenie Laplace'a, zmieniają typ równania na eliptyczny, po czym do odpowiedniej całki wariacyjnej stosują jedną z metod bezpośrednich.

W niniejszej pracy przedstawiono przykłady ilustrujące możliwość zastosowania klasycznej metody Kantorowicza do rozwiązania zagadnień opisywanych równaniem przewodnictwa.

2. STAN NIEUSTALONY POLA CIEPLNEGO W PROSTOKĄTNYM PRZEWODZIE WIODĄCYM PRĄD

Nieskończenie długi, prostokątny przewodnik elektryczny ($-a \leqslant x \leqslant a$, $-b \leqslant y \leqslant b$), począwszy od chwili $t = 0$, przewodzi stały prąd. Dla czasu $t = 0$ każdy punkt przewod-

nika jest źródłem ciepła o stałej wydajności $g = J^2\varrho$, gdzie: J — gęstość prądu, ϱ — rezystywność.

Przyjmujemy, że materiał przewodnika jest jednorodny i izotropowy oraz że wszystkie parametry charakteryzujące ten materiał są stałe. W chwili $t = 0$ temperatura wszystkich punktów jest równa temperaturze otoczenia i wynosi zero.

Należy wyznaczyć rozkład temperatury w przewodniku dla przypadków:

a) powierzchnia przewodnika utrzymywana jest w stałej temperaturze,

b) przewodnik oddaje ciepło zgodnie z prawem Newtona ze stałym współczynnikiem h.

Funkcja $T(x, y, t)$ opisująca pole temperaturą spełnia zgodnie z [2, 3] równanie przewodnictwa

$$\nabla^2 T - c\frac{\partial T}{\partial t} = -q, \tag{2.1}$$

gdzie: $c = \frac{c_w \gamma}{\lambda}$, $q = \frac{g}{\lambda}$, przy czym λ jest przewodnością cieplną właściwą, c_w ciepłem właściwym, a γ gęstością materiału. Funkcja $T(x, y, t)$ spełnia ponadto warunek początkowy

$$T(x, y, 0) = 0 \tag{2.2}$$

oraz warunki brzegowe: pierwszego rodzaju — dla przypadku a), i drugiego rodzaju — dla przypadku b).

2.1. Pierwsze zagadnienie brzegowe

Założenie, że powierzchnia przewodnika jest utrzymywana w stałej temperaturze wymaga rozwiązania równania (2.1) z warunkiem początkowym (2.2) i brzegowym:

$$T(-a, y, t) = T(a, y, t) = 0 \quad \text{dla} \quad -b \leqslant y \leqslant b, \tag{2.3}$$

$$T(x, -b, t) = T(x, b, t) = 0 \quad \text{dla} \quad -a \leqslant x \leqslant a. \tag{2.4}$$

Funkcjonał (1.6) przy uwzględnieniu warunków (2.3) i (2.4) przyjmuje postać

$$F[T] = \int_D\int \left[\operatorname{grad}^2 T + 2T\left(c\frac{\partial T}{\partial t} - q\right)\right] ds. \tag{2.5}$$

Przybliżone rozwiązanie równania (2.1) przewidujemy w takiej postaci, aby spełniony został jeden z warunków brzegowych, np. (2.4). Warunek ten spełnia funkcja

$$T(x, y, t) = f(x, t)(b^2 - y^2). \tag{2.6}$$

Funkcję $f(x, t)$ wyznacza się z warunku istnienia minimum funkcjonału (2.5), żądając jednocześnie, by znikała ona dla $x = \pm a$ oraz spełniała warunek początkowy $f(x, 0) = 0$. Po podstawieniu założonego rozwiązania (2.6) do funkcjonału (2.5) i scałkowaniu względem zmiennej y otrzymujemy funkcjonał

$$F[T] = \frac{16}{3} b^3 \int_0^a \left[\frac{2}{5} b^2 \left(\frac{\partial f}{\partial x}\right)^2 + f^2 + \left(\frac{4}{5} b^2 c \frac{\partial f}{\partial t} - q\right) f\right] dx, \tag{2.7}$$

który zgodnie z warunkiem Eulera osiąga minimum dla funkcji f spełniającej równanie

$$\frac{4}{5}b^2f''-2f=\frac{4}{5}b^2c\frac{\partial f}{\partial t}-q. \tag{2.8}$$

Dla rozwiązania tego równania wykonuje się przekształcenie Laplace'a. Przyjmując, że $F(x, s)=\mathcal{L}[f(x, t)]$ otrzymuje się równanie

$$\frac{4}{5}b^2F''-2\left(\frac{2}{5}b^2cs+1\right)F=-\frac{q}{s}. \tag{2.9}$$

Rozwiązanie powyższego równania z uwzględnieniem warunków brzegowych pozwala na wyznaczenie transformaty Laplace'a poszukiwanego rozkładu temperatury

$$T(x, y, s)=\frac{5q(b^2-y^2)}{4b^2cs\left(s+\dfrac{5}{2b^2c}\right)}\left[1-\frac{\operatorname{ch}\sqrt{c\left(s+\dfrac{5}{2b^2c}\right)}\,x}{\operatorname{ch}\sqrt{c\left(s+\dfrac{5}{2b^2c}\right)}\,a}\right]. \tag{2.10}$$

Korzystając ze wzorów podanych w [1], znaleźć można oryginał transformaty

$$T(x, y, t)=\frac{20qa^2}{\pi b^2}(b^2-y^2)\sum_{n=0}^{\infty}\frac{(-1)^n\cos\left(\dfrac{2n+1}{2a}\pi x\right)}{(2n+1)\left[(2n+1)^2\pi^2+10\dfrac{a^2}{b^2}\right]}\times$$

$$\times\left\{1-e^{-\left[\frac{5}{2b^2}+\left(\frac{2n+1}{2a}\pi\right)^2\right]\frac{t}{c}}\right\}. \tag{2.11}$$

Znaczne uproszczenie wzorów uzyskuje się przez wprowadzenie pojęcia uśrednionej stałej czasowej [8]. Jeżeli funkcja $U(t)$ określa pewien stan nieustalony, to uśrednioną stałą czasową dla tego przebiegu definiuje zależność

$$\tau=\int_0^{\infty}\frac{U(t)-U(\infty)}{U(0)-U(\infty)}\,dt=\lim_{s\to 0}\frac{[sU(s)]'}{U(0)-U(\infty)}, \tag{2.12}$$

gdzie $U(s)=\mathcal{L}[U(t)]$.

Zastosowanie uśrednionej stałej czasowej do równania (2.12) prowadzi do zależności

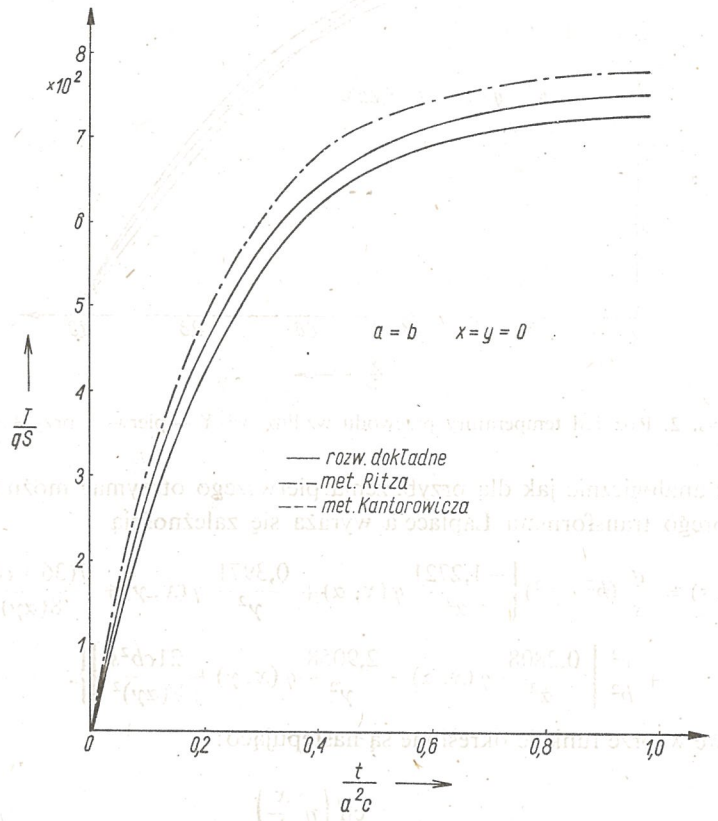
$$T(x, y, t)=\frac{q}{2}\left[1-\frac{\operatorname{ch}\left(\sqrt{\frac{5}{2}}\,\frac{x}{b}\right)}{\operatorname{ch}\left(\sqrt{\frac{5}{2}}\,\frac{a}{b}\right)}\right](b^2-y^2)\left(1-e^{-\frac{t}{\tau}}\right), \tag{2.13}$$

gdzie

$$\frac{\tau}{b^2c}=\frac{2}{5}-\frac{\dfrac{a-x}{b}\operatorname{sh}\left(\sqrt{\frac{5}{2}}\,\frac{a+x}{b}\right)+\dfrac{a+x}{b}\operatorname{sh}\left(\sqrt{\frac{5}{2}}\,\frac{a-x}{b}\right)}{2\sqrt{10}\operatorname{ch}\left(\sqrt{\frac{5}{2}}\,\frac{a}{b}\right)\left[\operatorname{ch}\left(\sqrt{\frac{5}{2}}\,\frac{a}{b}\right)-\operatorname{ch}\left(\sqrt{\frac{5}{2}}\,\frac{x}{b}\right)\right]}. \tag{2.14}$$

Otrzymane rozwiązanie zostało porównane z rozwiązaniem uzyskanym po zastosowaniu metody rozdzielenia zmiennych [5] oraz pierwszym przybliżaniem metody Ritza [11].

Do obliczeń przyjęto przekrój kwadratowy. Metoda Kantorowicza w porównaniu z metodą Ritza daje lepsze wyniki. Przebieg czasowy temperatury w osi przewodu (rys. 1) dla wybranych chwil $\frac{t}{a^2c} = 0{,}2;\ 0{,}6;\ 1{,}0$ obarczony jest błędem odpowiednio 7,5; 3,2;



Rys. 1. Zależność temperatury przewodu od czasu — pierwsze przybliżenie

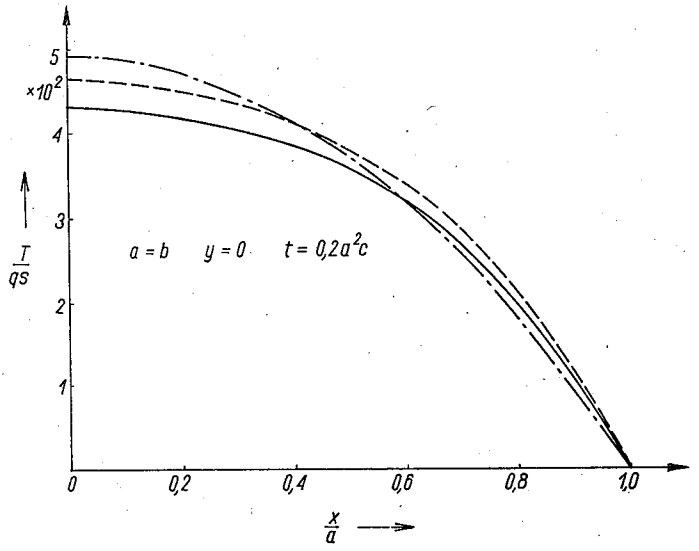
1,9%. Wyniki uzyskane metodą Rizta wykazują błędy odpowiednio 14,4; 7,2; 5,4%; błąd wyznaczono w stosunku do rozwiązania uzyskanego metodą rozdzielenia zmiennych jako rozwiązania dokładnego równania różniczkowego.

Rys. 2 ilustruje rozkład temperatury wzdłuż osi X dla wybranej chwili $\frac{t}{a^2c} = 0{,}2$. Rozkład ten charakteryzuje się stałym błędem około 7% wzdłuż osi X.

Założona postać (2.6) rozwiązania równania (2.1) może być traktowana jako pierwsze przybliżenie. Dla zwiększenia dokładności rozwiązania, kierując się zasadą by założona funkcja znikała dla $y = \pm b$, przyjąć należy postać bardziej złożoną (drugie przybliżenie)

$$T(x, y, t) = (b^2 - y^2)[f_1(x, t) + f_2(x, t)y^2], \tag{2.15}$$

gdzie funkcje f_1, f_2 spełniają warunki początkowe zerowe i znikają dla $x = \pm a$.



Rys. 2. Rozkład temperatury przewodu wzdłuż osi X — pierwsze przybliżenie

Postępując analogicznie jak dla przybliżenia pierwszego otrzymać można rozwiązanie równania, którego transformata Laplace'a wyraża się zależnością

$$T(x, y, s) = \frac{q}{s}(b^2-y^2)\left\{\frac{-1{,}2721}{\alpha^2}\varphi(x, \alpha)+\frac{0{,}3971}{\gamma^2}\varphi(x, \gamma)+\frac{7(36+cb^2s)}{8(\alpha\gamma)^2}+\right.$$

$$\left.+\frac{y^2}{b^2}\left[\frac{0{,}2808}{\alpha^2}\varphi(x, \alpha)-\frac{2{,}9058}{\gamma^2}\varphi(x, \gamma)+\frac{21cb^2s}{8(\alpha\gamma)^2}\right]\right\}. \tag{2.16}$$

Występujące we wzorze funkcje określone są następująco:

$$\varphi(x, \eta) = \frac{\operatorname{ch}\left(\eta\frac{x}{b}\right)}{\operatorname{ch}\left(\eta\frac{a}{b}\right)},$$

$$\alpha = \alpha(s) = \sqrt{14-\sqrt{133}+cb^2s},$$

$$\gamma = \gamma(s) = \sqrt{14+\sqrt{133}+cb^2s}.$$

Korzystając ze wzorów (2.12) i (2.15) wyznacza się uśrednioną stałą czasową

$$\frac{\tau}{b^2c} = \left\{1{,}2721\psi(x, \alpha_o)-0{,}3971\psi(x, \gamma_o)-0{,}2083+\frac{y^2}{b^2}[-0{,}2808\psi(x, \alpha_o)+\right.$$

$$\left.+2{,}9058\psi(x, \gamma_o)+0{,}0417]\right\}:\left\{0{,}5156\varphi(x, \alpha_o)-0{,}0156\varphi(x, \gamma_o)-0{,}5+\right.$$

$$\left.+\frac{y^2}{b^2}[-0{,}1138\varphi(x, \alpha_o)+0{,}1138\varphi(x, \gamma_o)]\right\}, \tag{2.17}$$

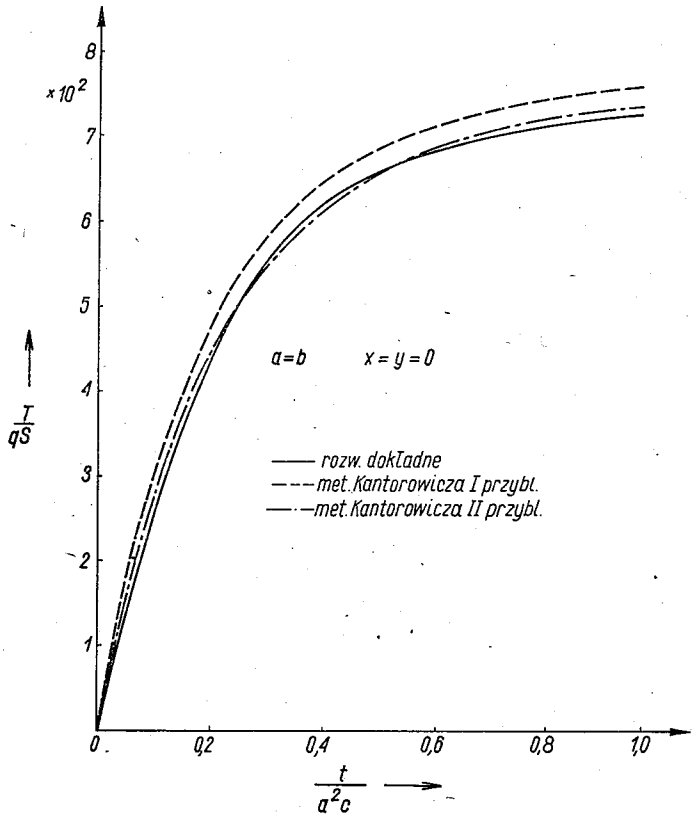
gdzie $\alpha_o = \alpha(0) = 1{,}5708079$, $\gamma_o = \gamma(0) = 5{,}0529756$ oraz

$$\psi(x, \eta) = \frac{1}{\eta^3}\left[\frac{(a-x)\operatorname{sh}\left(\eta\frac{a+x}{b}\right)+(a+x)\operatorname{sh}\left(\eta\frac{a-x}{b}\right)}{4b\operatorname{ch}\left(\eta\frac{a}{b}\right)}+\frac{1}{\eta}\varphi(x, \eta)\right].$$

Wówczas

$$T(x, y, t) = q(b^2-y^2)\left\{0{,}5156\varphi(x, \alpha_o)-0{,}0156\varphi(x, \varphi_o)-0{,}5+\right.$$
$$\left.+\frac{y^2}{b^2}[-0{,}1138\varphi(x, \alpha_o)+0{,}1138\varphi(x, \gamma_o)]\right\}\left(1-e^{-\frac{t}{\tau}}\right). \qquad (2.18)$$

W porównaniu z rozwiązaniem dokładnym, które wyraża się szeregiem podwójnym [3], otrzymane związki cechuje prostota, a jednocześnie zadowalająca dokładność. Wyniki obliczeń przedstawione są na rys. 3. Dla chwil czasowych jak poprzednio $\frac{t}{a^2c} = 0{,}2$; 0,6 i 1,0 rozkład temperatury obarczony jest błędem odpowiednio: 1,7; 0,8; 0,5%.



Rys. 3. Zależność temperatury przewodu od czasu. Porównanie wyników pierwszego i drugiego przybliżenia

2.2. Trzecie zagadnienie brzegowe

Założenie, że powierzchnia przewodnika oddaje ciepło zgodnie z prawem Newtona, wymaga rozwiązania równania (2.1) z warunkiem początkowym (2.2) i brzegowym trzeciego rodzaju

$$\left(\frac{\partial T}{\partial x}\right)_{x=-a} - hT(-a, y, t) = \left(\frac{\partial T}{\partial x}\right)_{x=a} + hT(a, y, t) = 0 \quad \text{dla} \quad -b \leqslant y \leqslant b, \tag{2.19}$$

$$\left(\frac{\partial T}{\partial y}\right)_{y=-b} - hT(x, -b, t) = \left(\frac{\partial T}{\partial y}\right)_{y=b} + hT(x, b, t) = 0 \text{ dla } -a \leqslant x \leqslant a, \tag{2.20}$$

gdzie h jest współczynnikiem oddawania ciepła.

Rozwiązanie równania (2.1) przewiduje się w postaci

$$T(x, y, t) = f(x, t)(\alpha - y^2), \tag{2.21}$$

przy czym α dobiera się tak, by dla $y = \pm b$ wypełnić warunek (2.20)

$$\alpha = \frac{b}{h}(2 + hb). \tag{2.22}$$

Rozwiązanie sprowadza się do minimalizacji funkcjonału

$$F[T] = \int\int_D \left[\text{grad}^2 T + 2T\left(c\frac{\partial T}{\partial t} - q\right)\right] dx\,dy + \int_\Gamma hT^2 ds, \tag{2.23}$$

który po uwzględnieniu założenia (2.21) i scałkowaniu względem zmiennej y przyjmuje postać

$$F[T] = 2\int_{-a}^{a} \left\{(f')^2\left(\alpha^2 b - \alpha b^3 + \frac{1}{5}b^5\right) + f^2\left[\frac{4}{3}b^3 + h(\alpha - b^2)^2\right] + \right.$$

$$+ c\left(\alpha^2 b - \frac{2}{3}\alpha b^3 + \frac{1}{5}b^5\right) f\frac{\partial f}{\partial t} - 2q\left(\alpha b - \frac{1}{3}b^3\right) f\Big\} dx +$$

$$+ 2h[f^2(-a, t) + f^2(a, t)]\left(\alpha^2 b - \frac{2}{3}\alpha b^3 + \frac{1}{5}b^5\right). \tag{2.24}$$

Warunkiem koniecznym istnienia minimum powyższego funkcjonału jest, by funkcja f spełniała równanie Eulera, które w postaci operatorowej jest następujące

$$F''\left(\alpha^2 b - \frac{2}{3}\alpha b^3 + \frac{1}{5}b^5\right) - F\left[\frac{4}{3}b^3 + h(\alpha - b^2)^2 + \text{cs}\left(\alpha^2 b - \frac{2}{3}\alpha b^3 + \frac{1}{5}b^5\right)\right] =$$

$$= -\frac{q}{s}\left(\alpha b - \frac{1}{3}b^3\right), \tag{2.25}$$

przy warunku

$$F'(-a, s) - hF(-a, s) = F'(a, s) + hF(a, s) = 0. \tag{2.26}$$

Można zauważyć, że warunki brzegowe (2.26) i (2.19) są równoważne. Ostatecznie, roz-

wiązanie zagadnienia uwzględniające uśrednioną stałą czasową przyjmuje postać

$$T(x, y, t) = \frac{qb^2}{2}\left(\frac{2+bh}{bh} - \frac{y^2}{b^2}\right)\left(1 - \frac{\operatorname{ch} K\frac{x}{b}}{\frac{K}{bh}\operatorname{sh} K\frac{a}{b} + \operatorname{ch} K\frac{a}{b}}\right)\left(1 - e^{-\frac{t}{\tau}}\right), \quad (2.27)$$

gdzie

$$\frac{\tau}{b^2 c} = \frac{15+10bh+2b^2h^2}{15bh+5b^2h^2} -$$

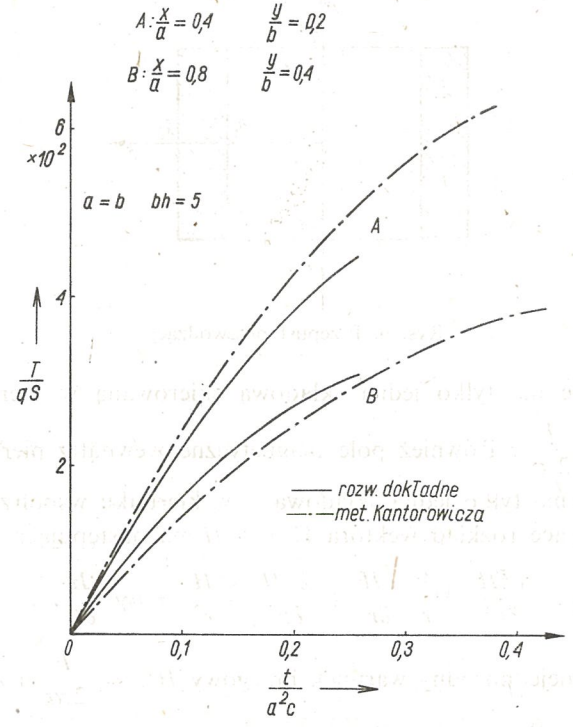
$$+ \frac{K}{bh}\left(\frac{a-x}{b}\operatorname{ch} K\frac{a+x}{b} + \frac{a+x}{b}\operatorname{ch} K\frac{a-x}{b}\right) + \frac{a-x}{b}\operatorname{sh} K\frac{a+x}{b} +$$

$$+ \frac{a+x}{b}\operatorname{sh} K\frac{a-x}{b} - \frac{1}{bh}\operatorname{sh} K\frac{a}{b}\operatorname{ch} K\frac{x}{b}$$

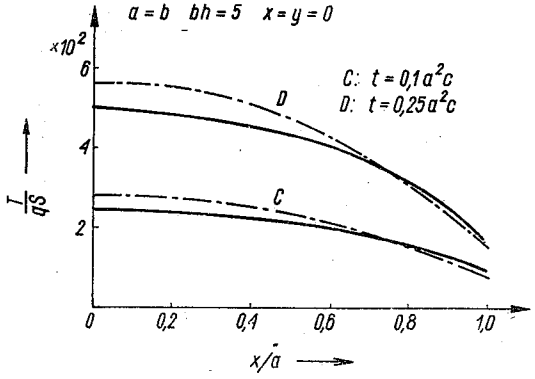
$$- \frac{}{4K\left(\frac{K}{bh}\operatorname{sh} K\frac{a}{b} + \operatorname{ch} K\frac{a}{b}\right)\left(\frac{K}{bh}\operatorname{sh} K\frac{a}{b} + \operatorname{ch} K\frac{a}{b} - \operatorname{ch} K\frac{x}{b}\right)},$$

$$K = \sqrt{\frac{5bh(3+bh)}{15+10bh+2b^2h^2}}.$$

Wyniki obliczeń dokonanych dla przewodu o przekroju kwadratowym, przy założeniu $bh = 5$, przedstawione są na rys. 4 i 5 wraz z wykresami sporządzonymi w oparciu o wzory z [2], gdzie zagadnienie takie rozwiązano stosując metodę rozdzielenia zmiennych.



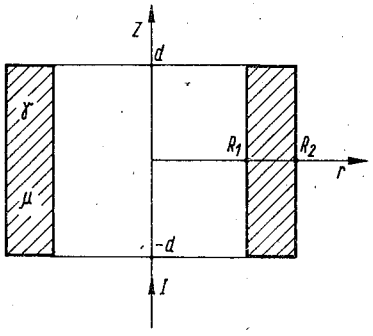
Rys. 4. Rozkład temperatury przewodu od czasu w przypadku swobodnej wymiany ciepła



Rys. 5. Rozkład temperatury przewodu wzdłuż osi X w przypadku swobodnej wymiany ciepła

3. POLE ELEKTROMAGNETYCZNE W PRZEPUŚCIE PRZEWODZĄCYM

Pierścień przewodzący o dowolnych, stałych parametrach: przenikalności magnetycznej μ i konduktywności γ oraz prostokątnym przekroju otacza umieszczony wzdłuż osi przewód (rys. 6). Począwszy od chwili $t = 0$ w przewodzie zaczyna płynąć prąd o natężeniu I wywołujący pole magnetyczne. Należy wyznaczyć nieustalony przebieg pola magnetycznego w pierścieniu.



Rys. 6. Przepust przewodzący

Pole wzbudzające ma tylko jedną składową skierowaną w kierunku współrzędnej φ o wartości $H = \frac{I}{2\pi r}$. Również pole magnetyczne wewnątrz pierścienia wytworzone przez prądy wirowe ma tylko jedną składową — w kierunku współrzędnej φ. Równanie różniczkowe określające rozkład wektora $\mathbf{H} = \mathbf{1}_\varphi H$ ma następującą postać [2]

$$\frac{\partial^2 H}{\partial r^2} + \frac{1}{r}\frac{\partial H}{\partial r} + \frac{\partial^2 H}{\partial z^2} - \frac{H}{r^2} = \mu\gamma\frac{\partial H}{\partial t}. \tag{3.1}$$

Funkcja H spełnia niejednorodny warunek brzegowy $H|_\Gamma = \frac{I}{2\pi r}$ i zerowy warunek początkowy $H(r, z, 0) = 0$.

Można wykazać, że równanie (3.1) jest równoważne równaniu (1.1). Dokonując podstawienia

$$u(r, z, t) = rH - \frac{I(t)}{2\pi}, \tag{3.2}$$

równanie (3.1) sprowadzić można do postaci (1.1)

$$\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial u}{\partial r}\right) + \frac{\partial}{\partial z}\left(\frac{1}{r}\frac{\partial u}{\partial z}\right) = \frac{\mu\gamma}{r}\left(\frac{\partial u}{\partial t} + \frac{I'(t)}{2\pi}\right). \tag{3.3}$$

Funkcja $u(r, z, t)$ spełnia jednorodne warunki brzegowe i podobnie jak $H(r, z, t)$ zerowy warunek początkowy.

Zakłada się następującą postać rozwiązania

$$u(r, z, t) = f(z, t)(r - R_1)(r - R_2). \tag{3.4}$$

Dla przyjętej postaci funkcji $u(r, z, t)$ po wykonaniu całkowania względem zmiennej z, funkcjonał (1.6) zapisze się wzorem

$$F[u] = \int_{-d}^{d} \left\{B(f')^2 + Df^2 + \mu\gamma\left[\frac{I'(t)}{2\pi} E + 2B\frac{\partial f}{\partial t}\right] f\right\} dz, \tag{3.5}$$

gdzie:

$$B = \frac{1}{12}(R_2^2 - R_1^2)(R_2^2 - 8R_1R_2 + R_1^2) + (R_1R_2)^2 \ln\frac{R_2}{R_1},$$

$$D = (R_1 + R_2)^2 \ln\frac{R_2}{R_1} - 2(R_2^2 - R_1^2),$$

$$E = 2R_1R_2 \ln\frac{R_2}{R_1} - R_2^2 + R_1^2.$$

Równanie Eulera, dla powyższego funkcjonału, wyraża się zależnością

$$Bf'' - Df - \mu\gamma B\frac{\partial f}{\partial t} = \mu\gamma\frac{I'(t)}{4\pi} E. \tag{3.6}$$

Dokonując transformacji Laplace'a oraz rozwiązując równanie (3.6), ze wzoru (3.4) otrzymuje się operatorową postać rozwiązania

$$H(r, z, s) = \frac{I(s)}{2\pi r} + \frac{\mu\gamma s I(s)}{4\pi(D + B\mu\gamma s)} \frac{(r - R_1)(R_2 - r)}{r}\left[1 - \frac{\operatorname{ch}\sqrt{\frac{D}{B} + \mu\gamma s}\, z}{\operatorname{ch}\sqrt{\frac{D}{B} + \mu\gamma s}\, d}\right], \tag{3.7}$$

gdzie stałe B i D jak we wzorze (3.5).

Rozwiązanie dokładne tego problemu znaleźć można w [8, 14], gdzie podana jest transformata Laplace'a natężenia pola wyrażająca się szeregiem, którego wyrazy o złożonej postaci zawierają zmodyfikowane funkcje Bessela. Autorzy wspomnianych prac posłużyli się pojęciem uśrednionej stałej czasowej. Zakładając, że prąd wzbudzający

ma następujący przebieg

$$I(t) = I_o\left(1-e^{-\frac{t}{\tau_o}}\right), \tag{3.8}$$

gdzie τ_o jest stałą czasową wymuszenia, wyznaczyli uśrednioną stałą czasową τ natężenia pola. W rezultacie, otrzymali oni następującą postać stałej czasowej

$$\frac{\tau-\tau_o}{\mu\gamma d^2} = \frac{d^2-z^2}{2d^2} + \\ -\frac{2r}{R_1 R_2 d^3}\sum_{n=0}^{\infty}\frac{(-1)^n}{m^3}\cos mz\,\frac{[R_2K_1(mR_2)-R_1K_1(mR_1)]I_1(mr)+[R_1I_1(mR_1)-R_2I_1(mR_2)]K_1(mr)}{I_1(mR_1)K_1(mR_2)-I_1(mR_2)K_1(mR_1)}, \tag{3.9}$$

gdzie $m = \frac{\pi}{2d}(2n+1)$.

Postępując analogicznie, wyznacza się stałą czasową dla wyrażenia (3.7)

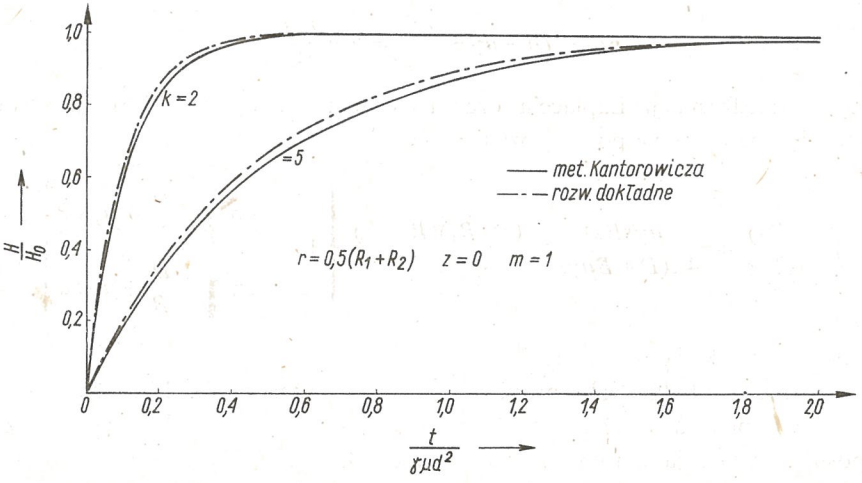
$$\frac{\tau-\tau_o}{\mu\gamma d^2} = kM^2\,\frac{k^2-2k\ln k-1}{(1+k)^2\ln k-2k^2+2}\left(\frac{r}{R_1}-1\right)\left(1-\frac{r}{R_2}\right)\left[1-\frac{\operatorname{ch}\left(K\frac{z}{d}\right)}{\operatorname{ch}K}\right], \tag{3.10}$$

gdzie:

$$k = \frac{R_2}{R_1}, \qquad M = \frac{R_1}{d},$$

$$K = \frac{1}{M}\sqrt{\frac{(1+k)^2\ln k-2k^2+2}{\frac{1}{12}(k^2-1)(k^2-8k+1)+k^2\ln k}}.$$

Na rys. 7 porównane zostały wyniki uzyskane metodą Kantorowicza (wzór (3.10)) i rozdzielenie zmiennych (wzór (3.9)). Przyjęto $z = 0$, $r = 0{,}5(R_1+R_2)$, $M = 1$. Z rys. 7 i wzorów (3.9) i (3.10) wynika, że kosztem niewielkiego błędu uzyskano znaczne uprosz-

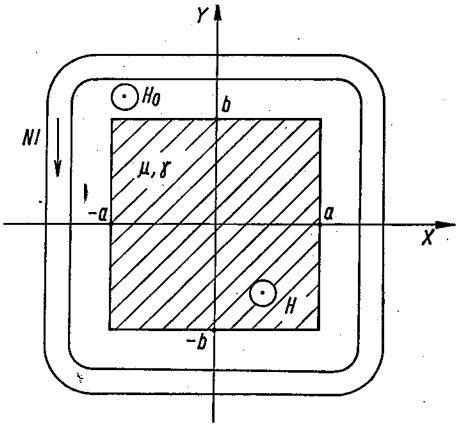


Rys. 7. Przebieg zmian pola magnetycznego w pierścieniu w funkcji czasu

czenie wyrażenia określającego uśrednioną stałą czasową dyfuzji pola magnetycznego. Wielkość błędu dla chwili $\frac{t}{\mu\gamma d^2} = 0{,}1;\ 0{,}3;\ 0{,}6$ wynosi dla $k = 2$ odpowiednio: 7,0; 5,7 i 3,9%, dla $k = 5$: 4,3; 1,4 i 0,2%.

4. POLE MAGNETYCZNE W RDZENIU O PRZEKROJU PROSTOKĄTNYM

Rdzeń o stałej przenikalności magnetycznej μ i konduktywności γ jest umieszczony wewnątrz cewki wzbudzającej pole magnetyczne (rys. 8). Długość cewki jest dostatecznie duża, aby w przypadku braku rdzenia pole wzbudzenia było jednorodne i równe ampe-



Rys. 8. Cewka wzbudzenia z rdzeniem prostokątnym

rozwojom NJ na jednostkę długości. Należy obliczyć rozkład pola magnetycznego w rdzeniu. Pole magnetyczne wytworzone przez zaindukowane prądy wirowe ma ten sam kierunek co pole wzbudzenia i skierowane jest wzdłuż osi OZ.

Pole magnetyczne opisane jest równaniem

$$\frac{\partial^2 H}{\partial x^2} + \frac{\partial^2 H}{\partial y^2} = \mu\gamma \frac{\partial H}{\partial t}. \tag{4.1}$$

Funkcja $H(x, y, t)$ spełnia warunek brzegowy $H|_\Gamma = H_0(t)$ i zerowy warunek początkowy $H(x, y, 0) = 0$.

Natężenie pola H poszukuje się w postaci sumy

$$H(x, y, t) = u(x, y, t) + H_0(t). \tag{4.2}$$

Ze wzorów (4.1) i (4.2) otrzymuje się

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \mu\gamma \left[\frac{\partial u}{\partial t} + H_0'(t)\right], \tag{4.3}$$

gdzie funkcja $u(x, y, t)$ spełnia zerowe warunki graniczne. Ostatnie równanie różni się od równania (2.1) jedynie wartością współczynników. Przyjmując funkcję $u(x, y, t)$

w postaci (2.6) otrzymuje się rozwiązanie analogiczne do (2.10) o następującej postaci

$$u(x, y, s) = -\frac{5\mu\gamma s H_o(s)}{10+4b^2\mu\gamma s}(b^2-y^2)\left[1-\frac{\operatorname{ch}\left(\sqrt{\frac{5}{2b^2}+\mu\gamma s}\,x\right)}{\operatorname{ch}\left(\sqrt{\frac{5}{2b^2}+\mu\gamma s}\,a\right)}\right]. \tag{4.4}$$

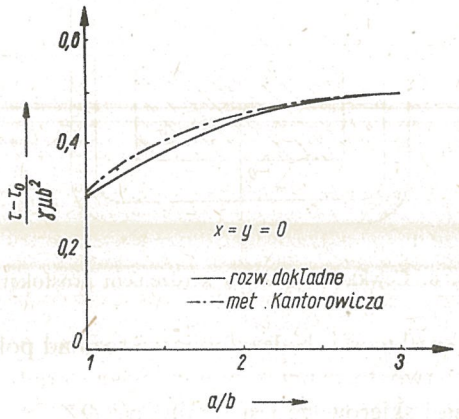
Ze wzorów (4.2) i (4.4) wyznacza się transformatę natężenie pola magnetycznego.

Zakładając, że pole wzbudzenia narasta ze stałą czasową τ_o

$$H_o(t) = H_o\left(1-e^{-\frac{t}{\tau_o}}\right),$$

otrzymano

$$\frac{\tau-\tau_o}{\mu\gamma b^2} = \frac{1}{2}\left(1-\frac{y^2}{b^2}\right)\left[1-\frac{\operatorname{ch}\left(\sqrt{\frac{5}{2}}\,\frac{x}{b}\right)}{\operatorname{ch}\left(\sqrt{\frac{5}{2}}\,\frac{a}{b}\right)}\right]. \tag{4.5}$$



Rys. 9. Zależność uśrednionej stałej czasowej od stosunku wymiarów a/b

Na rys. 9 przedstawiono zależność stałej czasowej w punkcie $x = y = 0$ od stosunku wymiaru boków a/b. Jednocześnie porównano wyniki metody Kantorowicza (wzór (4.4)) z wartością dokładną uzyskaną metodą rozdzielenia zmiennych [8]. Błąd nie przekracza 6%.

5. PODSUMOWANIE

W pracy zastosowana została metoda Kantorowicza do rozwiązywania problemów prowadzących do równania przewodnictwa. Konfrontacja uzyskanych wyników z rozwiązaniami dokładnymi świadczy o skuteczności metody. Prosta, zwarta forma otrzymanych zależności, przy zadowalającej dokładności, świadczy o jej przydatności w praktyce inżynierskiej.

LITERATURA

1. W. A. Ditkin, A. P. Prudnikow, *Intiegralnyje prieobrazywanija i opieracjonnye i szcislenije*, Moskwa 1974.
2. E. Kącki, *Równania różniczkowe cząstkowe w elektrotechnice*, WNT, Warszawa 1971.
3. E. Kącki, *Termokinetyka*, WNT, Warszawa 1967.
4. M. G. Kogan, *Rieszenije nieliniejnych zadacz tieorii tiepłoprowodnosti mietodom Kantorowicza*, Inż.-Fiz.-Żurnał, 1967, t. XII, nr 1.
5. B. Krajewski, *Application of variational methods to problems of unsteady head flow*, Arch. Mech. Stosow., 5, 20, 1968.
6. B. Krajewski, *Ein direktes Variationsverfahren zur Behandlung der Wärmeübertragungsprobleme für erzwungene Konvektion*, Int. J. Head Mass Transfer, 16, 1973.
7. B. Krajewski, *Ujęcie wariacyjne zagadnień związanych z nieliniowym przewodzeniem ciepła w stanach nieustalonych*, Biul. Inf. ITC Pol. Warsz., 1974, nr 42.
8. W. Lipiński, *Analiza dynamiki pola elektromagnetycznego w środowiskach przewodzących*, Rozprawa habilitacyjna PNPS, nr 63, IAP, Szczecin 1976.
9. J. Łach, *Nieustalone pole temperatury przy nieliniowym warunku brzegowym*, Arch. Bud. Maszyn, 1973, t. XX, zeszyt 3.
10. J. Łach, *Zastosowanie zmodyfikowanej metody Kantorowicza do wyznaczania nieustalonego pola temperatury*, Biul. Inf. ITC Pol. Warsz., 1972, nr 36.
11. J. Purczyński, A. Brykalski, J. Trzaska, *Metody bezpośredniej analizy stanów nieustalonych w obiektach o parametrach rozłożonych PNPS* (w druku).
12. P. Rolicz, *Zastosowanie bezpośrednich metod analizy funkcjonalnej do analizy nieustalonych procesów elektromagnetycznych w maszynach i urządzeniach elektrycznych*, Rozprawa habilitacyjna Pol. Pozn., Rozprawy, nr 85, Poznań 1977.
13. S. D. Savakar, *On a variational formulation of a class of thermal entrance problems*, Int. J. Head Mass Transfer, 13, 1970.
14. R. Sikora, J. Purczyński, W. Lipiński, *Stany nieustalone przebiegów elektromagnetycznych w przepuście przewodzącym i żłobku maszyny indukcyjnej*, Arch. Elektr., 1976, t. XXV, zeszyt 4.
15. W. J. Smirnow, *Kurs wysszej matiematiki*, t. II, Moskwa 1948.
16. O. C. Zienkiewicz, *Metoda elementów skończonych*, Warszawa 1972.

J. PURCZYŃSKI, A. BRYKALSKI

KANTOROWICZ' METHOD APPLIED TO SOLVE CONDUCTIVITY EQUATION

Summary

Some problems of heat conduction (in the transient state) have been calculated in the paper. Kantorowicz' method has been applied. The distribution of temperature field has been obtained in conductors whose cross-sections are rectangular and sector of ring. The first and third kind of boundary conditions have been taken into account. Simple expressions are given for the temperature field in this case. Also, the distribution of magnetic field has been calculated in a ring enclosing a conductor with current and core inside the coil.

J. PURCZYŃSKI, A. BRYKALSKI

APPLICATION DE LA MÉTHODE DE KANTOROWICZ À LA RÉSOLUTION DE L'ÉQUATION DE CONDUCTIBILITÉ

Résumé

Dans l'étude on a appliqué la méthode de Kantorowicz à la résolution de quelques problèmes liés à la conductibilité thermique à l'état instable. On a déterminé la répartition des températures dans le conducteur à section rectangulaire parcouru par le courant. Prenant en considération le premier et le troisième problèmes aux limites, on obtient de simples formules définissant la répartition des températures. En outre on a déterminé la répartition du champ magnétique dans l'anneau conducteur embrassant le conducteur avec le courant et dans le noyau placé à l'intérieur de la bobine.

J. PURCZYŃSKI, A. BRYKALSKI

ANWENDUNG DER KANTOROVITSCH-METHODE ZUR LÖSUNG DER LEITUNGSGLEICHUNG

Zusammenfassung

In der Arbeit wurde die Methode von Kantorovitsch zur Lösung einiger Probleme der Wärmeleitung im instationären Zustand angewandt. Die Temperaturverteilung in der stromleitenden Leitung mit rechteckigem Querschnitt und einem Querschnitt mit ringförmigem Ausschnitt wurde bestimmt. Es wurden die erste und die dritte Randbedingung berücksichtigt, und einfache Formeln für die Temperaturverteilung gewonnen. Außerdem wurde die Verteilung des magnetischen Feldes im leitenden, den Stromleiter umfassenden Ring sowie in einem im Spuleninnern befindlichen Kern bestimmt.

Я. ПУРЧИНЬСКИ, А. БРЫКАЛЬСКИ

ПРИМЕНЕНИЕ МЕТОДА КАНТОРОВИЧА К РЕШЕНИЮ УРАВНЕНИЯ ПРОВОДИМОСТИ

Резюме

Применен метод Канторовича к решению нескольких проблем теплопровидимости в неустановившемся режиме. Определено распределение температуры в проводнике с током имеющем прямоугольное сечение. Учтены первое и третье краевое условие. Получены простые формулы определяющие распределение температуры. Кроме того определено распределение магнитного поля в проводящем кольце обнимающем проводник и в сердечнике находящемся внутри катушки.