

można efektywny współczynnik załamania zapisać następująco [25]:

$$n_{R,ef} = [b_0 n_{R2}^2 + (1 - b_0) n_{R1}^2]^{1/2}, \quad (45)$$

gdzie b_0 jest rozwiązaniem równania (42), w którym znormalizowaną częstotliwość ν zastąpiono liczbą rzeczywistą ν_0 :

$$\nu_0 = \frac{1}{2} k_0 d \sqrt{n_{R2}^2 - n_{R1}^2}. \quad (46)$$

Wówczas znormalizowaną stałą propagacji można przedstawić w postaci [25]:

$$b = b_0 + ib_i = b_0 + \Delta\nu \left. \frac{db}{d\nu} \right|_{\nu=\nu_0}, \quad b_i \ll b_0, \quad (47)$$

w której

$$\Delta\nu = i \frac{\nu_0}{n_{R2}^2 - n_{R1}^2} \frac{n_{R2}g + n_{R1}\alpha}{2k_0}. \quad (48)$$

W powyższej zależności przez g i α oznaczono odpowiednio wzmocnienie promieniowania w obszarze czynnym oraz jego straty w warstwach ograniczających, tj. przyjęto zespolone stałe dielektryczne w tych obszarach w postaci (por. tablicę):

$$\varepsilon_{R1} = n_{R1}^2 - i \frac{n_{R1}\alpha}{k_0} \quad (49a)$$

$$\varepsilon_{R2} = n_{R2}^2 + i \frac{n_{R2}g}{k_0}. \quad (49b)$$

Wprowadźmy efektywny współczynnik wzmocnienia g_{ef} :

$$N_{R,ef}^2 = n_{R,ef}^2 + i \frac{n_{R,ef}g_{ef}}{k_0} \quad (50)$$

który, podstawiając (47) i (48) do (41), można wyrazić w postaci:

$$g_{ef} = 2k_0 \operatorname{Im}(N_{R,ef}) = \frac{1}{n_{R,ef}} [\Gamma n_{R2}g - (1 - \Gamma)n_{R1}\alpha]. \quad (51)$$

W zależności tej występuje współczynnik wypełnienia Γ , zdefiniowany w sposób następujący [44]:

$$\Gamma = b_0 + \frac{\nu_0}{2} \left. \frac{db}{d\nu} \right|_{\nu=\nu_0} = b_0 + (1 - b_0) \frac{\nu_0 \sqrt{b_0}}{1 + \nu_0 \sqrt{b_0}}. \quad (52)$$

Uwzględniając, że

$$n_{R1} \leq n_{R,ef} \leq n_{R2}; \quad (53) \quad n_{R2} - n_{R1} \ll n_{R1} \quad (54)$$

można równanie (51) uprościć do postaci:

$$g_{ef} \approx \Gamma g - (1 - \Gamma)\alpha. \quad (55)$$

Powyższe zależności znacznie się upraszczają dla przypadków granicznych. Na przykład dla $\nu_0 \leq 0.5$ można stosować następujące rozwinięcie [31]:

$$b_0 = \nu_0^2 - \frac{4}{3}\nu_0^4 + \frac{92}{45}\nu_0^6 - \frac{1072}{315}\nu_0^8 + \dots \quad (56)$$

Wówczas zależność na współczynnik wypełnienia upraszcza się do postaci:

$$\Gamma = 2\nu_0^2 - 4\nu_0^4 + \frac{368}{45}\nu_0^6 - \frac{1072}{63}\nu_0^8 + \dots \quad (57)$$

Dla dużych wartości ν_0 zarówno b_0 , jak i Γ , dążą asymptotycznie do jedności. Wówczas użyteczne wzory przybliżone można zapisać następująco [45, 46]:

$$b_0 \approx 1 - \frac{1}{2\nu_0^2} \ln(1 + 2\nu_0^2); \quad (58)$$

$$\Gamma \approx \frac{2\nu_0^2}{1 + 2\nu_0^2}. \quad (59)$$

W przypadku struktur o zmiennej wartości współczynnika załamania $N_{R2}(y)$ w obszarze czynnym, postępowanie jest zasadniczo analogiczne. Potraktujmy zmiany współczynnika załamania jako zaburzenie:

$$N_{R2}^2(y) = N_{R2}^2(0) + f(y), \quad (60)$$

gdzie w ogólności $f(y)$ jest zespolone. Wówczas kolejne zależności zmieniają się w sposób następujący [31]:

$$N_{R,ef}(y) = \{b[N_{R2}^2(0) + f(y)] + (1-b)N_{R1}^2\}^{1/2} \quad (61)$$

$$b = b_0 + \frac{db_0}{d\nu_0} [\nu(y) - \nu_0] + \frac{1}{2} \frac{d^2b_0}{d\nu_0^2} [\nu(y) - \nu_0]^2 + \dots \quad (62)$$

$$\begin{aligned} \nu(y) - \nu_0 &= \frac{1}{2} k_0 d \left(\sqrt{N_{R2}^2(0) + f(y)} - \sqrt{N_{R2}^2(0) - N_{R1}^2} \right) = \nu_0 \left(\sqrt{1 + \frac{f(y)}{\Delta\epsilon_0}} - 1 \right) \approx \\ &\approx \nu_0 \left[\frac{1}{2} \frac{f(y)}{\Delta\epsilon_0} - \frac{1}{8} \left(\frac{f(y)}{\Delta\epsilon_0} \right)^2 + \dots \right] \end{aligned} \quad (63)$$

gdzie

$$\Delta\epsilon_0 = N_{R2}^2(0) - N_{R1}^2. \quad (64)$$

Wstawiając (62) i (63) do (61) oraz zaniedbując wyrazy wyższych rzędów, otrzymujemy stąd:

$$N_{R,ef}(y) \approx [b_0 N_{R2}^2(0) + (1-b_0)N_{R1}^2 + \Gamma f(y)]^{1/2}. \quad (65)$$

4. EFEKT FALOWODOWY W STRUKTURZE PLANARNEJ

W płaszczyźnie złącza lasera z kontaktem paskowym efekt falowodowy, a więc i kształt funkcji $f(y)$, zależy od zmian obu części: rzeczywistej i urojonej, zespolonego współczynnika załamania $N_{R2}(y)$.

Przyjmijmy następujące oznaczenia:

$$n_{R2}(y) = n_{R2}(0) + \Delta n_{R2}(y) \quad (66)$$

$$k_{e2}(y) = k_{e2}(0) + \Delta k_{e2}(y) \quad (67)$$

$$f(y) = f_r(y) - if_i(y). \quad (68)$$

Wstawiając (66) i (67) do (60) otrzymujemy stąd dla małych zaburzeń Δn_{R2} i Δk_{e2} :

$$f_r(y) = 2[n_{R2}(0)\Delta n_{R2}(y) - k_{e2}(0)\Delta k_{e2}(y)] \quad (69)$$

$$f_i(y) = 2[n_{R2}(0)\Delta k_{e2}(y) + k_{e2}(0)\Delta n_{R2}(y)]. \quad (70)$$

Zmiana stałej ekstynkcji następuje w wyniku zmian absorpcji promieniowania. Generalnie absorpcja ta jest ujemna w centrum obszaru czynnego (wzmocnienie), a staje się dodatnia (straty) w pewnej odległości od centrum. Rozkład wzmocnienia lokalnego $g(y)$ wiąże się z rozkładem $k_{e2}(y)$ następująco (por. zależności (33) i (37)):

$$g(y) = -\alpha_2(y) = -\frac{4\pi}{\lambda} k_{e2}(y), \quad (71)$$

gdzie $\alpha_2(y)$ jest rozkładem współczynnika absorpcji w obszarze czynnym. Zmianę stałej ekstynkcji można stąd wyrazić w następującej postaci:

$$\Delta k_{e2}(y) = \frac{\lambda}{4\pi} [g(0) - g(y)]. \quad (72)$$

Współczynnik załamania $n_{R2}(y)$ zmienia się z kolei w wyniku zmiany absorpcji na swobodnych nośnikach i zmiany temperatury:

$$n_{R2}(y) = \Omega_{fc}[n_{fc}(0) - n_{fc}(y)] - \Omega_T[T(0) - T(y)], \quad (73)$$

gdzie n_{fc} oznacza całkowitą koncentrację swobodnych nośników, T — temperaturę, obie te wielkości dotyczą obszaru czynnego, a $\Omega_{fc} = (4 \div 11)10^{-21} \text{ cm}^3$ i $\Omega_T = 4,9 \cdot 10^{-4} \text{ K}^{-1}$ są współczynnikami proporcjonalności. Ich wartości liczbowe podano dla GaAs.

Ponieważ zmiany Δn_{R2} pod wpływem zmian temperatury i koncentracji swobodnych nośników są przeciwnego znaku, sumaryczny efekt jest przeważnie niewielki. Na przykład w standardowych biheterozłączowych laserach z kontaktem paskowym ($n_{R2} = 3.6$,

$$n_{R1} = 3.4, n_{fc}(0) - n_{fc}\left(\frac{S}{2}\right) = 10^{18} \text{ cm}^{-3}, T(0) - T\left(\frac{S}{2}\right) = 1.5 \text{ K}, \lambda = 870 \text{ nm}) \text{ oba efekty}$$

całkowicie się równoważą dla grubości obszaru czynnego $d = 0.08 \mu\text{m}$ [7]. W laserach z cieńszym obszarem czynnym występuje pewien niewielki rzeczywisty efekt falowodowy, natomiast w laserach, w których $d > 0.08 \mu\text{m}$, ma miejsce rzeczywisty efekt antyfalowodowy (Rys. 3a).

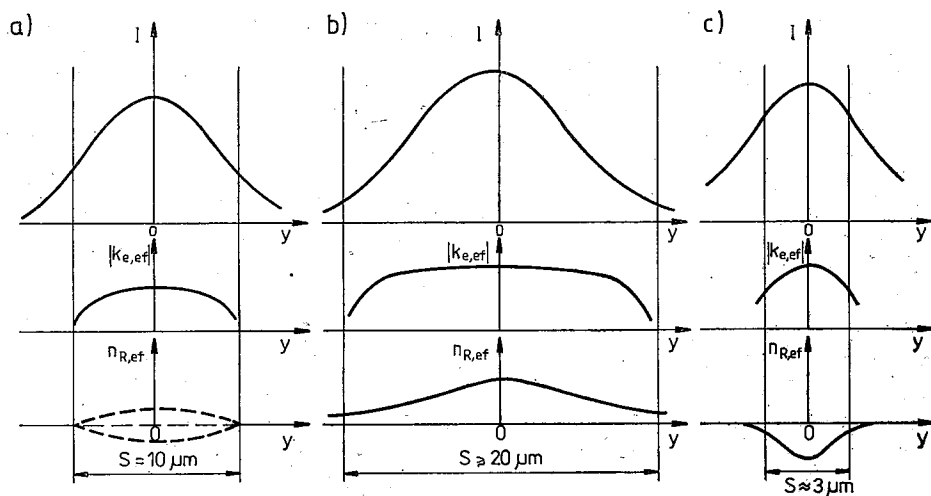
Podsumowując analizę własności falowodowych planarnych struktur laserów paskowych można stwierdzić, że podobnie jak to ma miejsce w laserach szerokokontaktowych, również w biheterozłączowych laserach z kontaktem paskowym — w płaszczyźnie prostopadłej do złącza — efekt falowodowy jest spowodowany przez skokową zmianę rzeczywistej części zespolonego współczynnika załamania na granicach wnęki rezonansowej (falowód rzeczywisty).

W płaszczyźnie złącza laserów z kontaktem paskowym o standardowej szerokości obszaru czynnego ($S \approx 10 \mu\text{m}$) dominujące znaczenie ma natomiast przestrzenny rozkład

wzmocnienia lokalnego $g(y)$, czyli zmiana stałej ekstynkcji $\Delta k_{e2}(y)$ — falowód wzmocnieniowy, a zmiana $\Delta n_{R2}(y)$ daje jedynie niewielki, dodatkowy efekt falowodowy lub antyfalowodowy, zależnie od grubości obszaru czynnego. Powyższe wnioski zostały potwierdzone przez pomiary Cooka i Nasha [5], a konkretnie po stwierdzeniu astygmatyzmu promieniowania emitowanego przez lasery z kontaktem paskowym.

Wzmocnieniowy efekt falowodowy dominuje również w płaszczyźnie złącza laserów z szerokimi kontaktami paskowymi ($S \geq 20 \mu\text{m}$). Dodatkowo wspomaga go jednak niewielki rzeczywisty efekt falowodowy (Rys. 3b) spowodowany rozkładem temperatury. Wpływ prawie jednorodnego rozkładu koncentracji nośników jest w tym przypadku pomijalny.

Całkiem odmienna sytuacja istnieje w laserach z wąskimi kontaktami paskowymi ($S \approx 3 \mu\text{m}$) [47]. W tym przypadku pole modu promieniowania w znacznym stopniu penetruje warstwy bierne (Rys. 3c) i nie powoduje istotnego obniżenia koncentracji



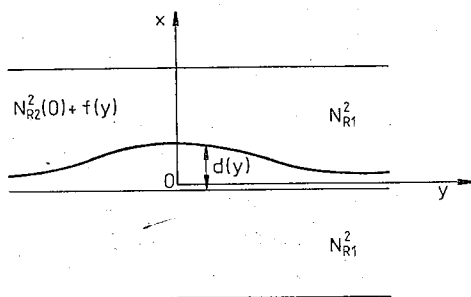
Rys. 3. Efekt falowodowy w płaszczyźnie złącza laserów z kontaktem paskowym: a) laser standardowy ($S \approx 10 \mu\text{m}$), b) laser z szerokim obszarem czynnym ($S \geq 20 \mu\text{m}$), c) laser z wąskim obszarem czynnym ($S \approx 3 \mu\text{m}$). $I(y)$ — rozkład emitancji energetycznej promieniowania w rezonatorze

nośników wewnątrz obszaru czynnego. Wytwarza się wówczas silny antyfalowód rzeczywisty konkurujący z falowodem wzmocnieniowym. Podczas pracy impulsowej omawianych laserów ich efekt falowodowy staje się więc bardzo niestabilny w wyniku wzrostu temperatury w rezonatorze. Prowadzi to do znacznie większych zmian charakterystyk kierunkowych promieniowania [48], niż to ma miejsce w przypadku laserów szerokokontaktowych.

5. EFEKT FALOWODOWY W STRUKTURZE NIEPLANARNEJ

W przypadku struktur nieplanarnych (Rys. 4), w których grubość warstwy czynnej, $d(y)$, jest funkcją położenia, funkcjami położenia stają się również: znormalizowana częstotliwość $\nu(y)$, znormalizowana stała propagacji $b(y)$, współczynnik wypełnienia $\Gamma(y)$ oraz

efektywny współczynnik wzmocnienia $g_{ef}(y)$. Problem komplikuje się jeszcze bardziej, gdy mamy do czynienia zarówno ze zmianą grubości warstwy czynnej, jak i ze zmianą zespolonego współczynnika załamania w tym obszarze. Ostatnia z tych zmian może być stosunkowo duża, ponieważ jest ona między innymi związana z koncentracją nośników swobodnych w warstwie czynnej, która z kolei może być silną funkcją położenia z uwagi na zmiany grubości warstwy czynnej. Znaczna zmiana zespolonego współczynnika załamania może też być narzucona konstrukcyjnie (wbudowany falowód). W powyższych przypadkach nie można więc pominąć w rozwinięciach (62) i (63) wyrazów wyższych rzędów.



Rys. 4. Rezonator nieplanarnego lasera paskowego

Analiza tego przypadku przeprowadzona przez J. Buusa w pracy [31] doprowadziła do następującego wyniku:

$$N_{R,ef}^2(y) = b_0 N_{R2}^2 + (1 - b_0) N_{R1}^2 + \Gamma f(y) + \Delta_0(y) \frac{f^2(y)}{\Delta \varepsilon_0} + \\ + E_0(y) \frac{f^3(y)}{\Delta \varepsilon_0^2} + Z_0(y) \frac{f^4(y)}{\Delta \varepsilon_0^3} + \dots \quad (74)$$

$$z \Delta_0(y) = \frac{3}{8} v_0(y) \frac{db_0}{dv_0} + \frac{1}{8} v_0^3(y) \frac{d^2 b_0}{dv_0^2} \quad (75)$$

$$E_0(y) = -\frac{1}{16} v_0(y) \frac{db_0}{dv_0} + \frac{1}{16} v_0^2(y) \frac{d^2 b_0}{dv_0^2} + \frac{1}{48} v_0^3(y) \frac{d^3 b_0}{dv_0^3} \quad (76)$$

$$Z_0(y) = \frac{3}{128} v_0(y) \frac{db_0}{dv_0} - \frac{3}{128} v_0^2(y) \frac{d^2 b_0}{dv_0^2} + \frac{1}{192} v_0^3(y) \frac{d^3 b_0}{dv_0^3} + \frac{1}{348} v_0^4(y) \frac{d^4 b_0}{dv_0^4}. \quad (77)$$

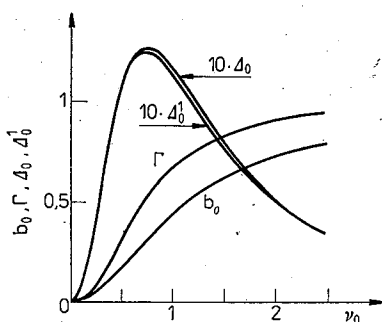
Zależność (75) można przekształcić do prostszej postaci:

$$\Delta_0 = \frac{1}{4} v_0 \frac{d\Gamma}{dv_0}, \quad (78)$$

a stosując zależność przybliżoną (59), otrzymujemy ostatecznie:

$$\Delta_0 \approx \Delta_0^1 = \frac{v_0^2}{(1 + 2v_0^2)^2}. \quad (79)$$

Na Rys. 5 pokazano zależność znormalizowanej stałej propagacji b_0 , współczynnika wypełnienia Γ oraz Δ_0 i Δ_0^1 od znormalizowanej częstotliwości ν_0 .

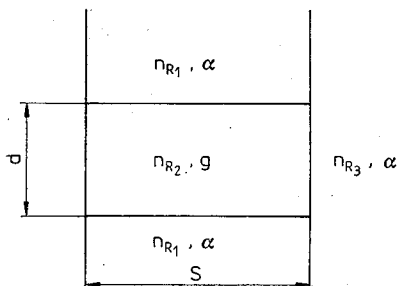


Rys. 5. Zależność znormalizowanej stałej propagacji b_0 , współczynnika wypełnienia Γ oraz parametrów Δ_0 (78) i Δ_0^1 (79) od znormalizowanej częstotliwości ν_0 w nieplanarnym laserze złączowym [31]

Zwykle w rozwinięciu (74) pomija się wyrazy z funkcją $f(y)$ w potęgze wyższej niż 2 i zastępuje się Δ_0 przez Δ_0^1 . W przypadku umiarkowanego zaburzenia ($f \approx -0,2$) powoduje to powstanie błędu mniejszego od 1%. Jednakże dla większych bezwzględnych wartości f , błąd ten może nawet przekroczyć 5%. Należy wówczas uwzględnić następny wyraz rozwinięcia (z f^3), co spowoduje zmniejszenie się tego błędu do wartości poniżej 0,2% nawet dla $f = -0,4$.

6. EFEKT FALOWODOWY W LASERACH Z ZAGRZEBANYM OBSZAREM CZYNNYM

W laserach z zagrzebanym obszarem czynnym o strukturze pokazanej schematycznie na Rys. 6, należy stosować metodę efektywnego współczynnika załamania w formie nieco zmodyfikowanej. Dla uproszczenia analizy przyjmujemy jednorodny rozkład współczynnika wzmocnienia g w obszarze czynnym oraz jednakowy współczynnik strat α poza nim.



Rys. 6. Rezonator lasera złączowego z zagrzebanym obszarem czynnym

Rozpatrzmy wpięrk strukturę o nieskończenie dużej szerokości $S = \infty$ obszaru czynnego. Wówczas, podobnie jak w rozdziale 3, możemy wprowadzić znormalizowaną częstotliwość ν_∞ :

$$\nu_{\infty} = k_0 \frac{d}{2} \sqrt{n_{R2}^2 - n_{R1}^2} \quad (80)$$

dzięki czemu otrzymujemy równanie transcendentálne na znormalizowaną stałą propagacji b_{∞} :

$$\nu_{\infty} \sqrt{1 - b_{\infty}} = \arctan[\sqrt{b_{\infty}/(1 - b_{\infty})}] \quad (81)$$

oraz zależność na współczynnik wypełnienia Γ_{∞} :

$$\Gamma_{\infty} = b_{\infty} + (1 - b_{\infty}) \frac{\nu_{\infty} \sqrt{b_{\infty}}}{1 + \nu_{\infty} \sqrt{b_{\infty}}} \quad (82)$$

Wzmocnienie niezbędne do wzbudzenia akcji laserowej można w tym wypadku wyznaczyć z warunku:

$$g_{\infty} = \frac{1}{\Gamma_{\infty}} [(1 - \Gamma_{\infty}) \alpha + \alpha_{end}] \quad (83)$$

gdzie przez α_{end} oznaczono straty krawędziowe.

Dla symetrycznej struktury trójwarstwowej, składającej się z warstwy czynnej i dwóch, jednakowych warstw biernych, efektywny współczynnik załamania jest dany następującą zależnością:

$$n_{R,ef} = \sqrt{b_{\infty} n_{R2}^2 + (1 - b_{\infty}) n_{R1}^2} \quad (84)$$

W przypadku struktur o skończonej szerokości S obszaru czynnego, współczynnik ten należy stosować przy analizie efektu falowodowego w płaszczyźnie równoległej do złącza. Model ten został zaproponowany przez J. Buusa w pracy [19]. Zgodnie z nim, analogicznie do znormalizowanej częstotliwości ν_{∞} (zależność (80)), definiujemy boczną znormalizowaną częstotliwość ν_L .

$$\nu_L = k_0 \frac{S}{2} \sqrt{n_{R,ef}^2 - n_{R,3}^2}, \quad (85)$$

Wówczas znormalizowana stała propagacji b_M dla modu bocznego rzędu M ($M = 0, 1, 2, \dots$) wyraża się zależnością analogiczną do wzoru (81):

$$\nu_L \sqrt{1 - b_M} = \arctan[\sqrt{b_M/(1 - b_M)}] + M \frac{\pi}{2} \quad (86)$$

Dla każdego modu bocznego M otrzymujemy stąd inny boczny współczynnik wypełnienia Γ_L (por. zależność (82)):

$$\Gamma_L = b_M + (1 - b_M) \frac{\nu_L \sqrt{b_M}}{1 + \nu_L \sqrt{b_M}} \quad (87)$$

Całkowity współczynnik wypełnienia Γ_M modu M będzie oczywiście równy iloczynowi obu rozpatrywanych współczynników, czyli:

$$\Gamma_M = \Gamma_{\infty} \Gamma_L \quad (88)$$

Wówczas warunek progowy wzbudzenia modu bocznego rzędu M można przedstawić w postaci:

$$\Gamma_M g_M - (1 - \Gamma_M) \alpha = \alpha_{end} \quad (89)$$

skąd otrzymujemy warunek na progowe wzmocnienie w obszarze czynnym niezbędne dla wzbudzenia modu M :

$$g_M = \frac{1}{\Gamma_M} [(1 - \Gamma_M) \alpha + \alpha_{\text{end}}]. \quad (90)$$

Ponieważ $\Gamma_L < 1$, czyli $\Gamma_M < \Gamma_\infty$, więc zawsze wzmocnienie progowe g_M jest większe od g_∞ .

Sformułujemy teraz za J. Buusem [19] kryteria właściwego doboru parametrów struktury lasera z zagrzebanym obszarem czynnym. Są one następujące:

1) Projektowany laser powinien cechować się niską gęstością prądu progowego. Można to osiągnąć zwiększając efektywność ograniczenia promieniowania w obszarze czynnym. Przyjmujemy więc, że szerokość S tego obszaru jest na tyle duża, że progowe wzmocnienie podstawowego ($M = 0$) modu bocznego, czyli g_0 , spełnia warunek:

$$g_0 < 1,2g_\infty. \quad (91)$$

2) Aby uniknąć wzbudzenia modów bocznych wyższych rzędów ($M > 0$), ich progowe wzmocnienia powinny być wystarczająco duże:

$$g_1 > 1,2g_\infty, \quad (92)$$

gdzie g_1 — progowe wzmocnienie modu bocznego pierwszego rzędu ($M = 1$).

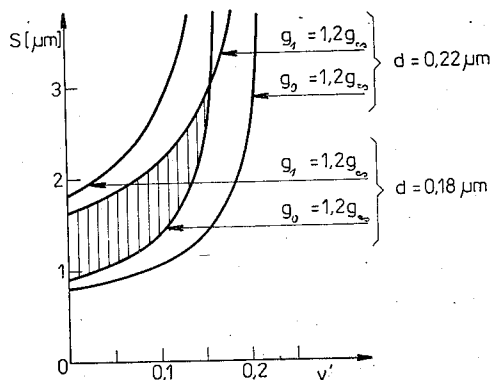
3) Powyższe dwa kryteria powinny być spełnione również dla struktur nieco różniących się od struktury wzorcowej. Przyjmujemy więc typową wartość grubości obszaru czynnego $d = 0,2 \mu\text{m}$ i żądamy, aby oba kryteria były spełnione przy 10% zmianach tego parametru, a więc dla $0,18 \mu\text{m} < d < 0,22 \mu\text{m}$.

W charakterze przykładu rozpatrzmy laser GaInAsP, w którym warstwa czynna została wykonana z $\text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y}$, obie warstwy ograniczające z InP, a warstwy boczne z $\text{Ga}_x, \text{In}_{1-x}, \text{As}_y, \text{P}_{1-y}$.

Wówczas przyjmując:

$$n_{R1} = 3,20; \quad n_{R2} = 3,52 \quad (\text{dla } y = 0,62) \quad (93)$$

$$n_{R3} = n_{R1} + (n_{R2} - n_{R1}) \frac{y'}{y}$$



Rys. 7. Ilustracja właściwego doboru (obszar zakreskowany) parametrów konstrukcyjnych rezonatora lasera GaInAsP z zagrzebanym obszarem czynnym

oraz $\alpha = 15 \text{ cm}^{-1}$ i $\alpha_{\text{end}} = 30 \text{ cm}^{-1}$, otrzymujemy wyniki pokazane na rys. 7. Obszar zakreskowany pokazuje zakres zmian szerokości S obszaru czynnego w funkcji ułamka molowego y' w warstwach bocznych zapewniający spełnienie kryteriów 1–3 wymienionych powyżej. Rysunek ten może być więc użyteczny przy projektowaniu konstrukcji laserów GaInAsP z zagrzebanym obszarem czynnym.

7. PODSUMOWANIE

W pracy przedstawiono analizę własności falowodowych paskowych laserów złączowych. Pozwala ona badać efekt falowodowy w laserach o różnych konstrukcjach, ponieważ uwzględniono w niej zarówno zmianę grubości obszaru czynnego w miarę oddalania się od jego centrum, jak i ewentualną obecność wbudowanego falowodu (gdy kształt funkcji $f(y)$ w równaniu (60) jest narzucony konstrukcyjnie).

Osobno przeanalizowano efekt falowodowy w laserach z zagrzebanym obszarem czynnym. W tym przypadku, zarówno w płaszczyźnie prostopadłej do złącza, jak i w płaszczyźnie złącza, zdecydowanie dominuje rzeczywisty efekt falowodowy wynikający z istoty konstrukcji lasera, stąd więc pominięto całkowicie w tej analizie wzmocnieniowy efekt falowodowy.

Opisana w pracy metoda efektywnego współczynnika załamania jest (zgodnie ze swą nazwą) niezwykle efektywna i pozwala w stosunkowo łatwy sposób analizować ograniczenie pola elektromagnetycznego w falowodzie aktywnym lasera paskowego. Jest to jednakże metoda przybliżona. W przypadku, gdy jej dokładność okazuje się niewystarczająca, należy wrócić do tradycyjnej metody analizy pola elektromagnetycznego w falowodach dielektrycznych przedstawionej np. w monografiach [42, 50–56]. Pozycje [53 i 56] są tłumaczone na język rosyjski, pozycja [54] jest dostępna w języku polskim.

DODATEK

MODY PROMIENIOWANIA W PLANARNYM FALOWODZIE AKTYWNYM [44]

Rozpatrzmy planarny falowód aktywny pokazany na rys. 2. Przyjmijmy, że rozkład pola elektromagnetycznego wewnątrz warstwy centralnej falowodu, tj. dla $|x| \leq d/2$, ma następującą postać określoną parametrem u :

$$E(x) \sim \frac{\sin}{\cos} \left(u \frac{x}{d/2} \right) \quad |x| \leq d/2. \quad (\text{A1})$$

Analogicznie pole w warstwach ograniczających ($d/2 \leq |x|$) będzie zależne od parametru w :

$$E(x) \sim \exp \left(-w \frac{|x|}{d/2} \right) \quad d/2 \leq |x|. \quad (\text{A2})$$

Oba parametry u i w są związane wzorem:

$$u^2 + w^2 = [(1/2)k_0 d]^2 (N_{R2}^2 - N_{R1}^2). \quad (\text{A3})$$

Równania własne dla modów TE i TM wyznacza się z warunku brzegowego dla $x = d/2$ dla pól odpowiednio magnetycznego i elektrycznego (por. [41] rozdz. 4.2.). Są one następującej postaci:

$$w = u \tan u \quad \text{Mody TE} \quad (\text{A4})$$

$$(N_{R2}/N_{R1})w = u \tan u \quad \text{Mody TM.} \quad (\text{A5})$$

Formalnie związki te są identyczne z zależnościami otrzymanymi dla falowodu biernego, z tym że obecnie, ponieważ współczynniki załamania N_{R1} i N_{R2} są zespolone, to również rozwiązania tych równań będą zespolone.

Wprowadźmy nowe wielkości: v i b zwane odpowiednio znormalizowaną częstotliwością (43) i znormalizowaną stałą propagacji (38). Wiążą się one z parametrami u i w następująco:

$$u = v(1-b)^{1/2} \quad (\text{A6})$$

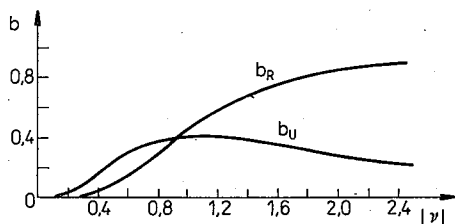
$$w = vb^{1/2} \quad (\text{A7})$$

$$u^2 + w^2 = v^2. \quad (\text{A8})$$

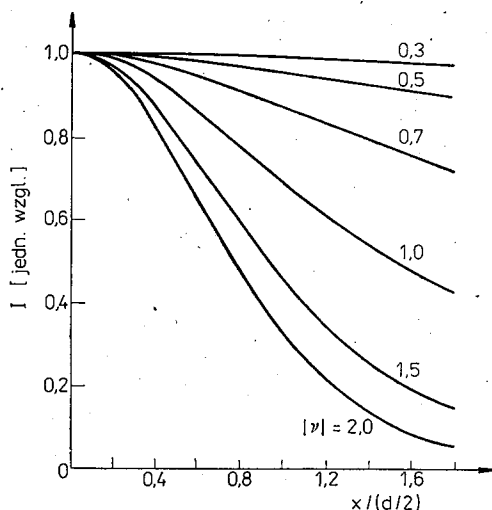
Wówczas równania (A4) i (A5) zredukują się do jednego równania charakterystycznego:

$$v(1-b)^{1/2} = \arctan[b/(1-b)]^{1/2}. \quad (\text{A9})$$

Równanie to nie doczekało się dotąd rozwiązania analitycznego. Rozwiązanie numeryczne dla modu podstawowego, otrzymane za pomocą komputera, przedstawiono na rys. A1., a rozkłady pola promieniowania dla różnych wartości — na rys. A2.



Rys. A1. Zależność obu części: rzeczywistej (b_R) i urojonej (b_U) znormalizowanej stałej propagacji modu podstawowego od znormalizowanej częstotliwości v dla przypadku $n_{R1} = n_{R2}$, tj. dla czystego falowodu wzmacnieniowego [44]



Rys. A2. Rozkłady (w poprzek falowodu) emitancji energetycznej I promieniowania lasera dla różnych wartości znormalizowanej częstotliwości [44]

Warunek odcięcia modu odpowiada przypadkowi, dla którego parametr w staje się urojony, tj.:

$$\operatorname{Re}(w) = 0. \quad (\text{A10})$$

Wówczas pole nie zanika w warstwach ograniczających [por. równanie (A2)]. Powyższa zależność jest równoważna równości:

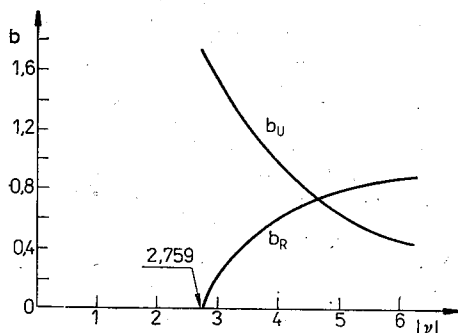
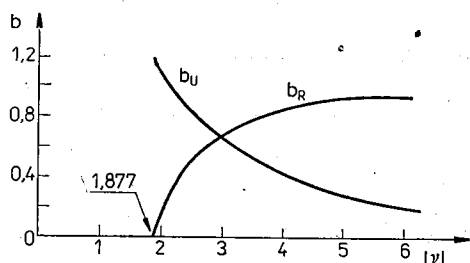
$$\operatorname{Re}(b) = 0, \quad (\text{A11})$$

z której warunki odcięcia, otrzymane również za pomocą komputera, dla modów pierwszego i drugiego rzędu są odpowiednio postaci:

$$|\nu_1| = 1.877; \quad (\text{A12})$$

$$|\nu_2| = 2.759 \quad (\text{A13})$$

Na rys. rys. A3 i A4 wykreślono zależności znormalizowanej stałej propagacji b od znormalizowanej częstotliwości ν odpowiednio dla modów pierwszego i drugiego rzędu.



Rys. A3. Zależność obu części: rzeczywistej (b_R) i urojonej (b_u) znormalizowanej stałej propagacji od znormalizowanej częstotliwości ν dla modu pierwszego rzędu [44]

Rys. A4. Zależność obu części: rzeczywistej (b_R) i urojonej (b_u) znormalizowanej stałej propagacji od znormalizowanej częstotliwości ν dla modu drugiego rzędu [44]

BIBLIOGRAFIA

1. W. Nakwaski: *Samoistny efekt falowodowy w laserach paskowych bez wbudowanego falowodu wewnętrznego*. Rozprawy Elektrotechniczne, 1984, vol. 30, ss. 245—260
2. W. Nakwaski: *Charakterystyki emisyjne laserów paskowych z samoistnym efektem falowodowym*. Rozprawy Elektrotechniczne, 1984, vol. 30, ss. 261—285
3. G. B. Hocker, W. K. Burns: *Mode dispersion in diffused channel waveguides by the effective index method*, Appl. Opt., 1977, vol. 16, ss. 113—118
4. T. L. Paoli: *Waveguiding in a stripe-geometry junction laser*, IEEE J. Quantum Electron., 1977, vol. QE-13, ss. 662—668
5. D. D. Cook, F. R. Nash: *Gain-induced guiding and astigmatic output beam of GaAs lasers*, J. Appl. Phys. 1975, vol. 46, ss. 1660—1672
6. W. Streifer, E. Kapon: *Application of the equivalent-index method to DH diode lasers*. Appl. Opt., 1979, vol. 18, ss. 3724—3725
7. J. Buus: *Detailed field model for DH stripe lasers*. Optical and Quantum Electron., 1978, vol. 10, ss. 459—474

8. J. Buus, *The theory for the dielectric slab waveguide with complex refractive index applied to GaAs lasers*. Proc. 7th European Micro-Wave Conf. 1977, ss. 29—33
9. J. Buus, M. Danielsen: *Carrier diffusion and higher order transversal modes in spectral dynamics of the semiconductor laser*. IEEE J. Quantum Electron., 1977, vol. QE-13, ss. 669—674
10. J. Buus: *Multimode field theory explanation of kinks in the characteristics of DH lasers*. Electron. Lett., 1978, vol. 14, ss. 127—128
11. J. Buus: *A model for the static properties of DH lasers*. IEEE J. Quantum Electron., 1979, vol. QE-15, ss. 734—739
12. J. Buus, K. Stubkjaer: *Characteristic and near field for a broad SiO₂ insulated d.h. laser: Experiment and theory*. IEE Solid-State and Electron Devices, 1979, vol. 3, ss. 210—214
13. J. Buus: *A theoretical investigation of semiconductor lasers*, Ph. D. thesis, Electromagnetics Institute, Technical University of Denmark, Lyngby, 1979, LD 39
14. J. Buus, M. J. Adams: *Phase and group indices for double heterostructure lasers*. IEE Solid-State and Electron. Devices, 1979, vol. 3, ss. 189—195
15. J. Buus: *Threshold currents for GaInAsP lasers emitting at 1.55 μm* . Optics Lett., 1980, vol. 5, ss. 211—213
16. J. Buus: *Dispersion of TE modes in slab waveguides with reference to double heterostructure semiconductor lasers*. Appl. Opt., 1980, vol. 19, ss. 1987—1989
17. P. Danielsen, J. Buus, E. Nicolaisen: *Estimation of the index profile in a single-mode fiber by means of farfield and pulse response measurements*. J. Optical Commun., 1981, vol. 2, ss. 20—23
18. J. Buus: *Analytical approximation for the reflectivity of DH lasers*. IEEE J. Quantum Electron., 1981, vol. QE-17, ss. 2256—2257
19. J. Buus: *Buried heterostructure lasers in the GaInAsP system: Design*. Appl. Opt., 1981, vol. 20, ss. 1884—1885
20. J. Buus: *A systematic design method for lasers with built-in optical guiding*. Proc. 7th European Conf. on Optical Commun., Copenhagen, 1981, P. 19, ss. 80—83
21. J. Buus: Comment on „Equivalent network analysis of dielectric thin-film waveguide with trapezoidal cross section”. Electron. Lett., 1981, vol. 17, s. 497
22. J. Buus: *Beamwidth for asymmetric and multilayer semiconductor laser structures*. IEEE J. Quantum Electron., 1981, vol. QE-17, ss. 732—736
23. J. Buus: *Modelling the behaviour of stripe geometry lasers*. Conf. on Semiconductor Injection Lasers and their Applications, SILA'82, Cardiff 1982, Paper 16
24. R. G. Plumb, J. Buus: *Effects of cavity length on 20 μm stripe laser waveguiding*. IEE Proc., 1982, vol. 129, Pt. I (Solid. State and Electron Devices), ss. 301—305
25. J. Buus: *The effective index method and its application to semiconductor lasers*. IEEE J. Quantum Electron., 1982, vol. QE-18, ss. 1083—1089
26. J. Buus: *Theory for self-sustained pulsations in semiconductor lasers*. Electron. Lett., 1982, vol. 18, ss. 718—719
27. J. Buus: *Models of the static and dynamic behavior of stripe geometry lasers*. IEEE J. Quantum Electron., 1983, vol. QE-19, ss. 953—960
28. J. Buus: *Comparison of two recent semiconductor laser models*. IEEE J. Quantum Electron., 1983, vol. QE-19, ss. 1356—1358
29. M. J. Adams, J. Buus: *Two-segment cavity theory for the mode selection in semiconductor lasers*. IEEE J. Quantum Electron., 1984, vol. QE-20, ss. 99—103
30. J. Buus, M. J. Adams: *Influence of gap width on mode selectivity of cleaved-coupled lasers*. Electron. Lett., 1984, vol. 20, ss. 579—580
31. J. Buus: *Application of the effective index method to nonplanar structures*. IEEE J. Quantum Electron., 1984, vol. QE-20, ss. 1106—1109
32. J. Buus: *Principles of semiconductor laser modelling*. IEE Proc., 1985, vol. 132, Pt. J. Optoelectronics, ss. 42—51

33. J. Buus: *Dynamic line broadening of semiconductor lasers modulated at high frequencies*. Electron. Lett., 1985, vol. 21, ss. 129—131
34. J. Buus: *Mode selectivity in DFB lasers with cleaved facets*. Electron. Lett., 1985, vol. 21, ss. 179—180
35. J. Buus, R. Plastow: *A theoretical and experimental investigation of Fabry-Perot semiconductor laser amplifiers*. IEEE J. Quantum Electron., 1985, vol. QE-21, ss. 614—618
36. J. Buus: *Linewidth enhancement, frequency modulation and instabilities of gain-guided semiconductor lasers*. Electron. Lett., 1985, vol. 21, ss. 1115—1116
37. J. Buus: *Propagation of chirped semiconductor laser pulses in monomode fibers*. Appl. Opt., 1985, vol. 24, ss. 4196—4198
38. J. Buus: *Investigation of semiconductor lasers and waveguiding in active and passive dielectric structures*. Doctor Technices Thesis, Electromagnetics Institute, Technical University of Denmark, Lyngby 1985
39. M. Osiński, J. Buus: *Linewidth broadening factor in semiconductor lasers*. 10th IEEE International Semiconductor Laser Conference, Kanazawa 1986, C-6
40. W. Nakwaski: *Modelowanie zjawisk fizycznych w nowoczesnych laserach złączowych*. Rozprawy Elektrotechniczne, 1988, vol. 34, nr 3, ss. 965—970
41. B. Mroziwicz, M. Bugajski, W. Nakwaski: *Lasery Półprzewodnikowe*. Warszawa: PWN, 1985
42. D. Marcuse: *Light Transmission Optics*. Van Nostrand Reinhold Co., New York, 1972
43. D. Gloge: *Weakly guiding fibres*. Appl. Opt., 1971, vol. 10, ss. 2252—2258
44. W. O. Schlosser: *Gain-induced modes in planar structures*. BSTJ, 1973, vol. 52, ss. 887—905
45. D. Botez: *Analytical approximation of the radiation confinement factor for the TE_0 mode of a double heterojunction laser*, IEEE J. Quantum Electron., 1978, vol. QE-14, ss. 230—232
46. K. L. Chen, S. Wang: *An approximate expression for the effective refractive index in symmetric DH lasers*. IEEE J. Quantum Electron., 1983, vol. QE-19, ss. 1354—1356
47. W. Streifer, R. D. Burnham, D. R. Scifres: *Symmetrical and asymmetrical waveguiding in very narrow conducting stripe lasers*. IEEE J. Quantum Electron., 1979, vol. QE-15, ss. 136—141
48. H. S. Sommers, Jr: *Experimental properties of injection lasers: VIII. Narrow stripe lasers with induced waveguide*. J. Appl. Phys. 1979, vol. 50, ss. 6630—6642
49. H. C. Casey, Jr, M. B. Panish: *Heterostructure Lasers*. Academic Press, New York 1978
50. E. A. J. Marcatili: *Dielectric rectangular waveguide and directional coupler for integrated optics*. Bell Syst. Techn. J., 1969, vol. 48, ss. 2071—2102
51. D. Marcuse: *Theory of Dielectric Optical Waveguides*. Academic Press, New York, London, 1974
52. M. K. Barnoski (Ed.): *Introduction to Integrated Optics*. Plenum Press, New York, London, 1974
53. H. Kogelnik: *Theory of dielectric waveguides in integrated optics*. [w] Integrated Optics, ed. by T. Tamir, Springer Verlag, Heidelberg, New York, 1975, chapter 2
54. H.-G. Unger: *Planar Optical Waveguides and Fibres*. Clarendon Press, Oxford, 1977
55. M. J. Adams: *An Introduction to Optical Waveguides*, J. Wiley and Sons, Chichester, New York, Brisbane Toronto, 1981
56. R. G. Hunsperger: *Integrated Optics: Theory and Technology*. Springer Verlag, Berlin, Heidelberg, New York, Tokyo, 1984

W. NAKWASKI

WAVEGUIDING EFFECT IN STRIPE-GEOMETRY FABRY-PEROT LASER DIODES

Summary

A theoretical analysis of the waveguiding effect in stripe lasers of various design, i.e. planar structure, nonplanar structure and buried-heterostructure, is presented.

W. NAKWASKI

EFFET DE GUIDE D'ONDES DANS LES LASERS LONGITUDINAUX
AVEC RÉSONATEURS DE FABRY — PEROT

Résumé

L'étude ci-présentée contient une analyse théorique de l'effet de guide d'ondes dans les laser longitudinaux à constructions diverses: à procédé planar, non-planar et à zone active enfouie (bury-in).

W. NAKWASKI

WELLENLEITEREFFEKT IN STREIFENDIODENLASERN MIT DEM FABRY-PEROT-
RESONATOR

Zusammenfassung

Im vorliegenden Aufsatz wird die theoretische Analyse der Wellenleitereffekte in Streifendiodenlasern verschiedener Konstruktionen geschildert: Planarstruktur, Nichtplanarstruktur und vergrabene Doppel-Heterostruktur.

В. НАКВАСКИ

ВОЛНОВОДНЫЙ ЭФФЕКТ В ПОЛОСКОВЫХ ЛАЗЕРНЫХ ДИОДАХ
С РЕЗОНАТОРОМ ФАБРИ-ПЕРО

Резюме

Представлен теоретический анализ волноводного эффекта в полосковых лазерных диодах с разнотипными конструкциями планарной структуры, непланарной структуры и зарытой мезаполоской.

Wpływ czynników konstrukcyjnych i technologicznych na temperaturowy współczynnik rezystancji scalonych sieci rezystywnych grubowarstwowych

STANISŁAW NOWAK

Akademia Górniczo-Hutnicza, Kraków

Otrzymano 1988.06.20

Autoryzowano do druku 1989.02.03

Przeanalizowano ważniejsze wpływy konstrukcji i technologii na temperaturowy współczynnik rezystancji (TWR) rezystywnych sieci grubowarstwowych. Dokonano optymalizacji procesu wytwarzania i zaprezentowano metodę wyboru warunków wypalania warstwy ze względu na minimalną wartość bezwzględną TWR. Uzasadniono możliwość kompensacji termicznej zintegrowanych dzielników i tłumików wykonanych technikami: grubowarstwową i cienkowarstwową. Ponadto dokonano przeglądu opisów matematycznych mechanizmu przewodzenia warstw grubych i cienkich w aspekcie zależności temperaturowej oraz wskazano na niedoskonałość tych opisów.

1. WSTĘP

Z rozwojem scalonych wzmacniaczy operacyjnych oraz przetworników C/A i A/C wzrasta zainteresowanie stabilnymi rezystorami, a szczególnie stabilnymi sieciami rezystywnymi skompensowanymi termicznie. O niestabilności rezystora decydują głównie zmiany długoterminowe, starzeniowe, nieodwracalne oraz zmiany termiczne odwracalne, wywołane zmianami temperatury otoczenia. Zmiany termiczne rezystorów i sieci rezystywnych są wyrażane przez temperaturowy współczynnik rezystancji (TWR).

Celem pracy jest analiza czynników konstrukcyjnych, technologicznych i fizycznych na wartość i znak TWR oraz próby jego minimalizacji. Ponadto dyskutuje się niedoskonałość teoretycznego określenia TWR w rezystorach cienkowarstwowych i grubowarstwowych. Przedmiotem badań są rezystory w sieciach rezystywnych.

W procesie wytwarzania scalonych sieci rezystywnych rozpowszechniły się dwie technologie:

a) cienkowarstwowa — rezystory powstają przez naparowanie cienkiej warstwy stopu rezystywnego na podłożu szklanym,

b) grubowarstwowa — rezystory powstają przez naniesienie sitodrukiem odpowiedniej kompozycji na podłoże z ceramiki alundowej i wypalenie naniesionej warstwy w temperaturze od 950 K do 1200 K.

2. PROBLEMY Z MATEMATYCZNYM OPISEM ZALEŻNOŚCI REZYSTANCJI OD TEMPERATURY

Zmiany termiczne rezystorów warstwowych są trudne do ujęcia teoretycznego i dotąd nie udało się analitycznie określić zależności rezystancji od temperatury otoczenia ani warstw cienkich wykonanych, ze stopów metali, ani tym bardziej warstw grubych wykonanych z kompozycji ziaren o przewodnictwie metalicznym zawieszonych w izolacyjnym szkle. Ze względu na komplementarne stosowanie sieci rezystywnych grubowarstwowych i cienkowarstwowych oraz na ich podobieństwa i różnice zostaną przedstawione próby opisu matematycznego obu warstw rezystywnych: cienkiej i grubej.

2.1. OPIŚ WARSTW CIENKICH

Wychodząc z elektronowej teorii metali i przyjmując, że elektrony swobodne podlegają statystyce Fermiego-Diraca, Grüneisenn [6, 41] określił zależność temperaturową dla metali niemagnetycznych:

$$\text{TWR} = \frac{1}{293 - 0,15\Theta} \quad \text{dla } T \geq 0,15\Theta \quad (1)$$

gdzie Θ — temperatura Debye'a — jest określona wzorem

$$\Theta = \frac{h\nu_{\max}}{k} \quad (2)$$

h — stała Plancka = $6,62517 \cdot 10^{-34}$ Js

k — stała Boltzmanna = $1,38044 \cdot 10^{-23}$ J/K

$\nu_{\max}$ — maksymalna częstotliwość drgań sprężystych siatki krystalicznej danego ciała.

Powyżej temperatury granicznej $0,15 \Theta$ (niektórzy autorzy przyjmują całkowitą wartość Θ), a więc w zakresie temperatur eksploatacji rezystora, wzór (1) dla metali niemagnetycznych obowiązuje, a TWR jest stały dodatni i wynosi 3500 do 4300 ppm/K dla różnych metali, natomiast zależność rezystancji od temperatury jest określona wzorem (3)

$$R(T) = R_{m0}(1 + \alpha T) \quad (3)$$

gdzie $\alpha = \text{TWR}$, R_{m0} — rezystancja zmierzona w temperaturze $+20^\circ\text{C}$. Poniżej temperatury granicznej $0,15 \Theta$ występują znaczne odchylenia od zależności (3) i w pobliżu zera bezwzględnej rezystancja jest proporcjonalna do piątej potęgi temperatury T^5 a nie wprost proporcjonalna do temperatury, jak we wzorze (3). Wzory (2) i (3) nie obowiązują jednak dla metali magnetycznych, w których zamiast zależności (3) wprowadzono silniejszą niż liniową zależność od temperatury, uwzględniając ograniczenia narzucone temperaturą Curie

$$R(T) = c \cdot T^{1,7} \quad \text{dla } T \leq T_c \quad (4)$$

gdzie T_c — temperatura Curie

c — stała materiałowa.

Różna dla różnych metali temperatura Curie komplikuje problem $T_c = 1043$ K dla Fe; $T_c = 1400$ K dla Co; $T_c = 635$ K dla Ni.

Pomierzone TWR w pobliżu temperatury pokojowej dla metali magnetycznych wynoszą $TWR \approx +4500$ ppm/K, a obliczone wartości sięgają $+5800$ ppm/K.

Wychodząc z równania Boltzmanna ujmującego transport elektronów można by bliżej określić zależności temperaturowe, lecz jego rozwiązanie nastęrcza jak dotąd duże trudności. Istotnym członem jest tzw. całka zderzeń [6]

$$\left(\frac{\partial f}{\partial t} \right)_{zd} = \frac{1}{4\pi^3} \int W(k, k') [f(k') - f(k)] d\Omega_{k'} \quad (5)$$

Po linearyzacji i wprowadzeniu pojęcia czasu relaksacji wyrażenie (5) można uprościć do wyrażenia (6). Słuszność równania (6) jest często krytykowana, lecz z braku lepszego przybliżenia jest ono powszechnie używane:

$$\left(\frac{\partial f}{\partial t} \right)_{zd} = \frac{f - f_0}{\tau} \quad (6)$$

$W(k, k')$ — prawdopodobieństwo prostych przejść elektronów,

k stan przed k' stan po przejściu elektronu,

$f(r, k, t)$ — funkcja rozkładu,

$f_0(r, k)$ — równowagowa funkcja rozkładu,

$d\Omega_{k'}$ — elementarna objętość, wektor k jest zależny od temperatury

$f - f_0$ — odchylenie od równowagi, zmiana funkcji rozkładu,

τ — czas relaksacji, zależny od temperatury.

Następnie określono przewodność właściwą

$$\sigma = \frac{N_{eff} e^2}{m} \tau(k_F) \quad (7)$$

$\tau(k_F)$ — czas relaksacji, stała czasowa elektronu na powierzchni Fermiego,

N_{eff} — ilość elektronów biorących udział w przewodzeniu,

m, e — masa i ładunek elektronu.

Całkowita ilość elektronów przewodnictwa jest dana funkcją gęstości stanów i prawdopodobieństwem zajęcia stanów. Przewodność jest proporcjonalna do czasu relaksacji, natomiast rezystywność odwrotnie proporcjonalna.

Mathiessen [6] podał regułę superpozycji rezystywności przyjmując założenie stałych czasów relaksacji dla rozpraszania elektronów na:

- fononach sieci metalu podstawowego,
- domieszkach,
- defektach sieci.

Ponieważ rezystywność jest proporcjonalną odwrotnością do czasu relaksacji uzyskano zależność zwaną regułą Mathiessena

$$\frac{1}{\tau} = \frac{1}{\tau_s} + \frac{1}{\tau_i} + \frac{1}{\tau_a} \quad (8)$$

i odpowiednio

$$\varrho = \varrho_s + \varrho_i + \varrho_a \quad (9)$$

ϱ_s — rezystywność odpowiadająca drganiom sieci — jest funkcją temperatury

ϱ_i — rezystywność powstała na atomach wtrąconych, które deformują pole wewnątrz sieci krystalicznej — praktycznie nie zależy od temperatury

ϱ_a — rezystywność powstała w wyniku rozpraszania elektronów na defektach sieci.

Reguła (9) wskazuje, że rezystywność stopu jest zawsze większa od rezystywności czystego metalu. Ponadto gdy dobozem domieszek spełni się warunek (10),

$$\varrho_i \gg \varrho_s \quad (10)$$

to wypadkowa rezystywność bardzo słabo zależy od temperatury i jest opisywana wzorem (3). Reguła (9) w przybliżeniu opisuje interesujące właściwości stopów np. stop konstantan o zawartości 60% Cu + 40% Ni ma $\varrho_{20} = 0,49$ ($\mu\Omega\text{m}$ i TWR = +5 ppm/K, natomiast stop kanthal o składzie 70% Fe, 23% Cr, 4,5% Al, 1% Co i 1,5% innych domieszek wykazuje $\varrho_{20} = 1,4/\mu\Omega\text{m}$ i TWR = +50 ppm/K. Wielu autorów krytykuje przyjmowanie założenia o stałości czasów relaksacji, lecz jak dotąd nie udało się rozwiązać równania Boltzmanna bez daleko idących założeń upraszczających. Otrzymane wyniki przy tych założeniach są jednak użyteczne. Rezystywność metalu zanieczyszczonego małą ilością domieszek wg Zimanna [40] jest proporcjonalna do koncentracji domieszek, natomiast przy dużych stężeniach składników stopu obowiązuje wzór Nordheima [28]

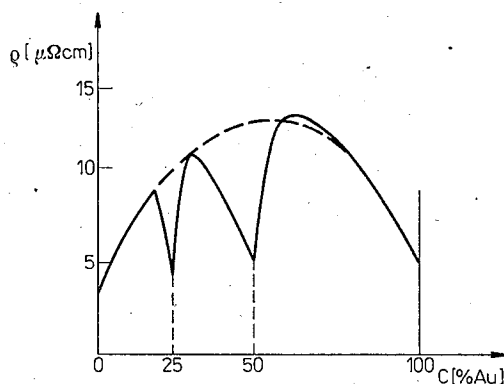
$$\varrho = k \cdot c(1-c) \quad (11)$$

gdzie c — stężenie jednego pierwiastka

$1-c$ — stężenie drugiego pierwiastka

k — współczynnik proporcjonalności.

W zależności tej występuje jednak anomalia, co przedstawia rys. 1. Gwałtowne spadki rezystywności na rys. 1 odpowiadają tworzeniu odpowiednich związków AuCu i AuCu₃ [6].



Rys. 1. Zależność rezystywności ϱ od stężenia c . Stężenie c określa zawartość złota w miedzi

Fuchs [10] i Sondheimer [36] wychodząc z czasu relaksacji określili TWR warstw bardzo cienkich w stosunku do TWR₀ stopu litego, z którego wykonano warstwę cieką. Przy obliczeniach uwzględniono, że ilość elektronów elastycznie odbijanych od dolnej i górnej ściany warstwy wynosi p , natomiast ilość elektronów pochłanianych przez obie

szorstkie ściany wynosi $1-p$. Ponadto przyjęto unormowaną grubość warstwy $\kappa = \frac{t}{\lambda}$

gdzie t — grubość warstwy,

λ — średnia droga swobodna elektronu w warstwie.

Wtedy, uwzględniając powyższe założenia, otrzymano wzory (12) i (13), które opisują efekt zwany efektem Fuchsa-Sondheimera

$$\frac{\text{TWR}}{\text{TWR}_0} = 1 - \frac{3(1-p)}{8\kappa} \quad \text{dla } \kappa \gg 1 \quad (12)$$

$$\frac{\text{TWR}}{\text{TWR}_0} = \frac{1}{\ln \frac{1}{\kappa} + 0,423} \quad \text{dla } \kappa \ll 1 \quad (13)$$

W przypadku zmniejszenia średnicy ziaren analogicznie zachodzi odbijanie i pochłanianie na granicach ziaren, co było przedmiotem prac Mayadasa i Shatzkes'a [24]. Przy bardzo małych ziarnach efekt rozpraszania i pochłaniania na granicach ziaren może być silniejszy od efektu Fuchsa-Sondheimera.

2.2. OPIS WARSTW GRUBYCH

Gruba warstwa rezystywna jest zbudowana z ziaren o przewodnictwie metalicznym, zawieszonych w izolacyjnym szklivię. Zależność temperaturowa rezystancji ziaren metalicznych jest opisana równaniem (3) i z tym opisem zgadzają się wszyscy autorzy analizujący mechanizm przewodnictwa warstw grubych. W ostatnich 15 latach prowadzi się intensywne prace nad wyjaśnieniem przewodnictwa i minimalizacji TWR. Zainteresowanie naukowe tymi problemami uzasadniają zalety warstw grubych w stosunku do cienkich. Z niezaprzeczalnych zalet warstw grubych przeważających nad cienkimi są: duża obciążalność warstwy około 15 W/cm², szeroki zakres rezystancji od 0,5 Ω do 500 MΩ, łatwość produkcji. Do zalet należy również zaliczyć: stabilność długoterminową i TWR, które są zbliżone do odnośnych parametrów warstwy cienkiej.

Rozwój opisu warstwy grubej przebiega w dwóch kierunkach:

1) zastosowania teorii perkolacji, która sprowadza zagadnienie do ilości efektywnych połączeń przewodzących między elektrodami rezystora; rezystor o małej rezystancji ma dużo połączeń ścieżek przewodzących, natomiast rezystor o dużej rezystancji ma mało połączeń;

2) analizy struktury fizycznej i mechanizmu przewodzenia wewnątrz połączenia, czyli wewnątrz ścieżki, przez którą płyną elektrony. Do pierwszej grupy prac należy zaliczyć prace z 1981 r. Evena i Robertsona [9], w której autorzy dążyli do uzyskania jak największej ilości ścieżek przewodzących przy minimalnej zawartości fazy przewodzącej V (V — stosunek objętości fazy przewodzącej do całej objętości warstwy). Dodając do pasty kulki szklane o temperaturze topnienia wyższej niż temperatura obróbki pasty, zredukowali V_{gr} (próg perkolacyjny) z 16% do 6%.

Optymalizację komputerową prowadzili wg wzoru (14).

$$\varrho = \frac{\varrho_c}{\left\{ \left[\left(\frac{V}{f_m(1-f_g)} \right)^n + p_{kr}^n \right]^{1/n} - p_{kr} \right\}^\beta} \quad (14)$$

ϱ_c — rezystywność cząstki przewodzącej

$$p_{kr} = \frac{V_{kr}}{f} \text{ — prawdopodobieństwo krytyczne}$$

f — współczynnik maksymalnego upakowania kul

β — wykładnik krytyczny, przyjęta wartość $\beta = 1,7$

f_g — współczynnik wypełnienia kulek szklanych

f_m — współczynnik wypełnienia cząstek przewodzących

n — wykładnik mówiący o uporządkowaniu kul.

Gdy $n = \infty$ — panuje chaos, gdy n jest małe, rzędu kilka — następuje uporządkowanie. Z pracy tej wynika, że jeśli chce się uzyskać stosunkowo małą i regulowaną rezystywność przy jak najmniejszej zawartości fazy metalicznej (mały ułamek wartości V_{kr}) to n musi być małe i f_g duże. Narzuca to uporządkowanie wewnątrz warstwy. Problemem uporządkowania zajmował się A. Kusy [19], który do wzoru perkolacyjnego na rezystywność

$$\varrho = \frac{\varrho_c}{V \cdot V^{-2/3} \cdot C^2(V)} \quad (15)$$

wprowadził funkcję uporządkowania $\psi(k, m)$

$$\psi(k, m) = 3 \frac{(k \cdot m)^{2/3}}{1 + k^2 + m^2} \quad (16)$$

modyfikując (15) do postaci (17)

$$\varrho = \frac{\varrho_c}{V \cdot V^{\psi(k, m)} V^{-2/3} \cdot C^2(V)} \quad (17)$$

gdzie k, m — stałe dodatnie

$C^2(V)$ — pole podstawy prostopadłościanów utworzonych z łańcuchów przewodzących jednostkowego sześciangu.

W granicy dla sześciangu i kuli $k = m = 1$ wartość funkcji ψ wynosi 1. Odchylenie kształtu cząstki przewodzącej do tych brył prowadzi do zmniejszenia ψ , czyli do zwiększenia uporządkowania w warstwie. W praktyce R. Vest [39] uzyskał krytyczny ułamek objętościowy $V_{kr} = 2\%$ i na tej bazie produkuje się rezystory o rezystancjach powierzchniowych $1 \text{ M}\Omega/\square \div 10 \text{ M}\Omega/\square$, wykazujące dobre właściwości eksploatacyjne. O ile to zagadnienie można uznać za opisane z dość dużą dokładnością, to opis struktury fizycznej gałęzi przewodzącej jest jedynie jakościowy.

Największy wkład w ten opis należy przypisać Pike'owi i Seager'owi którzy szczególnie w swej pracy [35] odrzucili 3 modele opisywane wcześniej przez innych autorów, a mianowicie:

- model jednorodnego półprzewodnika,
- model ciągłych łańcuchów przewodzących,
- model barier nietunelowych,

a zaproponowali model łańcuchów z rezonansowym tunelowaniem przez barierę. Polega on na tym, że każda cząstka jest otoczona warstewką szkliva i łańcuchy są utworzone z MIM-ów Metal-Isolator-Metal; przez metal rozumie się cząstkę rutenianu lub cząstkę tlenku ruteniu o przewodnictwie metalicznym. Główna część rezystancji łańcucha jest sumą re-

zystancji barier występujących w tym łańcuchu. Skorzystali oni ze wzorów Strattona [37] i Duke'a [7] na przewodność MIM

$$\sigma(U, T) = \frac{1}{U} \int D(Ex, U) N(Ex, U, T) dEx, \quad (18)$$

w którym: U — napięcie na złączu

Ex — składowa normalna energii elektronu

T — temperatura absolutna

$D(Ex, U)$ — funkcja transmisji, która opisuje prawdopodobieństwo, że elektron o energii Ex przejdzie drogą tunelowania z jednej strony bariery na drugą

$N(Ex, U, T)$ — funkcja zasilania, opisuje ilość elektronów w elektrodzie, które są zdolne wykonać przejście tunelujące.

Ponieważ powyższy wzór obowiązuje dla dostatecznie dużych płaskich elektrod, dlatego autorzy wprowadzili jego modyfikacje. Elektrody są bardzo małe i kuliste; są to ziarna przewodzące o średnicy około $0,5 \mu\text{m}$. Do przeniesienia ładunku elektronowego z jednej kulki na drugą, potrzebna jest dodatkowa energia. Zgodnie z prawami elektrostatyki energia ta [35] wynosi

$$Ec = \frac{1}{2} \frac{Q}{C} = \frac{e^2}{4\pi\epsilon\epsilon_0 \cdot rK(r, d)}, \quad (19)$$

gdzie: $K(r, d)$ — funkcja słabo zależna od r i d , dla warstw rutenowych wynosi około 3,

r — promień kulki przyjęto $0,2 \mu\text{m}$,

d — odstęp między kulkami = szerokość bariery, rzędu 2 do 4 nm,

ϵ — stała dielektryczna szkliva.

Po wstawieniu danych liczbowych uzyskano energię aktywacji $Ec = 340 \mu\text{eV}$, która według eksperymentu [35] zmieniała się od $60 \mu\text{eV}$ dla małych rezystywności, do $800 \mu\text{eV}$ dla dużych rezystywności, co przyjęto za zadawalającą zgodność. Należy rozumieć, że praktycznie średnice kul i odstęp między kulkami nie są stałe i należy się liczyć z pewnym pasmem energii E_c .

Analiza handlowych past wykazała, że w szklwie występują zanieczyszczenia typu Nb, Zr, Zn, Mn, Cd, Sn. Ponadto domieszką może być ruten, który rozpuszcza się w szklwie z ostrych krawędzi cząstek przewodzących. Mechanizm działania tych domieszek nie jest jeszcze zbadany i autorzy zaproponowali, aby były one centrami tunelowania rezonansowego. Te centra zwiększają prawdopodobieństwo tunelowania dla wszystkich elektronów, ale szczególnie dla tych, których poziom energii jest bliski poziomowi energetycznemu elektronów metali zanieczyszczających, zlokalizowanych wewnątrz szkliva, tworzącego barierę.

Uwzględniając powyższe okoliczności przyjęto, że elektron przechodzi z poziomu Fermiego na poziom o około 10^{-4} eV wyższy $\left(\frac{1}{2} Ec\right)$, a przejście przez barierę wysokości około 3 eV (szklivo) jest ułatwione przez centra rezonansowe wynikające z domieszek. Według Hilla [14] prąd płynący przez barierę zależy zarówno od wysokości bariery Φ jak i jej efektywnej szerokości d i opisany jest wzorem

$$J(V, T) = \frac{8\pi me}{h^3 B^2} \cdot \sinh\left(\frac{eV}{kT}\right) \frac{\pi BkT}{\sin \pi BkT} \cdot \exp(-Md\Phi^{-1/2}) \cdot \exp\left(-\frac{\delta E}{kT}\right) \quad (20)$$

$$\text{gdzie } \delta E = \delta E_0 - gF^{1/2} \quad (21)$$

V, F — napięcie i natężenie pola na przerwie między ziarnami przewodzącymi

B, M, g — stałe, k — stała Boltzmann

δE_0 — energia aktywacji bez pola zewnętrznego

m, e, h — masa, ładunek elektronu i stała Plancka.

Wysokość bariery szkliva jest duża $\Phi \approx 3$ eV i dlatego Pike i Seager [35] przyjęli przewodzenie tylko drogą tzw. tunelowania rezonansowego i pominęli człon zawierający wysokość bariery. Dla tunelowania rezonansowego przewodność wynosi [35]

$$G(T) = G(0) \frac{aT}{\sin aT} \text{ przy } aT < 3. \quad (22)$$

Uwzględniając powyższe i wynik prac Neugebauera i Webba [27] a także Abelesa [1] wzór na R_{MIM} ma postać

$$R_{\text{MIM}} = R_{b0} \frac{\sin aT}{aT} \left(1 + \exp \frac{\Delta E}{kT}\right). \quad (23)$$

Natomiast uwzględniając dane liczbowe dla warstwy grubej $\Delta E/kT \ll 1$, a $\approx (5 \div 10) \cdot 10^{-4}/\text{K}$ a więc $aT \ll 1$, wzór ten może być uproszczony

$$R_{\text{MIM}} = \frac{1}{2} R_{b0} \left(1 - \frac{1}{6} a^2 T^2\right) \left(2 + \frac{\Delta E}{kT}\right). \quad (24)$$

Łatwo zauważyć, że w funkcji temperatury R_{MIM} maleje. W rezultacie można napisać wzór na rezystancję elementarnej gałęzi sieci perkolacyjnej, który ma postać

$$R(T) = R_m(T) + R_{\text{MIM}}(T) \quad (25)$$

$$R(T) = R_{m0}(1 + aT) + \frac{1}{2} R_{b0} \left(1 - \frac{1}{6} a^2 T^2\right) \left(2 + \frac{\Delta E}{kT}\right). \quad (26)$$

Pike i Seager obliczyli T_m dla różnych past w zakresie rezystancji powierzchniowej $10^3 \Omega/\square$ do $10^6 \Omega/\square$. Za wyjątkiem past do potencjometrów serii DP 1200 uzyskali zgodność polegającą na tym, że ze wzrostem rezystancji powierzchniowej $R_\square$ (maleniem ułamka objętościowego V) rosła temperatura T_m . Dla past D-Proto od 212 K do 276 K, dla C-500 od 169 K do 247 K (253 K) [35].

Opis ten znalazł wielu zwolenników, aczkolwiek niektórzy autorzy go rozwijali dalej, np. Halder zaproponował, że obok przewodnictwa tunelowego może być przewodnictwo przeskokowe [11]. Bardziej ogólne określenie podano w pracy [25], że prąd przez barierę może być sumą prądu tunelowego, prądu przeskokowego i termoemisyjnego. Występuje duża komplikacja, gdy ogniwa w których płyną poszczególne prądy występują równoległe lub szeregowo-równoległe. Człon R_m we wzorze (25) rośnie z temperaturą, natomiast człon R_{MIM} ze wzrostem temperatury maleje, a więc zależność $R(T)$ ma ekstremum. Obliczenie tego ekstremum jest łatwe przez zróżniczkowanie

$$\frac{1}{R} \frac{\partial R}{\partial T} = - \frac{R_{b0}}{R_{b0} + R_{m0}} \left[\frac{a^2 T}{3} \left(1 + \frac{\Delta E}{4kT}\right) + \frac{\Delta E}{2kT^2} \right] + \left(1 - \frac{R_{b0}}{R_{b0} + R_{m0}}\right) \alpha. \quad (27)$$

Druga pochodna względem temperatury jest dodatnia, a więc jest to minimum, zaś temperaturę T_m przy której to minimum krzywej $R(T)$ występuje można znaleźć przyrównując wyrażenie (27) do zera; z tego też wyrażenia można obliczyć energię aktywacji E_A wystarczającą do przewodzenia przez barierę przy temperaturze T_m

$$E_A = \frac{2k \left[6\alpha T_m^2 \left(\frac{R_{b0} + R_{m0}}{R_{b0}} - 1 \right) - 2a^2 T_m^3 \right]}{a^2 T_m^2 + 6} \quad (28)$$

Autor niniejszej pracy sądzi, że przy dużych ułamkach objętościowych V tzn. przy rezystancjach powierzchniowych poniżej $10^2 \Omega/\square$ oprócz powyższych mechanizmów wystąpią styki bezpośrednie ziaren, gdyż zawartość szkliva jest tu mała i jest duże prawdopodobieństwo, że ziarna stykają się bezpośrednio, a wtedy należałoby wyjść z teorii Holma [18] jak to czynili Brady [3] i Kusy [19, 20].

Jakościowo to jest dość przekonujące, lecz dla ilościowego obliczenia zarówno T_m jak i wartość TWR należałoby znać udziały poszczególnych składowych w odpowiednich warstwach. Dlatego zrodził się nowy iloczynowy model zaproponowany przez Licznarskiego i współpracowników [22], który jest modelem bardziej statystycznym, niż fizycznym, a udziały poszczególnych składowych podają wykładniki występujące przy poszczególnych członach. Model ten został oparty na teorii Efros'a i Shklovski'ego [8], w której rezystancja systemu metal-szkło jest iloczynem rezystancji metalu R_m i rezystancji szkła R_g według wzoru

$$R_{mg}(T) = [R_m(T)]^{1-s} \cdot [R_g(T)]^s \quad (29)$$

gdzie s jest związane z wykładnikiem β występującym we wzorze (14) poprzez wyrażenie

$$s = \frac{\beta}{\beta + q} < 1. \quad (30)$$

Dla trójwymiarowej sieci przyjmuje się $q = 1$; $\beta = 1,6 \div 1,7$; $s = 0,62 \div 0,63$.

Podstawiając do wzoru (29) wzór (3) i uproszczony wzór (23) otrzymuje się [22] wzór (31)

$$R(T) = R_{m0}^{1-s} R_{g0}^s (1 + \alpha T)^{1-s} \left(\exp \frac{\Delta E}{kT} \right)^s, \quad (31)$$

$$\text{TWR} = \frac{1}{R} \frac{dR}{dT} = \frac{(1-s)\alpha}{1 + \alpha T} - \frac{s\Delta E}{kT^2}. \quad (32)$$

Wzór ten jest stosunkowo prosty, lecz niewystarczający, gdyż jak wiadomo, na granicy ziaren przewodzących występuje warstwa dyfuzyjna domieszkowana (dyfuzja zachodzi podczas spiekania warstwy w temperaturze 850°C) i to zmusza do komplikacji wzoru.

Jeżeli przyjmiemy, że warstwa dyfuzyjna ma charakter półprzewodnikowy, to należałoby uwzględnić $R(T)_{ism}$ (izolator-półprzewodnik-metal) i wtedy wzór na $R(T)_{ism}$ ma postać

$$R(T)_{ism} = R_{m0}^{1-s} R_{s0}^{(1-s)s} R_{s0}^s (1 + \alpha T)^{1-s} \exp \left[\frac{E_s}{kT} (1-s)s \right] \cdot \exp \frac{\Delta E}{kT} s^2 \quad (33)$$

$$\text{TWR} = \frac{1}{R} \frac{dR}{dT} = \frac{(1-s)\alpha}{1+\alpha T} - \frac{(1-s)sE_s}{kT^2} - \frac{s^2\Delta E}{kT^2}. \quad (34)$$

Zdarza się jednak, że równocześnie w klastercie o przewodzeniu decydują dwie ścieżki, w których w jednej decyduje mechanizm tunelowy R_T , w drugiej mechanizm przeskokowy R_H [11]. Wtedy wypadkowa rezystancja tych ścieżek wynosi:

$$R_{TH} = \frac{R_T R_H}{R_T + R_H}. \quad (35)$$

Według teorii Motta [26] R_H wyraża się wzorem

$$R_H = R_{H0} \exp\left(2\alpha_r \cdot r + \frac{E_H}{kT}\right), \quad (36)$$

gdzie E_H — średnia energia aktywacji przeskoków elektronów.

r — odległość między pułapkami

α_r — funkcja z teorii Motta o wymiarze m^{-1} .

Ostatecznie w tym modelu [22] wzór na rezystancję jest dość złożony

$$R_{THM}(T) = (R_{m0})^{1-s} [R_{H0} \exp \alpha_r r]^{1-s} [A \exp(Bd\Phi^{1/2})]^{s^2} \cdot (1 + \alpha T)^{1-s} \left(\frac{\sin \alpha T}{\alpha T} \right)^{s^2} \cdot \exp\left[(1-s) \cdot s \frac{E_H}{kT}\right], \quad (37)$$

w którym Φ — wysokość bariery,

d — skuteczna szerokość bariery,

A, B — stałe (por. wzór (20)).

3. WYBÓR CZYNNIKÓW WPŁYWAJĄCYCH NA TWR REZYSTORÓW W GRUBOWARSTWOWYCH SIECIACH REZYSTYWNYCH

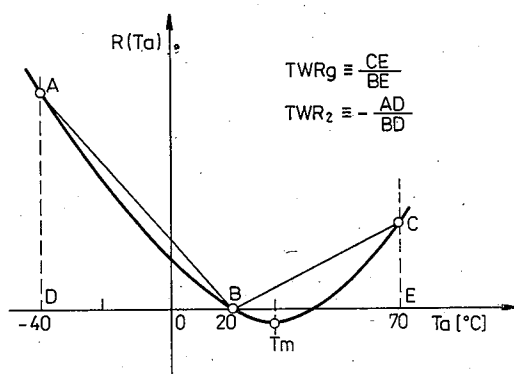
Wartość i znak TWR rezystorów grubowarstwowych w scalonych sieciach rezystywnych zależy od wielu czynników. Ilościowe związki między tymi czynnikami a TWR nie są znane, aczkolwiek istnieje wiele prób ich określenia.

Na podstawie literatury przedmiotu oraz własnych badań można wybrać ważniejsze czynniki, które mają istotny wpływ na TWR. Należą do nich: temperatura otoczenia, grubość i rodzaj warstwy, rodzaj podłoża, naprężenia w warstwie, temperatura wypalania i czas wypalania warstwy. Należy zwrócić uwagę, że wiele czynników konstrukcyjnych i technologicznych ma również wpływ, lecz na podstawie wieloletnich badań i obserwacji*) autor wybiera te najsilniej wpływające. Elementy konstrukcyjne takie jak długość rezystora, rodzaj materiału użytego na elektrody rezystora mają duży wpływ na TWR, ale w niniejszej pracy przyjęto do badań rezystor dostatecznie długi (powyżej 1 mm) i elektrody z materiałów ogólnie stosowanych.

*) Autor uruchomił licencyjną produkcję układów hybrydowych grubowarstwowych, odbył kilka krótkoterminowych szkoleń w zakresie konstrukcji i technologii w przodujących firmach, produkujących grubowarstwowe sieci rezystywne.

3.1. WPŁYW TEMPERATURY OTOCZENIA; TWR „ZIMNY”, TWR „GORĄCY”

Rezystor grubowarstwowy w przeciwieństwie do rezystorów metalowych i cienkowarstwowych wykazuje duże zmiany temperaturowego współczynnika rezystancji w funkcji temperatury otoczenia. W zakresie roboczych temperatur od -65°C do $+155^{\circ}\text{C}$ z reguły zmienia on znak przechodząc przez wartość zero. Fakt ten powoduje trudności w samej



Rys. 2. Sposób określania TWR metodą różnicową. Rezystancja rezystora grubowarstwowego w zależności od temperatury otoczenia

definicji TWR jak również pomiaru jego wartości. Na rys. 2 przedstawiono zależność rezystancji R od temperatury otoczenia T_a rezystora typowej sieci grubowarstwowej. Istnieją dwa sposoby określenia TWR:

a) różniczkowo

$$\text{TWR}_1 = \frac{1}{R_1} \frac{\partial R}{\partial T} \quad (38)$$

gdzie R_1 — rezystancja rezystora zmierzona w temperaturze $T_1 = +20^{\circ}\text{C}$

b) różnicowo

$$\text{TWR} = \frac{1}{R_1} \frac{\Delta R}{\Delta T} \quad (39)$$

Z rys. 2 widzimy, że zarówno TWR_1 jak i TWR zmieniają istotnie swe wartości w funkcji T . Przy różnicowym określeniu TWR należy rozróżnić TWR „gorący” określony w temperaturze granicznej dodatniej i TWR „zimny” określony w temperaturach ujemnych

$$\text{TWR}_g = \frac{1}{R_1} \frac{R_2 - R_1}{T_2 - T_1} \quad (40)$$

$$\text{TWR}_z = \frac{1}{R_1} \frac{R_3 - R_1}{T_3 - T_1} \quad (41)$$

gdzie R_2 — rezystancja zmierzona przy temperaturze T_2

T_2 — temperatura maksymalna graniczna np. 125°C lub maksymalna, w której rezystor można eksploatować obciążając go pełną mocą elektryczną $T_2 = 70^{\circ}\text{C}$

R_3 — rezystancja zmierzona przy temperaturze T_3

T_3 — najniższa temperatura dopuszczalna dla eksploatacji rezystora, np. -40°C .

Należy podkreślić, że aby uzyskać jednoznaczne wyniki pomiaru TWR_z i TWR_g koniecznym jest zachowanie uzgodnionych temperatur T_2 i T_3 . Ponadto rys. 3 wykazuje, że różnicowy pomiar stosowany powszechnie zawiera wielkie przybliżenie; TWR_z jest pochyleniem prostej AB, natomiast TWR_g jest pochyleniem prostej BC, czyli

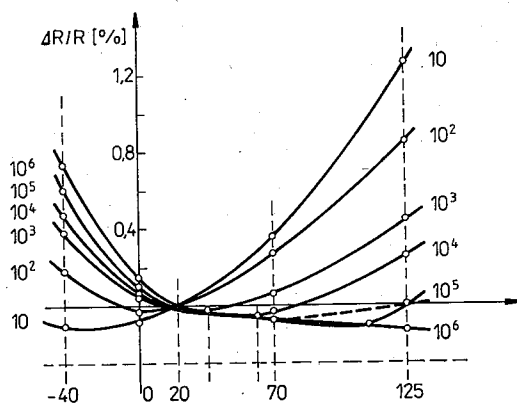
$$\text{TWR}_z \equiv \frac{AD}{DB} \quad (42)$$

oraz

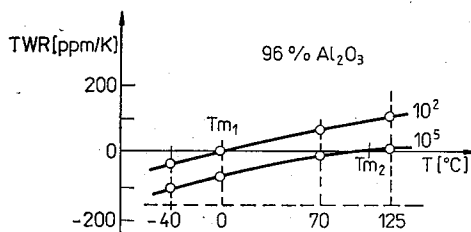
$$\text{TWR}_g \equiv \frac{CE}{BE}$$

Największy wpływ na wynik pomiaru ma wartość T_m (o której dużo mówiono w rozdz. 2.2). Używane szeroko pasty pokrywają zakres rezystancji powierzchniowych $R_\square$ od 10 do $10^6 \Omega/\square$, dla których wartość T_m zmienia się znacznie w zależności od $R_\square$. Z badań własnych uzyskano wyniki $R(T)$ jak na rys. 3. Z krzywych na rys. 3 określono TWR dla dwu rezystancji powierzchniowych 10^2 oraz $10^5 \Omega/\square$ i przedstawiono na rys. 4; wartościom T_m odpowiada wartość $\text{TWR} = 0$, czyli przecięcie osi poziomej przez odpowiednie krzywe.

Wieloletnie badania własne [29, 34] wykazały, że wiele czynników konstrukcyjnych np. grubość warstwy, czy technologicznych, np. niedopalenie warstwy, przesuwają krzywe



Rys. 3. Zależność rezystancji od temperatury otoczenia T_a rezystorów wykonanych o różnych rezystancjach powierzchniowych od $10\Omega/\square$ do $1\text{ M}\Omega/\square$



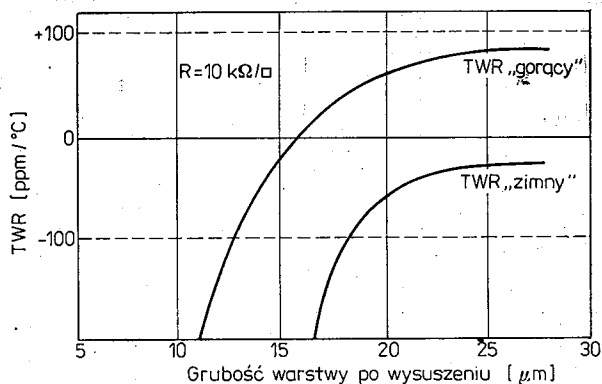
Rys. 4. Zależność TWR od temperatury otoczenia dla dwóch rezystancji powierzchniowych. Przecięcie osi poziomej odpowiada temperaturze T_m — rezystancja przy tej temperaturze osiąga minimum

w prawo, a więc T_m rośnie. W takich przypadkach duże rezystancje $R_{\square}$ nie osiągają wartości $TWR = 0$ w przedziale temperatur do $+125^{\circ}\text{C}$. Do badań optymalizacyjnych należy brać rezystory o największych $R_{\square}$ gdyż $TWR_z < 0$ i $TWR_g < 0$ i wtedy nie ma potrzeby badania równocześnie TWR_g jak i TWR_z . Reasumując, do badań wystarczy wziąć $R_{\square} = 10^6 \Omega/\square$ i badać tylko TWR_g .

3.2. WPŁYW GRUBOŚCI WARSTWY NA TWR

W procesie technologicznym wiele czynników wpływa na grubość warstwy nanoszonej sitodrukiem, dlatego znajomość wpływu grubości warstwy na jej właściwości ma zasadnicze znaczenie. Z ważniejszych czynników decydujących o grubości warstwy, to gęstość sita, lepkość pasty i odległość sita od podłoża. O ile gęstość sita i odległość sita od podłoża są świadomie dobierane przez wytwórcę sieci rezystywnych, to lepkość pasty może się zmieniać w zależności od temperatury otoczenia podczas sitodruku. Ponadto może się zmieniać w czasie, na skutek parowania rozcieńczalnika. W przypadku niewystarczającej jakości klimatyzacji pomieszczenia, w którym odbywa się druk sitowy, zmiany lepkości mogą być znaczne, gdyż zależność lepkości od temperatury jest w przybliżeniu hiperboliczna [34] i normalna temperatura sitodruku $+20^{\circ}\text{C}$ znajduje się na stromej części tej hiperboli.

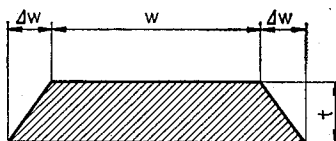
Anders [2] podał graficznie zależność TWR od grubości warstwy dla pasty PD 1400 o średniej rezystywności tj. $10 \text{ k}\Omega/\square$, co przedstawia rys. 5. Wyniki pomiarów grubości



Rys. 5. Zależność TWR od grubości warstwy [2]. Krzywe TWR „gorący” i TWR „zimny” stanowią odpowiednie wartości TWR

dotyczą warstw po wysuszeniu przed wypalaniem. Zmniejszenie grubości warstwy poniżej $20 \mu\text{m}$ powoduje przesunięcie krzywych w dół, tzn. TWR „gorący” przechodzi z wartości dodatnich do ujemnych, natomiast TWR „zimny” przyjmuje coraz większe wartości bezwzględne ze znakiem minus. Przy grubościach poniżej $15 \mu\text{m}$ zarówno TWR_z jak i TWR_g przyjmują duże wartości ujemne i takich warstw należy unikać podczas produkcji stabilnych sieci rezystywnych.

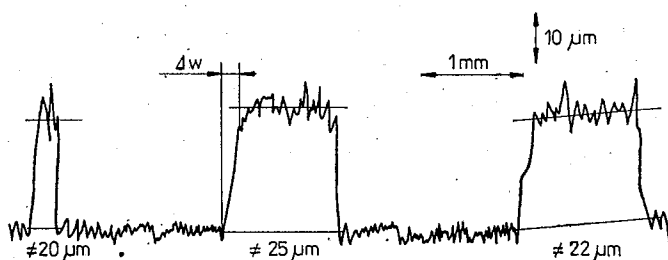
Przy analizie wpływu grubości należy rozpatrzyć wpływ szerokości ścieżki, gdyż przekrój poprzeczny ścieżki ma kształt trapezu i na obrzeżach grubość warstwy maleje, co



Rys. 6. Schemat przekroju poprzecznego warstwy rezystywnej

przedstawia rys. 6 w obszarze Δw . Szerokość warstwy projektowanej oznaczono przez w , natomiast rozplątanie warstwy przez $2\Delta w$. Tak więc krótsza podstawa trapezu wynosi w , natomiast dłuższa podstawa $w + 2\Delta w$. Prowadzone badania w OBR Mikroelektroniki Hybrydowej i Rezystorów (materiały niepublikowane) wykazały, że Δw wynosi około 0,1 mm przy grubości warstwy $20 \mu\text{m}$ gęstości sita 260 mesh (ok. 10000 oczek/ cm^2). Sito było wykonane z siatki nierdzewnej o symbolu 100 T, krawędzie rezystorów były równoległe do włókien siatki, naciąg siatki wynosił około 35 N/cm^2 . Maskowanie sita wykonano precyzyjną metodą tzw. metodą pośrednią [34].

Profilogram poprzeczny ścieżek o szerokości projektowanej 0,1 mm i 1 mm wykonano na przyrządzie Talysurf 4 i przedstawiono na rys. 7. Z rysunku można odczytać, że zreali-



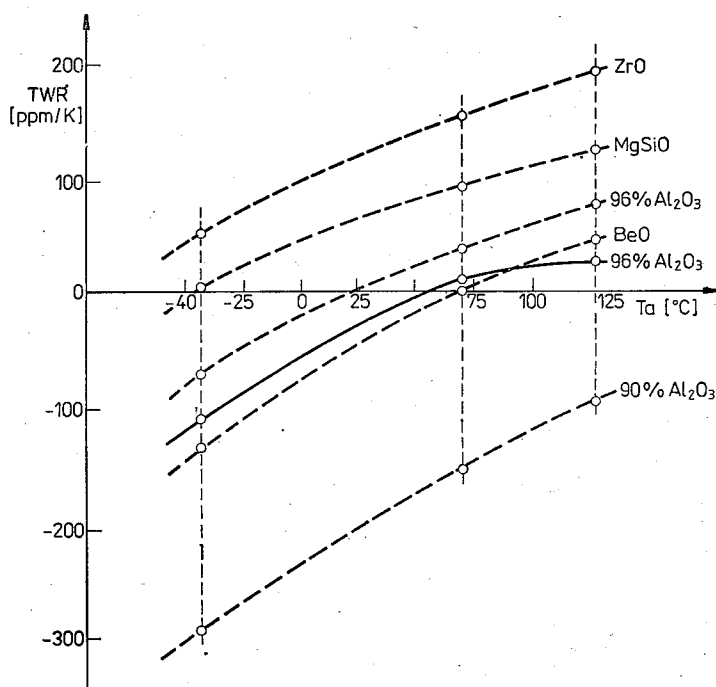
Rys. 7. Profilogram warstw rezystywnych. Z lewej ścieżka szerokości projektowanej $w = 0,1 \text{ mm}$, z prawej i w środku ścieżki szerokości projektowanej $w = 1 \text{ mm}$
Pod wykresami zanotowano zmierzoną grubość warstwy

zowane ścieżki mają większą szerokość od zaprojektowanej o $2\Delta w = 0,2 \text{ mm}$, co dla lewej ścieżki na rys. 7 stanowi 200% szerokości ($w = 0,1 \text{ mm}$) natomiast dla prawej ($w = 1 \text{ mm}$) stanowi tylko 20%. Rozlanie ścieżki Δw jest większe z lewej strony niż z prawej. Wynika to z kierunku ruchu rakli podczas sitodruku; kierunek tego ruchu był w tym przypadku od lewej do prawej. Określenie Δw nie jest zbyt łatwe i oszacowano go na 0,1 mm i przyjęto jednakową wartość po obu stronach ścieżki.

Mniejsze rozplątanie ścieżki lewej na rys. 7 o szerokości $w = 0,1 \text{ mm}$ nie jest w pełni zrozumiałe; należy podkreślić, że przy wytwarzaniu tych ścieżek rezystywnych zauważono niski stopień powtarzalności zarówno geometrii tych ścieżek, jak również ich cech elektrycznych. Na podstawie rys. 7 i wcześniejszych badań autora dotyczących stabilności rezystorów grubowarstwowych [30], można przyjąć, że przy spełnieniu warunku (44) efekt rozplątania obrzeży można pominąć

$$w \geq 2 \text{ mm}$$

(44)



Rys. 8. Zależność TWR od temperatury otoczenia warstw rezystywnych DP-1441 wytworzonych na podłożach z różnych materiałów; 1 — ceramika cyrkonowa, 2 — ceramika steatytowa, 3 i 5 — ceramika alundowa 96% Al_2O_3 , 4 — ceramika berylowa, 6 — ceramika alundowa ok. 90% Al_2O_3 (krzywe 5 i 6 — wyniki własne)

Wytwórcy testują swe pasty przy szerokości rezystorów $w = 5$ mm, natomiast szkodliwy wpływ obrzeży obserwuje się przy $w < 1$ mm. Dlatego wypływa stąd praktyczny wniosek: należy projektować rezystory o powierzchni większej niż 1 mm^2 tj. szerokości $w > 1$ mm oraz długości $L > 1$ mm.

3.3. WPŁYW PODŁOŻA NA TWR

Doświadczenia wykazują dwa rodzaje wpływu alundowego podłoża na TWR grubej warstwy rezystywnej: a) w wyniku różnic współczynników rozszerzalności liniowej podłoża i warstwy; b) w wyniku chemiczno-dyfuzyjnego oddziaływania podłoża na warstwę podczas jej wypału.

Wydrukowane i wypalone w tych samych warunkach warstwy na różnych podłożach wykazują różne wartości TWR, szczególnie zmienia się wartość T_m , co z punktu widzenia zastosowań sieci rezystywnych ma zasadnicze znaczenie. Na rys. 8 przedstawiono zależność TWR od temperatury otoczenia warstwy rezystywnej wykonanej na podłożach z różnych materiałów. Punkt przecięcia krzywych z osią poziomą odpowiada wartościom T_m (TWR = 0). Jako przykład można podać, że warstwa na steatycie (Mg Si O_2) ma $T_m = -40^\circ\text{C}$ na alundum 96% Al_2O_3 , $T_{m1} = 25^\circ\text{C}$ lub $T_{m2} = 60^\circ\text{C}$ (wynik własny),

na alundum 90% Al_2O_3 $T_m > 125^\circ\text{C}$. Początkowo próbowano tłumaczyć te różnice współczynnikami rozszerzalności liniowej różnych podłoży. Werma i Sharma [38] przyjęli TW_ϱ jako hipotetyczny temperaturowy współczynnik samej warstwy bez podłoża. Warstwa w zależności od współczynnika rozszerzalności warstwy α_f i współczynnika rozszerzalności podłoża α_s będzie ściskana lub rozciągana przy zmianach temperatury otoczenia. Ponieważ warstwy wykazują znaczny efekt piezorezystywny ściskanie warstwy daje zmniejszenie rezystancji a jej rozciąganie powoduje wzrost rezystancji; makroskopowo objawia się to zmianą TWR. Matematyczna analiza jest następująca [38]: jeżeli ϱ jest rezystywnością, L — długością, w — szerokością, t — grubością warstwy rezystywnej rezystora to rezystancja R ośrodka izotropowego wyraża się wzorem

$$R = \frac{\varrho L}{wt} \quad (45)$$

natomiast zmiany względne w funkcji temperatury można wyrazić przez pochodną względem temperatury, jak to przedstawia zależność (46)

$$\frac{1}{R} \frac{\partial R}{\partial T} = \frac{1}{\varrho} \frac{\partial \varrho}{\partial T} + \frac{1}{L} \frac{\partial L}{\partial T} - \frac{1}{w} \frac{\partial w}{\partial T} - \frac{1}{t} \frac{\partial t}{\partial T} \quad (46)$$

Dla materiału izotropowego

$$\frac{1}{L} \frac{\partial L}{\partial T} = \frac{1}{w} \frac{\partial w}{\partial T} = \frac{1}{t} \frac{\partial t}{\partial T} = \alpha_f \quad (47)$$

jest to współczynnik rozszerzalności liniowej warstwy rezystywnej. Podstawiając (47) do (46) i dokonując redukcji otrzymuje się temperaturowy współczynnik rezystancji samej warstwy bez podłoża

$$\frac{1}{R} \frac{\partial R}{\partial T} = \frac{1}{\varrho} \frac{\partial \varrho}{\partial T} - \alpha_f = \text{TW}_\varrho - \alpha_f \quad (48)$$

Po wprowadzeniu współczynnika rozszerzalności liniowej podłoża α_s , współczynnika Poissona warstwy ν oraz czułości piezorezystywnej γ wyrażonej wzorem (49)

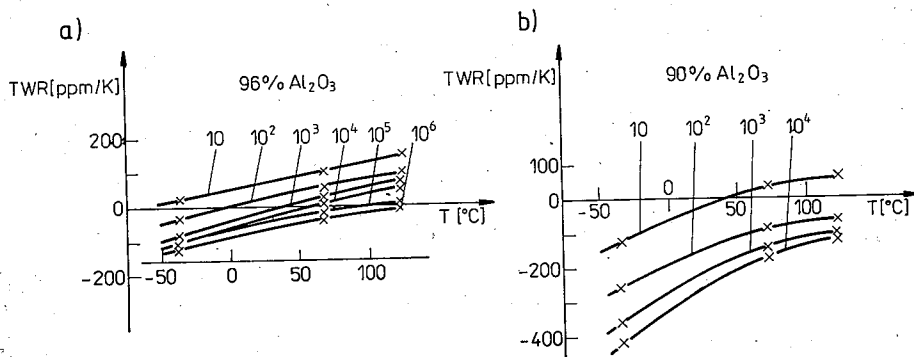
$$\gamma = \frac{\frac{\partial R}{R}}{\frac{\partial L}{L}} \quad (49)$$

uzyskano [4, 38] zależność wyrażającą wpływ naprężeń na TWR, a mianowicie

$$\text{TWR} = \text{TW}_\varrho - \alpha_f - \frac{2(\alpha_f - \alpha_s)(\gamma - 1 - \nu)}{1 - \nu} \quad (50)$$

Autor poddał badaniom [33] rezystory testowe całej rodziny past DP 1400 na podłożach alundowych o zawartości 96% Al_2O_3 , wypalanych w temperaturze 1700°C , których wyniki $R(T_a)$ przedstawia rys. 3, natomiast $\text{TWR} = \varphi(T_a)$ przedstawia rys. 9a. Dla porównania na rys. 9b przedstawiono wyniki własne [33] rezystorów wykonanych z tych samych past

i w tym samym procesie technologicznym, lecz na podłożach z tworzywa o zawartości około 90% Al_2O_3 , spiekanych w temperaturze 1400°C (zmniejszenie temperatury wypalania płytek podłożowych z 1700°C do 1400°C ma duże znaczenie ekonomiczne). Mimo, że pasta



Rys. 9. Zależność TWR od temperatury otoczenia: a) podłoże 96% Al_2O_3 o współczynniku rozszerzalności liniowej $\alpha_{s1} = 8 \cdot 10^{-6}/\text{K}$, b) podłoże ok. 90% Al_2O_3 o współczynniku rozszerzalności liniowej $\alpha_{s2} = 5 \cdot 10^{-6}/\text{K}$

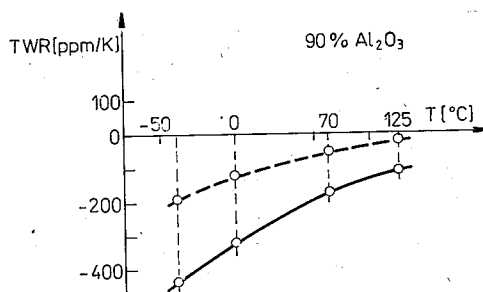
zastosowana była ta sama i warunki wypalania identyczne, to wyniki TWR na obu podłożach są zdecydowanie różne. Przesunięcia krzywych w prawo na rys. 9b nie tłumaczą różnice współczynników rozszerzalności liniowej α_{s1} i α_{s2} .

Do obliczeń przyjęto $\alpha_f = 15 \cdot 10^{-6}/\text{K}$ współczynnik Poissona $\nu = 0,22$ na podstawie literatury [4].

$\alpha_{s1} = 8 \cdot 10^{-6}/\text{K}$ dla podłoża 96% Al_2O_3 (z pomiaru)

$\alpha_{s2} = 5 \cdot 10^{-6}/\text{K}$ dla podłoża 90% Al_2O_3 (z pomiaru).

Dokonano porównania wyników pomierzonych z wynikami obliczonymi według wzoru (50) i przedstawiono na rys. 10 dla jednej pasty o rezystancji powierzchniowej



Rys. 10. Porównanie zależności TWR od temperatury otoczenia: ——— wyniki pomiarów; - - - - - wyniki obliczeń wg wzoru (47), ceramika 90% Al_2O_3 , warstwa 10 $\text{k}\Omega/\square$ (DP 1441)

10 $\text{k}\Omega/\square$ (DP 1441). Pomierzone wartości TWR warstw na podłożu 90% Al_2O_3 zupełnie nie zgadzały się z wyliczonymi. Próbowano wyliczyć, jaka musiałaby być czułość piezorezystywna γ , żeby uzyskać zadawalającą zgodność wyników z eksperymentem. W rezultacie otrzymano γ od 1 do 50, natomiast z pomiarów tensometrycznych oszacowano

$\gamma = 10\text{--}15$. Tak więc wykazano, że przesunięcie T_m nie da się wytłumaczyć mniejszym współczynnikiem α_{s2} podłoża 90% Al_2O_3 od współczynnika α_{s1} podłoża 96% Al_2O_3 . Wytłumaczenia tych różnic należy szukać w drugim rodzaju wpływu podłoża wymienionym na początku tego rozdziału, oznaczonym literą b). Wpływ ten można wyjaśnić większym udziałem przewodnictwa aktywowanego termicznie wzory (23) i (35). Atomy wnika-
jące do szkliwa między ziarna przewodzące warstwy mogą stawać się centrami rezonansowymi tunelowania, lub pułapkami elektronów w prądzie przeskokowym. Obie te składowe dają ujemny TWR.

Nie wyklucza się też, że obszary między ziarnami w wyniku dyfundowania do nich domieszek mogą uzyskać właściwości półprzewodnika, co dałoby również TWR ujemny o dużej wartości bezwzględnej.

Materiał podłoża 90% Al_2O_3 jest tak dobrany, żeby można było wypalać płytki podłożowe w temperaturze 1400°C, a więc bez atmosfery wodorowej, w której jest wypalane podłoże 96% Al_2O_3 (1700°C). Tworzywo ceramiczne 90% Al_2O_3 jest znacznie tańsze i łatwiejsze w obróbce, gdyż zawiera ok. 10% dodatków uplastyczniających. Niestety te dodatki mogą dyfundować do warstwy rezystywnej podczas jej wypalania w temperaturze 950°C w czasie ponad 10 min (pełny cykl około 1 h).

Inaczej przedstawia się problem w przypadku podłoża 96% Al_2O_3 . Po pierwsze dodatki są ściśle kontrolowane i stanowią tylko 4%. Wypalanie tych podłoży w temperaturze 1700°C prawdopodobnie bardziej wiąże te dodatki z siecią polikrystaliczną alundum i przez to ich dyfuzja jest utrudniona. Problem poprawy podłoża 90% Al_2O_3 jest problemem istotnym z punktu widzenia podłoża o złożonych kształtach (np. do potencjometrów grubowarstwowych). Jednym z kierunków prac nad tymi podłożami byłaby ścisła kontrola dodatków i eliminacja z nich pierwiastków o dużej ruchliwości.

3.4. WPŁYW TECHNOLOGII

Obserwacje wykazały, że na TWR największy wpływ mają: temperatura wypalania, czas wypalania warstwy, a także ilość dopływającego powietrza podczas tego procesu. Ilość powietrza jest ograniczona jednostronnie tzn. ilość ta musi być wystarczająco duża, aby proces utleniania zachodził do końca. Daje to dwie zalety warstw: a) proces utleniania jest skończony i dalsze utlenianie warstwy nie występuje podczas eksploatacji — warstwa jest stabilna w czasie, b) zarówno RuO_2 jak i $\text{Bi}_2\text{Ru}_2\text{O}_7$ wykazują TWR dodatni, co jest niezbędne dla uplasowania T_m w zakresie temperatur roboczych. Stosuje się tu dwa rozwiązania nawiewu powietrza:

a) regulowany nadmuch powietrza z kompresorów bezolejowych,

b) wymuszony przepływ powietrza, spowodowany pochyleniem tunelu grzewczego przez który przechodzą płytki z nadrukowaną warstwą. Pierwszy sposób gwarantuje wystarczające utlenianie, w drugim sposobie ilość utlenianych rezystorów na jednostkę czasu jest ograniczona, dlatego to rozwiązanie jest stosowane do pieców o małej wydajności.

Temperatura wypalania i czas wypalania wymagają optymalizacji w warunkach produkcyjnych dla konkretnego zestawu urządzeń.

4. OPTYMALIZACJA CZYNNIKÓW TECHNOLOGICZNYCH CELEM MINIMALIZACJI BEZWZGLĘDNEJ WARTOŚCI TWR

4.1. WYBÓR REZYSTANCJI POWIERZCHNIOWEJ DO BADAŃ

Badanie całego szeregu rezystancji jest bardzo pracochłonne, dlatego należy wybrać pastę i warstwę najbardziej reprezentatywną. Rys. 9 i doświadczenia autora wskazują, że łatwiej uzyskać TWR ujemny niż TWR dodatni. Warstwa o rezystancji powierzchniowej $10^6 \Omega/\square$ na rys. 9a posiada w całym zakresie temperatur roboczych (-40°C do $+125^\circ\text{C}$) TWR ujemny, a więc jeżeli zoptymalizujemy warunki technologiczne dla tej warstwy i uzyskamy TWR „gorący” ujemny o bardzo małej wartości bezwzględnej lub nawet dodatni, to wynik będzie obowiązywał dla wszystkich past tzn. wszystkie krzywe przesuną się w lewo na rysunku (por. również rozdz. 3.1). Jest to efekt pożądaný. Uwzględniając powyższe rozważania oraz troskę o oszczędność nakładów wystarczy przebadać sieci rezystywne z pasty $10^6 \Omega/\square$ i tylko TWR „gorący”. Przesunięcie TWR „gorącego” w kierunku wartości dodatnich spowoduje zmniejszenie wartości bezwzględnej TWR „zimnego”, gdyż krzywe przesuwają się w przybliżeniu równolegle. Tak więc z kilkudziesięciu wartości optymalizacji zostanie poddana jedna wartość rezystancji i wyniki będą rozciągnięte na całą rodzinę past przy użyciu jednego podłoża o zawartości 96% Al_2O_3 .

4.2. PLAN EKSPERYMENTU I JEGO WYKONANIE

Na podstawie wcześniejszych badań [29, 30, 34] jak również analiz poszczególnych wpływów autor wybrał trzy zmienne niezależne

x_1 — standaryzowana odległość sita od podłoża — istotny czynnik wpływający na grubość warstwy, a więc i na TWR (rozdz. 3.2),

x_2 — standaryzowany czas obróbki termicznej

x_3 — standaryzowana szczytowa temperatura wypalania.

Poszukiwaną funkcją jest

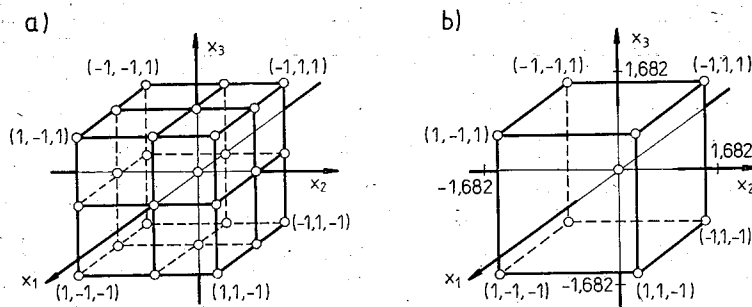
$$y = \text{TW}R = \varphi(x_1, x_2, x_3) \quad (51)$$

Początkowo autor próbował aproksymować tę funkcję liniowymi zależnościami [31] stosując analizę regresji I stopnia. Próby te wykazały, że są to wielowymiarowe przebiegi wypukłe i najprostszym przybliżeniem jest krzywa regresji drugiego stopnia podana wzorem (52)

$$y = b_0 + \sum_{i=1}^3 b_{ii} x_i^2 + \sum_{i=1}^2 \sum_{j=1}^3 b_{ij} x_i x_j + \sum_{i=1}^3 b_i x_i. \quad (52)$$

Niewiadomymi w tym równaniu są współczynniki regresji b_0 , b_{ii} , b_{ij} , b_i natomiast x_{ui} , y_{ui} są wartościami otrzymanymi z eksperymentu, gdzie u — jest numerem kolejnego doświadczenia zmieniający się od $u = 1$ do $u = 20 = N$, i zmienia się kolejno od 1 do 3.

Spośród wielu metod projektowania badań autor rozważa dwie: metodę rotatabilnego planowania i metodę czynników uwzględniając nieliniowość (wypukłość) badanych przebiegów. W obu metodach punktem wyjścia jest standaryzacja (Dodatek) zmiennych niezależnych (czynników), a następnie ułożenie programu badań. W metodzie czynnikowej eksperymenty mają współrzędne odpowiadające węzłom sieci przedstawionej na rys. 11a.



Rys. 11. Rozmieszczenie doświadczeń w przestrzeni 3-wymiarowej a) metoda czynnikowa, b) planowanie rotabilne

Dla trzech czynników x_1, x_2, x_3 ilość doświadczeń wynosi $N = 3^3 = 27$, a dla m czynników.

$$N = 3^m \quad (53)$$

Podstawa 3 wynika z faktu wypukłości (dla wyznaczenia krzywej wypukłej potrzeba minimum 3 punkty). W metodzie planowania rotabilnego przestrzeń sześciąta foremnego jest „rzadziej” przebadana, gdyż badania wykonuje się tylko w punktach o współrzędnych odpowiadających narożnikom sześciąta $n_c = 8$ i w środku sześciąta $N_0 = 6$ lecz dodatkowo wykonuje się badania poza sześciątem w tzw. punktach gwiazdnych (por. Dodatek). Odległość punktu gwiazdnego (promień gwiazdny ρ_g) od środka sześciąta zależy od ilości zmiennych niezależnych m : dla $m = 3$ $\rho_g = 1,682$.

Współrzędne doświadczeń rotabilnego planowania przedstawia rys. 11b. Ogólny wzór na ilość doświadczeń ma postać

$$N = 2^m + 2m + N_0 \quad (54)$$

a dla $m = 3$, $N = 8 + 6 + 6 = 20$, daje to oszczędność około 30% doświadczeń w stosunku do metody czynnikowej, ponadto otrzymane wyniki są mało wrażliwe na obrót układu współrzędnych i prawie wszystkie współczynniki otrzymanego równania regresji II stopnia nie są skorelowane ze sobą.

Uwzględniając powyższe zalety planowania rotabilnego zastosowano go do optymalizacji warunków wytwarzania warstw grubych rezystywnych, tak aby uzyskać zmniejszenie bezwzględnej wartości zmian termicznych TWR. Plan eksperymentu podano w Dodatku, natomiast opracowany program komputerowy, według którego obliczono współczynnik równania regresji, przedstawiono obok.

Uśrednione wyniki każdego punktu badań stanowiły macierz obserwacji. Ilość punktów badań wynosiła $N = 20$ przy różnych wartościach x_1, x_2, x_3 zgodnie z planem eksperymentu (Dodatek). Ta macierz obserwacji była podstawą obliczeń komputerowych. Mimo 10 niewiadomych czas obliczeń na komputerze JBM PC XT nie przekraczał 30 sekund. Z pierwszego etapu badań uzyskano współczynniki regresji dla TWR

$$\begin{aligned} b_0 &= 272,7 & B_{12} &= 19,6 & b_1 &= 38,7 \\ b_{11} &= 7,47 & b_{13} &= 68,3 & b_2 &= 197,8 \\ b_{22} &= -30 & b_{23} &= -156,8 & b_3 &= 311 \\ b_{33} &= -209,9 \end{aligned} \quad (55)$$

PROGRAM KOMPUTEROWY

```

rem współrzędne planu experimentu:
data -1,-1,-1, 1,-1,-1, -1,1,-1, 1,1,-1, -1,-1,1, 1,-1,1
data -1,1,1, 1,1,1, -1,682,0,0, 1,682,0,0, 0,-1,682,0
data 0,1,682,0, 0,0,-1,682, 0,0,1,682, 0,0,0, 0,0,0, 0,0,0
data 0,0,0, 0,0,0, 0,0,0
dim x(20,10),y(20),a(10,10),b(10),y1(10),i(10)
for i=1 to 20:read y(i):next
for i=1 to 20
  x(i,1)=1:read x(i,2),x(i,3),x(i,4):x(i,5)=x(i,2)^2:x(i,6)=x(i,3)^2
  x(i,7)=x(i,4)^2:x(i,8)=x(i,2)*x(i,3):x(i,9)=x(i,2)*x(i,4)
  x(i,10)=x(i,3)*x(i,4)
next

for i=1 to 10:i(i)=1
for j=1 to 10:a(i,j)=0
  for k=1 to 20:a(i,j)=a(i,j)+x(k,i)*x(k,j):next
next
y1(i)=0
  for k=1 to 20:y1(i)=y1(i)+y(k)*x(k,i):next
next
for z=1 to 9
  if a(z,z)=0 then
    l=z+1:while a(l,z)=0 and l<10:l=l+1:wend
    for k=z to 10:swap a(z,k),a(l,k):next
    swap i(z),i(l):swap y1(z),y1(l)
  else
    p=z+1:while a(p,z)=0 and p<10:p=p+1:wend
    for l=p to 10
      a=a(l,z)/a(z,z)
      for k=z to 10:a(l,k)=a(l,k)-a(z,k)*a:next
      y1(l)=y1(l)-y1(z)*a
    next
  end if
next

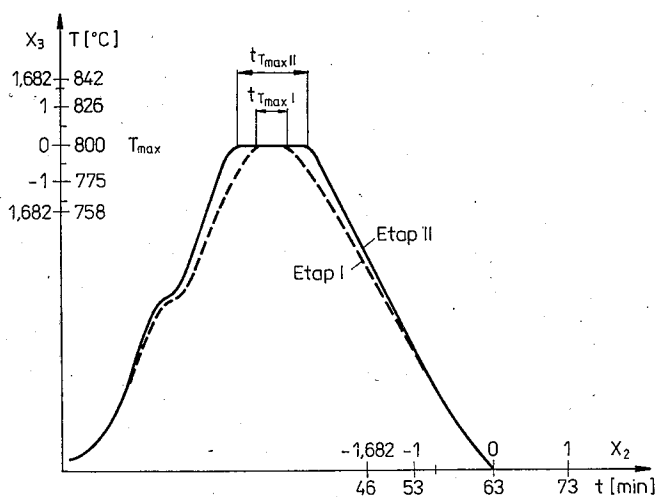
b(10)=y1(10)/a(10,10)
for i=9 to 1 step -1
  for j=i+1 to 10:y1(i)=y1(i)-a(i,j)*b(j):next
  b(i)=y1(i)/a(i,i)
next
print "Dla wyników eksperymentu:"
for i=1 to 20:print y(i):next:print
print "Otrzymano następujące współczynniki:"
for k=1 to 10:read b$:print b$:"=":b(i(k)):next
lprint "Dla wyników eksperymentu:"
for i=1 to 20:lprint y(i):next:lprint
lprint "Otrzymano następujące współczynniki:":restore wsp:
for k=1 to 10:read b$:lprint b$:"=":b(i(k)):next

wsp:data b0,b1,b2,b3,b11,b22,b33,b12,b13,b23

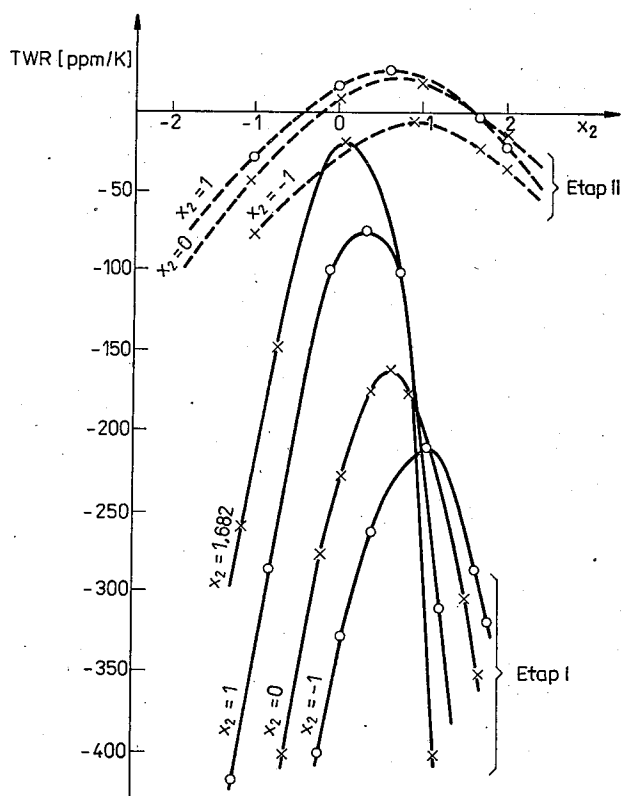
```

Analiza otrzymanej funkcji wykazała silną zależność TWR od temperatury wypalania x_3 i od czasu x_2 . Bezwzględna wartość TWR malała ze wzrostem czasu wypalania, co obrazuje poniższa tabela

x_2	-1	0	+1	+1,682
czas t [minuty]	53	63	73	80
$y_{opt.} = \text{TWR}$ [ppm/K]	-210	-154	-74,5	-21



Rys. 12. Temperaturowy profil pieca; ----- krzywa w etapie I, — krzywa w etapie II



Rys. 13. TWR „gorący” w funkcji temperatury wypalania; x_2 — standaryzowany czas wypalania, ----- etap I, — etap II

Postanowiono wydłużyć czas wypalania w strefie gorącej i wstępnie wykonane badania nazwano etapem I, natomiast przygotowano etap II, w którym czas przebywania warstw rezystywnych w strefie t_{TmaxII} był dwa razy dłuższy niż t_{TmaxI} . Różnice te przedstawia rys. 12, na którym przedstawiono profile pieca w obu etapach badań dla $x_2 = 0$ oraz $x_3 = 0$. Powtórzony eksperyment dla wydłużonego czasu wypalania t_{TmaxII} dał zdecydowane zmniejszenie wartości bezwzględnej TWR, obok wartości małych ujemnych w niektórych badaniach pojawiły się wartości dodatnie. Obliczone współczynniki regresji wyniosły:

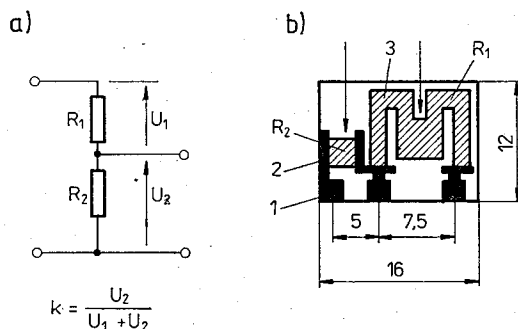
$$\begin{aligned} b_0 &= 3,83 & b_{12} &= -0,75 & b_1 &= 2,3 \\ b_{11} &= -14,7 & b_{13} &= -1,5 & b_2 &= 21,5 \\ b_{22} &= -14,3 & b_{23} &= -6 & b_3 &= 32,2 \\ b_{33} &= -20,8 \end{aligned} \quad (56)$$

Optymalny czas całkowity pozostał $x_2 = 0$ tj. 63 minuty a czas przebywania warstw rezystywnych w maksymalnej temperaturze $t_{TmaxII} = 11$ minut. Temperatury optymalne przesunęły się znacznie powyżej 800°C . Wyniki badań I i II etapu przedstawiono na rys. 13.

Etap II wykazuje jak dobrać czas i temperaturę wypalania, by uzyskać $|TWR|$ sieci rezystywnych poniżej 40 ppm/K.

5. KOMPENSACJA TERMICZNA SIECI REZYSTYWNYCH

W grubowarstwowych sieciach rezystywnych istnieje możliwość samoczynnej kompensacji wpływów temperatury otoczenia, gdyż rezystory znajdują się na wspólnej płytce



Rys. 14. Dzielnik rezystancyjny zintegrowany: a) schemat, b) rozmieszczenie rezystorów na płytce. 1 — kontakt lutowniczy, 2 — elektroda rezystora, 3 — warstwa rezystywna. Strzałki oznaczają kierunek korekcji rezystancji w procesie produkcyjnym

podłożowej o dużej przewodności cieplnej. Kompensacja ta jest skuteczna, gdy rezystory pracują w gałęziach tłumików lub dzielników napięcia. Na rys. 14a przedstawiono schemat najprostszego dzielnika składającego się z dwóch rezystorów R_1 i R_2 o współczynnikach TWR_1 i TWR_2 , natomiast na rys. 14b znajduje się topologia tych rezystorów na płytce

alundowej. Przy odpowiednich założeniach dotyczących wejścia i wyjścia współczynnik podziału napięcia k wynosi

$$k = \frac{R_2(T)}{R_1(T) + R_2(T)}. \quad (57)$$

Podstawiając najprostsze zależności otrzymuje się wyrażenie

$$k = \frac{R_2(1 + \text{TWR}_2 \cdot T)}{R_1(1 + \text{TWR}_1 \cdot T) + R_2(1 + \text{TWR}_2 \cdot T)}. \quad (58)$$

Dla wygody analizy matematycznej można przyjąć, $\text{TWR}_1 = a$, $\text{TWR}_2 = b$; wyrażenie (58) przyjmuje wtedy postać

$$k = \frac{R_2(1 + bT)}{R_1(1 + aT) + R_2(1 + bT)}. \quad (59)$$

Żeby określić wpływ temperatury należy znaleźć pierwszą pochodną względem temperatury

$$\frac{\partial k}{\partial T} = \frac{bR_2R_1(1 + aT) + R_2(1 + bT) - (aR_1 + bR_2)R_2(1 + bT)}{R_1[(1 + aT) + R_2(1 + bT)]^2} \quad (60)$$

Uwzględniając, że

$$|aT| \ll 1 \quad \text{ i } \quad |bT| \ll 1 \quad (61)$$

łatwo zauważyć, że mianownik jest zawsze dodatni i należy przeanalizować tylko licznik. Jeżeli mianownik oznaczy się przez M i dokona rozwinięcia oraz redukcji w liczniku otrzymuje się:

$$\frac{\partial k}{\partial T} = \frac{R_1R_2(b-a)}{M} = \frac{R_1R_2(\text{TWR}_2 - \text{TWR}_1)}{M} \quad (62)$$

$$\frac{\partial k}{\partial T} = \frac{R_1R_2\Delta\text{TWR}}{M}. \quad (63)$$

Jest to istotny wniosek, że o niestabilności termicznej współczynnika podziału k decyduje różnica — ΔTWR , a nie same ich wartości bezwzględne.

Aby obliczyć zmianę Δk przy zadanej zmianie ΔT należy dokonać dalszych uproszczeń i przekształceń wzoru (62)

$$\frac{\Delta k}{\Delta T} = \frac{R_1R_2\Delta\text{TWR}}{M} \quad (64)$$

Uwzględniając warunek (61) można uprościć mianownik

$$\frac{\Delta k}{T} = \frac{R_1R_2(b-a)}{(R_1 + R_2)^2} = \frac{R_1R_2(b-a)}{R_1^2 + 2R_1R_2 + R_2^2}. \quad (65)$$

Dzieląc licznik i mianownik przez R_1^2 otrzymuje się

$$\frac{\Delta k}{\Delta T} = \frac{k(b-a)}{1 + 2k + k^2}; \quad (66) \quad \Delta k = \frac{\Delta T k(b-a)}{1 + 2k + k^2}. \quad (67)$$

Dla większych stopni podziału dzielnika (mniejszy współczynnik k) mianownik wzoru (67) zmierza do jedności, wtedy

$$\Delta k \approx \Delta T \cdot k \cdot \Delta TWR \quad (68)$$

Dla dobrej kompensacji przyrost temperatury ΔT musi być taki sam dla R_1 jak i dla R_2 .

Biorąc powyższe pod uwagę można określić zalety zintegrowanych sieci rezystywnych w stosunku do rezystorów dyskretnych przeznaczonych na dzielniki lub tłumiki:

1. Rezystory w sieciach zintegrowanych są wytwarzane z tych samych surowców; to samo podłoże wspólne, ta sama pasta rezystywna, to samo pokrycie zabezpieczające.

2. Rezystory w sieci rezystywnej są wytwarzane w tym samym cyklu produkcyjnym; ten sam sitodruk, to samo suszenie ścieżek, ta sama temperatura wypalania, ten sam czas wypalania, identyczne wszystkie czynniki technologiczne znane i nieznane, ta sama historia rezystorów na wspólnym podłożu.

3. Między rezystorami w zintegrowanej sieci rezystywnej jest dobre sprzężenie termiczne, gdyż podłoże alundowe ma dobrą przewodność cieplną tylko o rząd gorszą od metali. Można tak zaprojektować topologię rezystorów na płytce, że niezależnie od obciążenia poszczególnych rezystorów mają one w stanie ustalonym tę samą temperaturę podczas pracy $\Delta T_1 = \Delta T_2 = \Delta T$.

Wykorzystując zoptymalizowany proces można uzyskać z pasty DP 1461 zintegrowaną sieć rezystywną dzielnika do oscyloskopu lub multimetru o dużych rezystancjach np. według rys. 14b rezystor R_1 o długości czynnej $L_1 = 20$ mm i szerokości $w_1 = 2$ mm będzie posiadał rezystancję $10 \text{ M}\Omega$, natomiast rezystor R_2 o długości $L_2 = 2$ mm i szerokości $w_2 = 2$ mm rezystancję $1 \text{ M}\Omega$. Ten wysokoomowy dzielnik przy $TWR_1 = -40 \text{ ppm/K}$ i $TWR_2 = -37 \text{ ppm/K}$ osiągnie przy $\Delta T = 30^\circ\text{C}$ wartość zmiany $\Delta k = 0,009\%$. Taka dobra stabilność nie jest osiągalna przy użyciu rezystorów dyskretnych o dużych rezystancjach. Wartość $\Delta TWR = 3 \text{ ppm/K}$ rezystorów na jednej wspólnej płytce jest łatwa do osiągnięcia w technologii grubowarstwowej, natomiast w technologii cienkowarstwowej osiąga się $\Delta TWR = 1 \text{ ppm/K}$, lecz przy znacznie węższym zakresie rezystancji znamionowych.

6. WNIOSKI

1. Na podstawie literatury i wieloletnich badań autora wykazano, że na TWR warstwy grubej oddziałują czynniki konstrukcyjne, jak wymiary warstwy, rodzaj warstwy, materiał podłoża, rodzaj i gęstość sita oraz technologiczne, z których najważniejsze to temperatura wypalania, czas wypalania i lepkość pasty w czasie sitodruku. Wpływ obrzeży jest widoczny przy szerokości ścieżki poniżej 1 mm.

2. Przedstawione opisy matematyczne mechanizmu przewodzenia nie ujmują ilościowej zależności między czynnikami konstrukcyjnymi i technologicznymi a temperaturowym współczynnikiem rezystancji warstwy.

3. Zależność rezystancji warstwy grubej od temperatury wykazuje minimum przy temperaturze T_m , co jakościowo wyjaśniają modele matematyczno-fizyczne. Wartość T_m

zależy jednak od wszystkich czynników wymienionych we wniosku Nr 1; ujęcie matematyczne ilościowe tej zależności napotyka na duże trudności i wymaga daleko idących założeń upraszczających.

4. Wyciągnięte wnioski optymalizacyjne z zaprojektowanego i przeprowadzonego przez autora eksperymentu, według rotatabilnego planu II stopnia, pozwoliły na zmniejszenie bezwzględnej wartości TWR prawie 10-krotnie.

5. Odpowiednie zaprojektowanie zintegrowanych sieci rezystywnych, przedstawione w rozdziale 5, pozwala na kompensację zmian rezystancji dzielników oraz tłumików i poprawę ich stabilności termicznej około 50 razy w stosunku do odpowiednich rozwiązań wykonanych z rezystorów dyskretnych.

DODATEK

W Dodatku podano w sposób tabelaryczny uzupełnienia dotyczące rotatabilnego planu eksperymentu II stopnia

Wartości średnie i zakresy zmienności mierzalnych zmiennych niezależnych

Zmienne standaryzowane $x_1, x_2, x_3^{*)}$	Zmienne mierzalne			Uwagi
	X_1 mm	X_2 minuty	X_3 °C	
-1,682	0,8	46	758	środek doświadczania
-1	1,5	53	775	
0	2,5	63	800	
1	3,5	73	825	
1,682	4,2	80	842	
—	1	10	25	ΔC_{max}

*) x_1 — standaryzowana odległość sita od płytki podłożowej,

x_2 — standaryzowany czas obróbki termicznej,

x_3 — standaryzowana temperatura obróbki wypalania.

Standaryzacja zmiennych polega na przekształceniu

$$x_i = \frac{X_i - \bar{X}_i}{\Delta C_{max}}$$

gdzie: X_i — wartość mierzona

$\bar{X}_i$ — średnia wartość z pomiarów i -tej zmiennej

$$\Delta C_{max} = X_{i,max} - \bar{X}_i.$$

Jest to maksymalny zakres zmienności X_i wewnątrz sześcianu na rys. 11.

Plan eksperymentu

u	x_1	x_2	x_3	Uwagi
1	-1	-1	-1	$n_c = 8$
2	1	-1	-1	
3	-1	1	-1	
4	1	1	-1	
5	-1	-1	1	
6	1	-1	1	
7	-1	1	1	
8	1	1	1	
9	-1,682	0	0	$n_q = 6$
10	1,682	0	0	
11	0	-1,682	0	
12	0	1,682	0	
13	0	0	-1,682	
14	0	0	1,682	
15	0	0	0	$N_0 = 6$
16	0	0	0	
17	0	0	0	
18	0	0	0	
19	0	0	0	
20	0	0	0	

Wartości liczbowe parametrów planu eksperymentu od wymiaru przestrzeni m

Oznaczenie wielkości	Wymiar przestrzeni m			
	2	3	4	5
Liczba prób N	13	20	31	52
Liczba punktów równo oddalonych od środka pierwszego rzędu n_0	4	8	16	32
Liczba punktów gwiazdnych n_q	4	6	8	10
Promień punktu gwiazdowego q	1,414	1,682	2,000	2,378
Liczba punktów w środku doświadczenia N_0	5	6	7	10

BIBLIOGRAFIA

1. B. Abeles: *Granular Metal Films*. Applied Solid State Science, vol. 6, Academic Press, New York 1976
2. J. Anders: *New Resistive Inks*. Bull. firmy Du Pont 1979
3. L. J. Brady: *Mechanism of Conduction in Thick Film Cermet Resistors*. Proc. Electron. Comp., Washington, D. C., May 1967, p. 238+246
4. A. Cattaneo, L. Pirozzi, B. Morten, M. Prudenziati: *Influence of the Substrate on the Electrical Properties of Thick Film Resistors*. Proc. of European Hybrid Microelectronics Conf., 1979, Gent., p. 146—150

5. D. S. Campbell: *The Use of Thin Film in Physical Investigations*. Academic Press inc., New York 1966
6. T. J. Coutts: *Active and Passive Thin Film Devices*. Academic Press, London, New York, San Francisco 1978
7. C. B. Duke: *Tunneling Phenomena in Solids*. Edited by E. Burstein and S. Lundquist Plenum, New York 1969
8. A. L. Efros, B. J. Shklovskii: *Phys. Stat. Solid.*, 1976, 76, 475
9. P. J. Even, J. M. Robertson: *A Percolation Model of Conduction Segregated Systems of Metallic and Insulating Materials: Application to Thick Film Resistors*. *J. Phys. D: Appl. Phys.* 14, 1981, p. 2253, G. Britain
10. K. Fuchs: *Proc. Cambridge Phil. Soc.* 34, 100, 1938
11. N. C. Halder: *Electron Tunneling and Hopping Possibilities in RuO₂ Thick Films*. *Electrocomponent Science and Technology*, 1983, vol. 11 p. 21
12. *Handbook of Thin Film Technology*. Edit. L. I. Maissel and R. Glang. Mc Graw Hill, New York, Toronto, London, Düsseldorf, 1970
13. W. R. Hardy, J. Schewchum: *Reactively Sputtered Tantalum Thin Film Resistors*. Part I. Physical and Electrical Properties. *Thin Solid Films*, 1971, vol. 8 p. 81
14. R. M. Hill: *Electrical Conduction in Discontinuous Metal Films*. *Contemp. Phys.*, 1969, vol. 10, No 3 p. 221
15. R. M. Hill: *Electrical Transport in Thick Film Resistors*. *Proc. ISHM Ghent 1979*
16. R. M. Hill, T. J. Coutts: *Charge Transfer in Discontinuous Thin and Cermet Films*. *Thin Solid Films*, 1977, vol. 42, p. 201—213
17. R. M. Hill, A. K. Jonscher: *AC and DC Conductivity in Hopping Electronic Systems*. *Journal of Non-crystalline Solids*, 1979, vol. 32
18. R. Holm: *Electric Contacts*. Springer Verlag 1958
19. A. Kusy: *Chains of Conducting Particles. That Determine the Resistivity of Thick Resistive Films*. *Thin Solid Films*, 1977, 43, p. 243—250
20. A. Kusy: *Struktura, mechanizm przewodnictwa oraz szumy typu 1/f rezystywnych warstw grubych*. Rzeszów 1979
21. B. W. Licznarski: *Conductivity of Discontinuous Chromium Thin Films*. *Thin Solid Films*, 1976, 36, 61
22. B. W. Licznarski: *Conduction Mechanism's Identification in Thick Film Resistors*. *Proceedings of 5th European Hybrid Microelectronics Conference 1985, Stresa, Italy*
23. L. J. Maissel: *Electrical Properties of Metallic thin Films*. 13th Chapter of *Handbook of Thin Film Technology* Mc Graw Hill, New York, London, Toronto, 1970
24. A. F. Mayadas, M. Shatzkes: *Phys. Rev. B* 1, 1970, p. 1382
25. J. D. Morochow, L. J. Trusow, W. N. Łapowok: *Fizyczne jawnienia w ultra-dispersyjnych sriedach*. Moskwa, Energatomizdat, 1984
26. J. H. Mott, E. Dewis: *Elektronnyye procesy w niekrystaliczeskich wieszczestwach*, Moskwa, Izd. Mir, 1974
27. C. A. Neugebauer, M. B. Webb: *Electrical Conduction Mechanism in Ultrathin Evaporated Metal Films*. *J. Appl. Phys.*, 1962, 33, p. 74
28. L. Nordheim: *Ann. Physil*, 1931, 9, 607
29. S. Nowak: *Rezystory warstwowe. Problemy optymalizacji i technologii ze względu na stabilność cech elektrycznych*. Zeszyty Naukowe AGH, Automatyka 40, Kraków 1986
30. S. Nowak: *Stabilność termiczna i długoterminowa rezystorów grubowarstwowych*. *Elektronika* 1984, nr 9, s. 12—16
31. S. Nowak: *Zastosowanie funkcji regresji do analizy procesu technologicznego warstw grubych*. II Sympozjum „Mikroelektroniczna technika grubowarstwowa”, Rzeszów 1977
32. S. Nowak: *TCR Optimization of Thick Film Resistors by Use Second Order Regression Function*. *Acta Polytechnica*, Praha, 1987, 1, III, 1
33. S. Nowak; D. Luśniak-Wójcicka: *Influence of the Substrates on the Temperature*

Resistance Coefficient of Thick Film Resistor Composition. ISHM Poland Chapter Conference, Kraków 1982

34. S. Nowak, D. Luśniak-Wójcicka, M. Cież: *Układy mikroelektroniczne grubo-warstwowe.* Warszawa, WNT, 1977
35. G.E. Pike, C.H. Seager: *Electrical Properties and Conduction Mechanisms of Ru-Based Thick Film (Cermet) Resistors.* J. Appl. Phys. 48/12 December, 1977, p. 5152—5168
36. E.H. Sondheimer: *Mean Free Path of Electron in Metals.* Advan. Phys., 1952, 1, 1
37. R. Stratton: *Volt-Current Characteristics for Tunneling through Insulating Films.* J. Phys. Chem. Solids 1962, 23, p. 1177—90
38. B.S. Verma, S.K. Sharma: *Effect of Thermal Strains on the Temperature Coefficient of Resistance.* Thin Solid Films, 1970, 5, R. 44—46
39. R.W. Vest: *Conduction Mechanism in Thick Film Microcircuits.* Final Technical Report, Purdue University Research Foundation Grant Numbers DARC-15-70-67, DARC-15-73-68, Arpa Order: 1642 December, 1975
40. J.M. Ziman: *Electrons and Fonons.* Oxford University Press, London, 1960
41. O. Zinke: *Widerstände, Kondensatoren, Spulen und ihre Werkstoffe.* Springer Verlag, Berlin, Heidelberg, New York, 1965

S. NOWAK

EFFECT OF STRUCTURE AND MANUFACTURING ON THE TEMPERATURE FACTOR OF INTEGRATED THICK-FILM RESISTANCE SYSTEMS

Summary

The main effects of the structure and of the manufacturing process of thick-film resistive networks on their temperature coefficient (TCR), are analysed. An optimisation of the manufacturing process is carried out and a method of selection of the firing conditions of the film from the aspect of a minimum absolute value of TCR is presented. The possible thermal compensation of integrated dividers and attenuators made with the use of thick-film and thin-film procedures, is proven. Moreover, the mathematical descriptions of the conduction mechanism of thick and thin films are reviewed, from the aspect of temperature dependence, and their inadequacy is displayed.

S. NOWAK

INFLUENCE DES AGENTS CONSTRUCTIFS ET TECHNOLOGIQUES SUR LE COEFFICIENT DE TEMPÉRATURE DE LA RÉSISTANCE DES RÉSEAUX INTÉGRÉS RÉSISTIFS À COUCHES ÉPAISSES

Résumé

On a analysé les plus importantes influences de la construction et de la technologie sur le coefficient de température de résistance (CTR) des réseaux résistifs à couches épaisses. On a optimisé le processus de production et on a proposé la méthode du choix des conditions de cuisson de la couche vu la valeur absolue minimale de CTR. On a motivé la possibilité de la compensation thermique des diviseurs intégrés et des atténuateurs exécutés avec des techniques à couches épaisses et à couches minces. En outre on a passé en revue les descriptions mathématiques du mécanisme de la conductibilité des couches épaisses et minces dépendant à la température et on a montré l'inexactitude de ces descriptions.

S. NOWAK

EINFLUß KONSTRUKTIVER UND TECHNOLOGISCHER FAKTOREN AUF DEN TEMPERATURKOEFFIZIENTEN DER RESISTANZ INTEGRIERTER DICKSCHICHTIGER RESISTIVNETZE

Zusammenfassung

In der Bearbeitung wurden Einflüsse von Konstruktion und Technologie auf TKW der Dickschichtwiderstandsnetzwerke analysiert. Der Herstellungsprozeß wurde optimiert und eine Bestimmungsmethode für die Einbrandbedingungen in Anbetracht des minimalen TKW-Absolutwertes dargestellt. Es wurde auch die Möglichkeit der thermischen Kompensation integrierter Dick- und Dünnschichtteile sowie der mittels Dick- und Dünntechniken ausgeführten Dämpfungsglieder begründet. Darüber hinaus hat man einen Überblick über die mathematischen Beschreibungen des Leitungsmechanismus in Dick- und Dünnschichten in Bezug auf das Temperaturverhalten gebracht, und auf die Unvollkommenheit diesbezüglicher Beschreibungen hingewiesen.

С. НОВАК

ВЛИЯНИЕ КОНСТРУКЦИОННЫХ И ТЕХНОЛОГИЧЕСКИХ ПАРАМЕТРОВ НА КОЭФФИЦИЕНТ СОПРОТИВЛЕНИЯ ИНТЕГРАЛЬНЫХ РЕЗИСТИВНЫХ СЕТЕЙ

Резюме

Проанализированы основные влияния конструкции и технологии на температурный коэффициент сопротивления резистивных толсто пленочных сетей. Проведена оптимизация производственного процесса и представлен метод избрания условий выжигания слоя с точки зрения минимизации абсолютной величины Т.К.С.. Обоснована возможность термической компенсации интегральных делителей и демпферов равно толсто как и тонко пленочных. Кроме того дан обзор математических описей механизма провозждения толстых и тонких пленок с точки зрения температурной зависимости и констатировано несовершенство этих описей.

A model describing piezoresistive effects in thick film resistors

STANISŁAW PASZCZYŃSKI

*Wydział Elektryczny, Politechnika Rzeszowska, Rzeszów**Received 1988.01.28**Authorized 1988.09.05*

An analytical model of the piezoresistive effect in thick film resistors is presented. The model of the resistor structure has the form of a chain composed of MIM (Metal-Insulator-Metal) unit barriers and in which the thin glass layer is the insulator and the conductive medium grains are electrodes. As the main mechanism responsible for current flow through the barrier is a tunnelling process depending on the presence of metal ions diffused from contacts and an effective change of barrier width caused by the strain of the resistor. Model parameters cover certain parameters of the resistor structure and of its manufacturing process, for example: resistive film length (x_F), amount of conductive medium in the paste (m_c), ratio of elastic moduli of conductive medium and glass, time of resistor holding at peak temperature. The difference of the electron work functions of both media (Φ_0) and the barrier width (s_0) are the adjusted, unmeasurable model parameters. The model fits well two sets of experimentally gained characteristics $G_L(x_F, m_c)$ and $r_S(x_F, m_c)$ where G_L is the longitudinal strain gauge factor of resistance and r_S is the relative sheet resistivity of ruthenate resistors, made basing on prototype paste series PS with $\text{Bi}_2\text{Ru}_2\text{O}_7$ amount ranging from 20 to 60 wt. %. For the verification of the model the $G_L(x_F)$ and $r_S(x_F)$ characteristics of ruthenate resistors produced with use of Dupont HS8049 paste is adopted. Thick film Pd-Ag paste is used (DP6120) for making terminals. Optimum values of fitted parameters Φ_0 and s_0 are reasonable, in good agreement with those recently reported in papers concerning the conduction of similar materials.

1. INTRODUCTION

The piezoresistive effects in thick film resistors (TFRs) started to be a matter of interest at the beginning of the seventies [1]. It was found that the strain gauge factor of resistance (G) defined as the fractional change in resistance (r) divided by the resistor strain (ε) is considerably greater in TFRs than in metal-film resistors, and its stability with time under the operating conditions as well as temperature stability is far better than in discontinuous metal films [2-5]. The value of gauge factor of TFRs remains constant with strain changes from about 10^{-3} to -10^{-3} i.e. in the range where the irreversible changes in the resistor structure are not observed [6]. These positive, promising piezoresistive properties of TFRs have caused both an increase of interest in the origin of the phenomenon [7-11], and a possibility of application of TFRs in strain-sensing devices such as e.g. pressure or force sensors [11-15].

For the last ten years, a number of research efforts have been focused on the modeling of microstructure and conduction mechanisms in TFRs. Some review reports on this research have been recently published [16—18]. A conduction mechanism which could account for all the observed properties of TFRs still remains unknown, although a large number of the properties, especially their piezoresistive effect may be interpreted in terms of tunnelling process [8, 16].

In the present study we intend to build an analytical model describing the piezoresistive phenomenon in TFRs. To achieve this aim, we have constructed a model of TFR structure.

The proposed model comprises the influence of many technological and structural parameters of TFRs such as e.g., conductive medium content in paste, resistor length, time of resistor holding at peak temperature, ratio of elastic moduli of conductive medium and glass on longitudinal strain gauge factor (G_L). The unmeasurable model parameters were obtained after fitting theoretical G_L values to the measured ones. In order to prepare test resistors, we produced a prototype series of ruthenate pastes which make possible evaluation of all the TFRs parameters necessary to verify the model.

In Chapter 2 we present some techniques of measuring strain gauge factors. Chapter 3 discusses the experimental data and the premises relevant for constructing the model of TFR structure but details of the proposed analytical model describing longitudinal gauge factor of TFRs are presented in Chapter 4. Chapter 5 gives a description of experimental procedures and Chapter 6—the results of measurements and calculations. Chapter 7 contains a discussion of the results and conclusions.

2. MEASUREMENT TECHNIQUES

Since fractional change in resistance ($r = \Delta R/R$) versus strain in TFRs is linear and without hysteresis [5, 6, 14] then the strain gauge factor of resistance defined as [7].

$$G_e = \frac{r(\epsilon)}{\epsilon}, \quad (2.1)$$

shows the „intensity” of the piezoresistive phenomenon and it is a dimensionless coefficient, convenient for comparison of different resistive materials. Taking into consideration

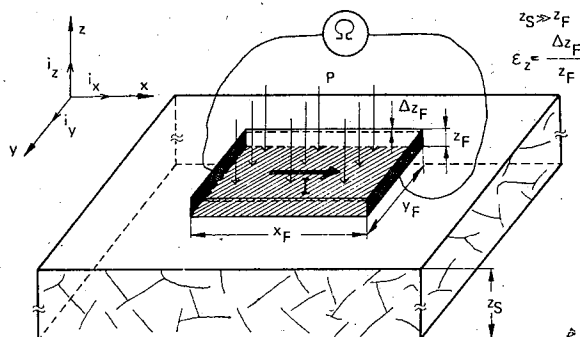


Fig. 1. Perpendicularly strained film resistor joined in xy -plane to ceramic substrate; P -pressure

some possible relations between directions of the strain and the current ($\vec{I}$) flowing through the film resistor joined to the ceramic substrate (Fig. 1) it is possible to define a few strain gauge factors [7, 19]:

$$\left. \begin{aligned} & \text{— longitudinal} && \text{— } G_L = r/\varepsilon_x \text{ if } \vec{I} \parallel \vec{i}_x \text{ and } \varepsilon = \varepsilon_x, \\ & \text{— transverse} && \text{— } G_T = r/\varepsilon_x \text{ if } \vec{I} \parallel \vec{i}_y \text{ and } \varepsilon = \varepsilon_x, \\ & \text{— perpendicular} && \text{— } G_p = r/\varepsilon_z \text{ if } \vec{I} \parallel \vec{i}_x \text{ or } \vec{I} \parallel \vec{i}_y \\ & \quad \text{(this case is shown in Fig. 1)} && \text{and } \varepsilon = \varepsilon_z, \\ & \text{— perpendicular} && \text{— } G_Q = r/\varepsilon_z \text{ if } \vec{I} \parallel \vec{i}_z \text{ and } \varepsilon = \varepsilon_z, \end{aligned} \right\} \quad (2.2)$$

(resistor contact films should be much the same as in a film capacitor).

If we assume that a resistive film is isotropic and if we take into account the fact that there is a rigid junction between the film and the ceramic substrate being significantly thicker than the resistive film, a following dependences describing the above defined gauge factors can be obtained [7, 11, 19]

$$\begin{aligned} G_L &= G_e(1-\nu_s-\nu_f)+1+\nu_s+\nu_f, \\ G_T &= G_e(1-\nu_s-\nu_f)-1-\nu_s+\nu_f, \\ G_p &= G_e-1, \\ G_Q &= G_e+1, \end{aligned} \quad (2.3)$$

where $G_e = \frac{1}{\rho} \frac{\partial \rho}{\partial \varepsilon}$ represents the so called „physical contribution” to the value of G_s dependent on the conduction mechanism which conditions the film resistivity ρ_s , ν_s and ν_f are Poisson's coefficient of the substrate and of the film respectively.

Assuming that $\nu_s = \nu_f = \nu$ and making use of Eq. (2.3) we receive

$$G_L - G_T = 2(1+\nu). \quad (2.4)$$

Expressions (2.3) and (2.4) are used to verify the rightness of the assumption concern the film isotropy. Since a typical value of G_L will be at least 3—4, thus, approximately, we can assume that strain gauge factors are proportional to G_e .

So far the literature has presented some methods of measurement of the longitudinal and transverse gauge factors [6, 8, 14, 20]. The principles of the most widely used three point bending method are shown in Fig. 2. The maximum deflection of the substrate (z_0)

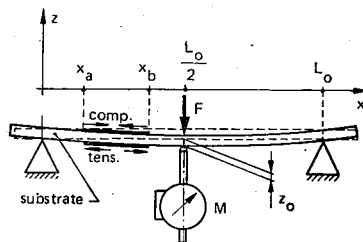


Fig. 2. Three point bending method cross-section for measuring strain gauge factors; M -dial micrometer, F -bending force

supported on parallel knife edges, with distance L_0 between them, is related to the strain $\varepsilon_x(x)$ at point x on the substrate surface by the expression [21, 22]

$$\varepsilon_x(x) = \frac{12xz_S z_0}{L_0^3} = cxz_0, \quad \text{for } z_0 < z_S \text{ and } x \leq L_0/2, \quad (2.5)$$

where $c = 12z_S/L_0^3$, z_S is substrate thickness, and strain sign is positive for tension and negative for compression.

Since the resistive film has a length $x_F = x_b - x_a > 0$ so its average strain $\bar{\varepsilon}_x$ is

$$\bar{\varepsilon}_x = \frac{\Delta x_F}{x_F} = \frac{1}{x_F} \int_{x_a}^{x_b} \varepsilon(x) dx = \frac{1}{2} c(x_a + x_b) z_0, \quad (2.6)$$

$$\text{for } 0 < x_a < x_b \leq \frac{L_0}{2}.$$

Making use of Eqs. (2.1), (2.2) and (2.6) we can obtain G_L and G_T factors as

$$G_L = \frac{r_x(z_0)}{\varepsilon_x(z_0)} \quad (\text{i.e. when contacts are parallel to } y \text{ axis}), \quad (2.7)$$

$$G_T = \frac{r_y(z_0)}{\varepsilon_x(z_0)} \quad (\text{i.e. when contacts are parallel to } x \text{ axis}). \quad (2.8)$$

The perpendiculars G_p and G_Q factors can be estimated by use of Eq. (2.3) or measured by adequate methods.

It is evident from Eqs. (2.5)–(2.8) that the accuracy of measurement of G_L and G_T factors is influenced mainly by errors of measurement of L_0 , z_S and z_0 since the resistance can easily be measured with higher accuracy than ± 0.01 percent. With the use of a properly made fixture which allows controlled three point substrate bending the total error in measured G_L and G_T values can be less than about 3–4 percent [6, 19].

An alternative method for the measurement of longitudinal and transverse gauge factors is shown in Fig. 3. One end of the resistor substrate is clamped and the opposite

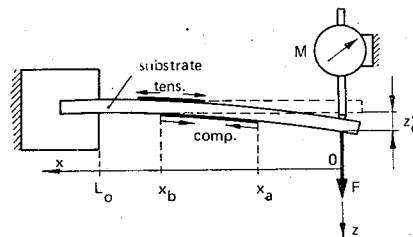


Fig. 3. Bending the substrate like a cantilever beam clamped at one end (cross-section); M —dial micrometer, F —bending force

end is deflected by a measurable distance (z'_0) by means of a dial micrometer. The expressions corresponding to Eqs. (2.5) and (2.6) are the following

$$\varepsilon_x(x) = \frac{3}{2} \frac{z_S x z'_0}{L_0^3} = c' x z'_0, \quad \text{for } x \leq L_0, \quad (2.9)$$

where $c' = 3z_S/2L_0^3$,

$$\bar{\varepsilon}_x = \frac{1}{2} c' z_0' (x_a + x_b) \quad \text{for} \quad x_a, x_b < L_0. \quad (2.10)$$

This method makes possible measurement of the gauge factors G_L and G_T with the total error less than about 3 percent [5, 8].

It should be noticed that the mentioned methods above of $\bar{\varepsilon}_x$ evaluation where previously experimentally checked with calibrated Ta_2N resistors [6] and silicon [8] strain gauges glued near the resistive films.

3. A MODEL OF THICK FILM RESISTOR STRUCTURE

Investigations on TFR structure carried out so far [16, 17, 23—30] have shown that it has the following features:

1. In general TFR is a multi phase ceramic with a high degree of dispersion in the media. One of them—glass in the form of a continuous matrix—joins the crystalline conductive particles together into one system with the ceramic substrate and contact films. In a majority of applications of TFRs an alumina ceramic substrate (96—99% Al_2O_3) with thickness from 0.5 to 1 mm is used. The thickness of the resistive films themselves are usually in the range of 10—25 microns.

2. The crystalline particles with low resistivity (10^{-2} — $10^{-5} \Omega\text{cm}$) dispersed in glass matrix are conductive oxides such as RuO_2 , IrO_2 , $\text{Bi}_2\text{Ru}_2\text{O}_7$, $\text{Pb}_2\text{Ru}_2\text{O}_{6.5}$, etc. Their sizes range from some nanometers to a few microns. The dispersion of the crystalline media in the film volume is not perfectly homogeneous. On a microscale, there can be observed some regions where conductive particles form natural clusters lengthwise the hypothetical boundaries of glass grains (i.e. such boundaries which could exist in the case when the film structure has been formed without melting the glass grains — an analogy of the results obtained by Inokuma et al [26] and Nishizawa et al [31] and Malliaris and Turner [32] is obvious). The phenomenon is also referred to as segregation or ordering.

3. The small spaces between conductive particles are filled, during the firing process, by melted glass which may „divide” clusters or aggregates of conductive powder due to the capillary forces. It is also possible to form necks between the particles during the firing process although the sintering process can be neglected in systems with particles greater than about 15 nm [30]. The results of investigations of the dependence of TFRs resistance (R) on the volume fraction of the conductive medium (v_c) shows a stabilized saturation when v_c tends to 1 (i.e. $dR/dv_c \rightarrow 0$ and $R \rightarrow R_{st}$) independently on TFR type. The R_{st} value is typically two or more orders greater than the value resulting from the conductive medium resistivity [9, 33—35]. Thus this proves that the sintering process of conductive particles is not intensive and the necks between particles does not condition TFRs properties. The resistance of resistors investigated by the Pike and Seager [16] decreased significantly (i.e. about 1—2 orders in dependence on their sheet resistivity) as their glass was leached away. Cattaneo and Prudenziati [36] showed that refiring processes caused no significant changes in the resistor properties. The attempt of modelling the $R(v_c)$ characteristic by the chains of series combination of the metal oxide particle and MIM

(Metal-Insulator-Metal) tunnelling barrier resistances [37] indicated that the MIM barriers determined the resistors resistance independently from the conductive-particle size. An analysis made by Prudenziati et al. [38] proved that in order to account for TFRs temperature characteristics it is enough to take into consideration the tunnelling MIM structures and their properties. Microscopic investigations of TFRs microstructure indicated that the mean distance between particles (s_0) is about 5–10 nm but one can believe that the electrical boundaries of the particles extend further than one can observe by the microscope. The observed solubility of the conductive medium in lead borosilicate glass (e.g. it is about 10^{-5} for RuO_2 at 800°C [39]) shows that the effective values of s_0 can be considerably below 5 nm.

4. The pores or microvoids are present as an important component of TFRs structure especially in the ruthenate resistors. The shape and size of the pores can differ within wide limits. The volume fraction of the pores increases with the increase in the volume fraction of crystalline particles [33] and this is in agreement with the general theory of sintering in such materials [40].

5. During the firing process, the metal migrates from the terminations into the resistive film and causes changes in the film properties. The migration process evolves most probably in three different phases [29]. Initially, the metal grains dissolve in the glass matrix of the resistor, then diffuse inside the film and eventually reprecipitate in smaller grains while the resistor is cooled down and concentration of metal ions in the film glass exceeds their solubility at room temperature. The most intensive migration process is observed for Ag based terminations. The diffusion of Ag from the terminations can be considered as diffusion from a constant source since silver solubility in lead borosilicate glass is high and equal to about 10 percent at room temperature [41]. It is important that the diffused silver does not react with the conductive crystalline particles in the film [29, 41].

6. Between the substrate and the resistive film there exists a transitional region — interface where dissolution of the substrate material (i.e. mainly Al_2O_3) in the film and

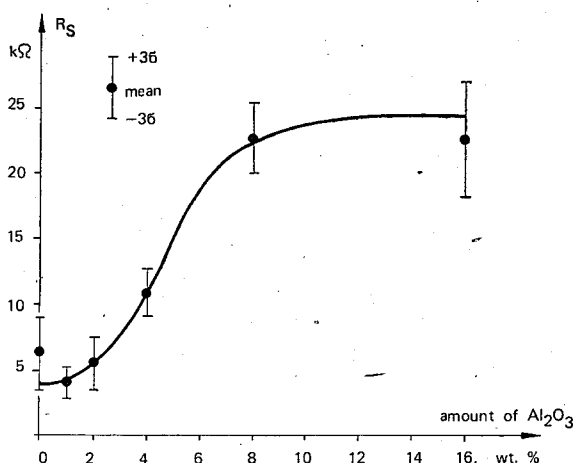


Fig. 4. Sheet resistance of prototype ruthenate TFRs as a function of amount of Al_2O_3 (i.e. powdered commercial alumina substrates) with constant content of bismuth ruthenate equal to 30 wt% (samples preparation technique was the same as described in Chapt. 5)

diffusion of glass components into the substrate is observed. Its thickness can be greater than ten microns [6, 42—44]. Since Al_2O_3 added to the resistive paste causes an increase in the resistor resistance [39] (see also Fig. 4) then the transitional region does not short the main part of the resistive film. This effect is likely to be responsible for an increase in the film resistivity with a decrease in its thickness below ~ 20 microns especially when the content of glass is relatively high [45].

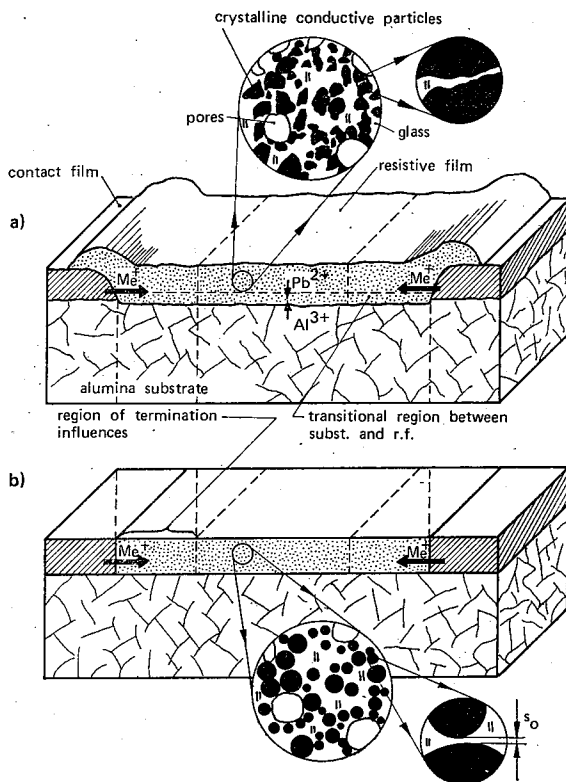


Fig. 5. Schematic cross-section of TFR showing the most important components of its macro- and micro-structure which influence electrical properties -(a) and its simplified form -(b)

Fig. 5a presents, schematic cross-section of TFR with the most important components of its structure which influence electrical properties. Fig. 5b shows the same scheme with the following simplifications:

- some deformations of the films near the contacts and the transitional region between the substrate and the resistive film are ignored,
- the irregularities in the film thickness are not taken into account either,
- the crystalline, conductive particles are replaced by ball-shaped particles of the same volume.

Basing on Fig. 5b a model of TFR structure was constructed (Fig. 6). It is characterized by the following features: —

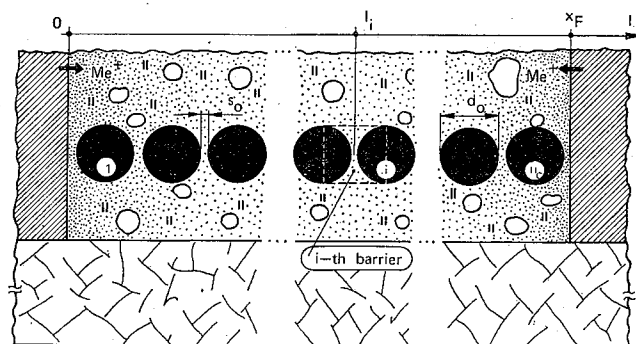


Fig. 6. Model of TFR structure showing one of the n_s parallel paths composed of n_c identical conductive particles separated by glass regions (dots represent concentration of ions diffused from terminations)

1. A conductive network of TFR structure is replaced by n_s of identical paths, parallel to current flow direction, (for the sake of clarity, we show only one path in Fig. 6) composed of n_c identical conductive ball-shaped particles, whose diameter is d_0 which corresponds to the value obtained from the powder surface area S_a measurement by B.E.T. method [46]. The electrical properties of the path are dependent on the properties of a unit barrier composed of two neighbouring halves of conductive particles and the glass region between them.

2. The distance between the neighbouring particles in the path (s_0) is constant, independent on the volume fraction of the conductive medium (v_c) and there is $s_0 \ll d_0$. (This assumption correlates with resistive film forming process in the case when the same powders were used in all considered pastes and final distance between conductive particles depends on powders form and initial stage of sintering process when glass is melting. In general conductive powder grains are agglomerates composed of many small crystalline particles — see also chapt. 3 p. 3). The distances between the neighbouring paths are significantly greater than s_0 thus electrical interactions between them can be ignored.

3. The volume fractions of the conductive medium and pores (v_p) are identical as in the real TFR and can be obtained from the paste composition and microscopic investigations of TFR structure. Some fractions of all conductive particles in the film form the conductive paths. It depends on many technological factors and segregation of conductive phase in the film. The pores are not situated in the region of barrier glass and their dispersion in the film is homogeneous.

4. The metal (Me^+ ions) migrates from the terminations into the film during the firing process by diffusion in the melted glass. This is the diffusion from constant source in the time t_p when the resistor remains at the peak temperature (in Fig. 6 the distribution of the diffused impurities is shown schematically by means of dots). We assume that the impurity in glass can be treated as an ionized donor although there can be many other mechanisms of the ions influence on barrier parameters e.g. they can change electrode electronic work function when located at interphase of crystalline grain-glass.

5. The n_c can be calculated as $n_c \cong x_F/d_0$ where x_F is the resistive film length. The resistor width remains constant.

The assumption concern the lack of electrical interactions between paths causes that the problem can be treated as one-dimensional one and thus the calculations are not so much complicated while simplification concern assumption about paralleism of current paths is reasonable because resistor electrical properties are the mostly dependent of the short-test paths with the lowest resistances.

4. ANALYTICAL MODEL OF LONGITUDINAL STRAIN GAUGE FACTOR

4.1. GENERAL EXPRESSIONS

The resistance R_{bi} of i -th unit barrier of the model shown in Fig. 6 is dependent mainly on s_0 , d_0 , a type of conduction mechanism of the barrier and the concentration of the impurity (Q_{oi}). To simplify the calculations, we assumed that the barriers with numbers $i = 1$ and $i = n_c + 1$ (i.e. the first and the last ones in the path), lienear the contacts, are identical to the others though one of their electrodes is flat. The resistance R_c of a single conductive path can be computed from the formula:

$$R_c = \sum_{i=1}^{n_c+1} R_{bi}(s_0, d_0, Q_{oi}). \quad (4.1)$$

If ε_g is the effective strain of the glass region in the unit barrier, then basing on Eqs. (1.1), (2.7) and (4.1) we can obtain the longitudinal gauge factor of the path (G_{LC}) as

$$G_{LC} = \frac{\sum_{i=1}^{n_c+1} \Delta R_{bi}[s_e(\varepsilon_g), d_0, Q_{oi}]}{R_c} \cong \frac{1}{\varepsilon} f_c(n_c+1, x_F, \varepsilon_g), \quad (4.2)$$

where

$$f_c(n_c+1, x_F, \varepsilon_g) = \frac{\sum_{i=1}^{n_c+1} \{R_{bi}[s_e(+\varepsilon_g), d_0, Q_{oi}] - R_{bi}[s_e(-\varepsilon_g), d_0, Q_{oi}]\}}{\sum_{i=1}^{n_c+1} \{R_{bi}[s_e(+\varepsilon_g), d_0, Q_{oi}] + R_{bi}[s_e(-\varepsilon_g), d_0, Q_{oi}]\}}, \quad (4.3)$$

and $s_e(\pm \varepsilon_g) = s_0(1 \pm \varepsilon_g)$, $Q_{oi} = Q_0(l_i)$, $l_i = x_F \frac{i-1}{n_c}$.

In Eq. (4.2) the influence of the geometrical effect dependent on the Poisson's coefficients has not been taken into consideration, and thus we assume that the shape of the barrier electrodes remain unchanged (the strain gauge factor of a metal film is close to one [2]).

In the case of an independency of barrier conductivity on barrier voltage drop there is $R_{bi} \sim 1/J_i$, where J_i is barrier current density and thus Eg. (4.3) can be easily changed to a form:

$$f_c(n_c+1, x_F, \varepsilon_g) = \frac{\sum_{i=1}^{n_c+1} \{1/J_i[s_e(+\varepsilon_g), d_0, Q_{oi}] - 1/J_i[s_e(-\varepsilon_g), d_0, Q_{oi}]\}}{\sum_{i=1}^{n_c+1} \{1/J_i[s_e(+\varepsilon_g), d_0, Q_{oi}] + 1/J_i[s_e(-\varepsilon_g), d_0, Q_{oi}]\}}. \quad (4.4)$$

It would be simple enough to prove that series or parallel combination of k such conductive paths does not change the G_{LC} value, thus G_{LC} is equal to the longitudinal strain gauge factor the TFR model. Considering a path in which k neighbouring barriers differ insignificantly in Q_0 i.e., $Q_0(l_i) \cong Q_0(l_{i+1}) \cong \dots \cong Q_0(l_{i+k})$ then, assuming that $(n_c + 1)/k = n_{ck}$ is an even number, the Eq. (4.2) can be reduced to

$$G_{LC} \cong \frac{1}{\varepsilon} f_C(n_{ck}, x_F, \varepsilon_g), \quad (4.5)$$

which includes the chain with barriers of numbers $i = kj$ where $j = 1, \dots, n_{ck}$.

Since the model deals with a rectangular resistor having a plane of symmetry — perpendicular to the direction of current flow so the number n_{ck} in Eq. (4.5) can be reduced to $n_{ck}/2$ to give

$$G_{LC} = \frac{1}{\varepsilon} f_C(n_{ck}/2, x_F, \varepsilon_g). \quad (4.6)$$

Let us assume that n_{ck} is constant for all the considered x_F values. In such a case it is easy to show that the resistance of the reduced path $R_{Ck}(x_F)$ is

$$R_{Ck}(x_F) = \sum_{i=1}^{n_{ck}/2} R_{bi}(s_0, d_0, Q_{0i}) = c_j \sum_{i=1}^{n_{ck}/2} [J_i(s_0, d_0, Q_{bi})]^{-1}, \quad (4.7)$$

where $c_j = \text{const}$, and may interpreted as a measure of the resistor sheet resistivity (R_s) and thus the dimensionless coefficient r_{SC} defined here as

$$r_{SC}(x_F) = \frac{R_{Ck}(x_F)}{R_{Ck}(x_{F_0})} \quad (4.8)$$

where x_{F_0} is the reference length, will be used simultaneously with G_{LC} for verification of the model.

4.2. THE CONDUCTION MECHANISM OF UNIT BARRIER

So far there are some conduction mechanism which are widely accepted for description of TFRs properties. These, the most possible conduction mechanisms which may occur in the unit barrier under consideration are: hopping transport and tunnelling [16, 17].

No reliable data has been presented so far to fully prove or to deny the hypothesis of the domination of hopping in TFR conduction [47, 48] but it is impossible to account for all the observable TFRs properties in terms of this mechanism. We suppose that the hopping „needs” a considerably greater degree of dispersion of the conductive medium nearly to the atomic scale. It should dominate in amorphous semiconductor [49] or heterogeneous systems with very high dispersion of the metallic medium [50, 51] (i.e. when a particle size is equal to about 1 nm) but such situations are not typical for TFRs.

So far, only tunnelling has been widely accepted as a mechanism which can be used for interpretation or modelling of TFRs characteristics [16, 37, 38, 52]. Thielemans et al [53] proved experimentally that the current of tunnelling MIM barrier is strongly dependent on the barrier thickness whereas Canali et al. [8] showed that the strain gauge factor of the tunnelling barrier can be compared to the gauge factors of TFRs. Since many other

TFRs properties are compatible with those of the MIM tunnelling barrier (e.g. nonlinearity of TFR voltage-current characteristic at high voltages [18]) we assume that tunnelling mechanism of MIM structure is adequate for a description of conduction in the unit of the model.

One should mention here that so far there were elaborated many models describing some electrical properties of resistive films which based on chains composed of region with different conduction mechanisms as current paths [16, 47, 54—56]. Some of the authors took into consideration apart from tunnelling, also hopping conduction in glass and metallic like conduction of grains [47], metallic and tunnelling conduction mechanisms only [16] or tunnelling and modified metallic conduction mechanisms of sintered contacts between adjacent grains [54]. In some papers more complicated models of resistive film structure have been created. They based on percolation theory and treated resistive films as percolative system [55, 56] or multipercolative system [57, 58]. In general these models describe resistivity (or resistance) and/or temperature coefficients of resistance versus concentration of conductive medium in the film and temperature characteristics. They include many fitparameters, difficult to measure but mostly they approximate very well experimentally obtained characteristics and allow to estimate the fractions of different conduction mechanisms in the resistive film. In this group of models one should mention Licznarski's generalized model [57, 58] which makes conduction mechanisms identification possible for any number of assumed mechanisms. None of these models takes into account influence of contact films on resistor properties and electro-mechanical interaction inside its structure.

Up till now some different expressions have been formulated to describe the tunnel current density (J_t) of an ideal tunnel junction composed of a simple rectangular potential barrier in the form of a thin dielectric film sandwiched between sheets of metal, versus the parameters of the barrier, its temperature (T) and voltage drop (V_b). Here one can mention the dependences obtained by Holm [59], Simmons [60] and Stratton [61] or expressions based on them as presented by Rowell [62] or Fisher and Giaever [63]. All the expressions are similar because they were obtained by means of WKB (Wentzel-Krammer-Brillouin) method. They cannot be used directly for the evaluation of J_t in analysed unit barrier since they are valid only for flat infinite electrodes.

In the unit barrier, particles-finite electrodes will be charged by flowing electron so that it should have some additional energy δE associated with a single electrode.

In our model we take into account an influence of δE on barrier tunnelling current density J_t by similar way as Pike and Seager have done [16]:

$$J_t = J_{ti} \left[\frac{1}{1 + \exp(\delta E/k_B T)} \right] \quad (4.9)$$

where expression in square brackets is Fermi factor connected with δE energy (k_B — Boltzman constant).

For calculating density of tunnelling current we used general formula [60, 61]:

$$J_{ti} = \frac{4\pi me k_B T}{h^3} \int_0^\infty P(E_x) \cdot \ln \left\{ \frac{1 + \exp[-(E_x/k_B T)]}{1 + \exp[-(E_x + eV_b)/k_B T]} \right\} dE_x, \quad (4.10)$$

where

$$\ln P(E_x) = \frac{-4\pi(2m)^{1/2}}{h} \int_{x_{1i}}^{x_{2i}} [\Phi_i(x) - E_x]^{1/2} dx, \quad (4.11)$$

and $P(E_x)$ — probability that an electron in an energy level E_x can penetrate the potential barrier between electrodes,

$\Phi_i(x)$ — barrier height at point x ,

x_{1i}, x_{2i} — limits of the barrier at Fermi level,

m — mass of tunnelling electron,

h — Planck's constant.

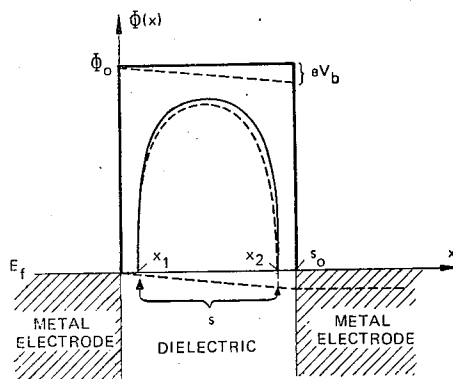


Fig. 7. Energy diagram for the symmetric metal-dielectric-metal structure showing influence of image forces and voltage drop across the barrier (for positive bias on right hand electrode) on barrier shape; E_f is Fermi level

We can find the barrier height Φ_i at point x (Fig. 7) from [60, 64—67]

$$\Phi_i(x) = \Phi_0 + p_i(x) + i_i(x) + q_{ci}(x) + q_{mi}(x), \quad (4.12)$$

where Φ_0 — theoretical barrier height resulting from the difference of the electrode work function and the electron affinity of the dielectric,

$p_i(x)$ — function including the influence of voltage v_b across the barrier [64],

$i_i(x)$ — function describing the image potential [64, 65],

$q_{ci}(x)$ — function describing the influence of constant space charge in the barrier, i.e. charge of ionized donors and acceptors. [65, 66],

$q_{mi}(x)$ — function showing the influence of a moving charge in the barrier volume [67].

In Eq. (4.12) we can ignore $q_{mi}(x)$ since Geppert [67] proved that changes in the barrier height caused by this component does not exceed about 10^{-18} eV. The other functions are the following [60, 64—66]

$$p_i(x) = \frac{-eV_b x}{s_0}, \quad (4.13)$$

$$i_i(x) \cong \frac{-e^2}{8\pi K K_0} \left\{ \frac{1}{2x} + \sum_{n=1}^{\infty} \left[\frac{ns_0}{(ns_0)^2 - x^2} - \frac{1}{ns_0} \right] \right\}, \quad (4.14)$$

$$q_{ci}(x) \cong -\frac{Q_{oi}s_0^2 e}{2KK_0} \left(-\frac{x^2}{s_0^2} + \frac{x}{s_0} + \frac{2K}{\sqrt{3}as_0} \right), \quad (4.15)$$

where $a = K_m/L_m$, K_m and L_m are permittivity and characteristic depth of potential penetration of electrode material respectively (typical values — $K_m \cong 1$, $L_m \cong 0.05$ nm [65]), K and K_0 are relative permittivity of dielectric and permittivity of free space respectively.

For evaluation of image potential one can use the same formula as for flat infinite electrodes (Eq. 4.14) because in our problem $s_0 \ll d_0$ and at neighborhood of tunnelling electron, electric field lines patterns are nearly identical in both systems.

In Eq. (4.14) we assumed that $Q_o(x) = Q_{oi}$ is constant since $s_0 \ll d_0 \ll x_F$. Recently Schmidlin [65] has shown that Eq. (4.15) being a solution to Poisson's equation is valid for $Q_{oi} \lesssim e/s_0^3 = Q_{oc}$ whereas for $Q_{oi} > Q_{oc}$, the strong local effect of the ions on the barrier shape (see Fig. 8c) and thus on the local current density can overlap and their resultant influence on the barrier current should be summed up for all the ions.

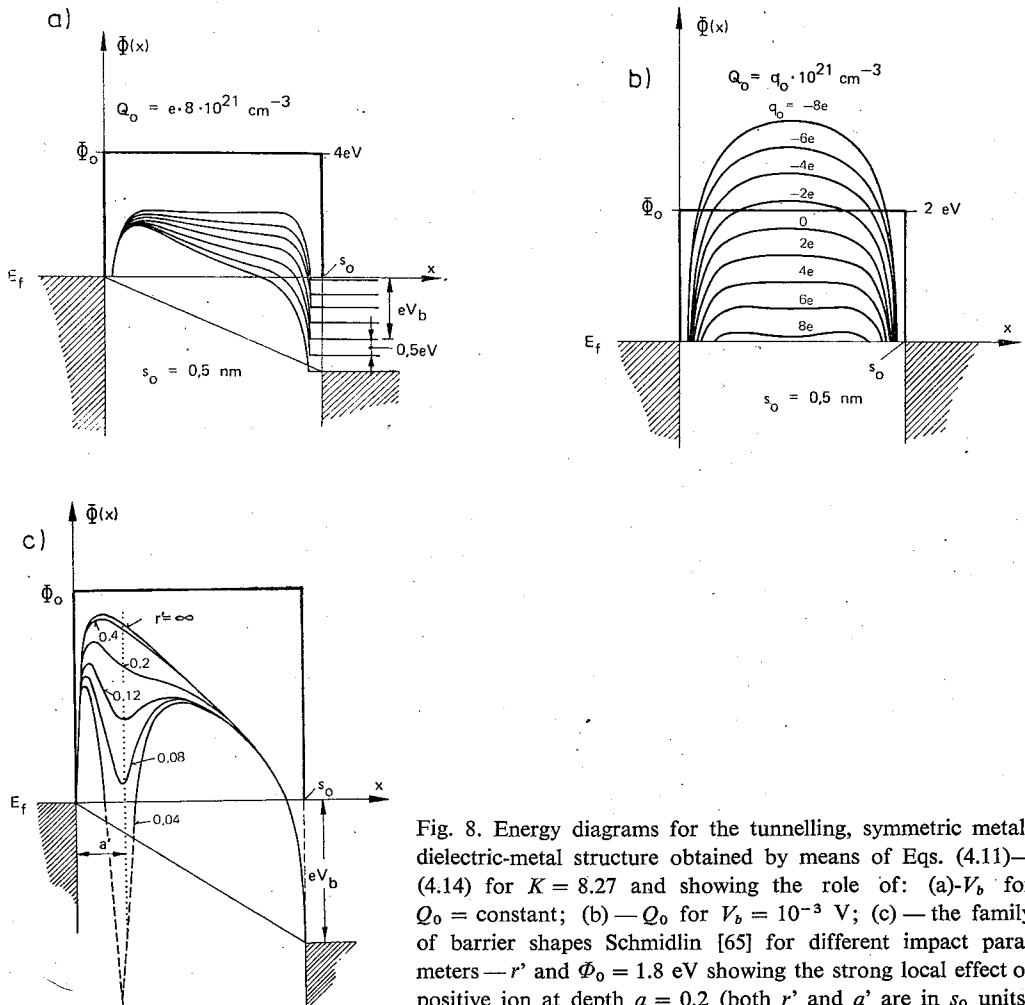


Fig. 8. Energy diagrams for the tunnelling, symmetric metal-dielectric-metal structure obtained by means of Eqs. (4.11)–(4.14) for $K = 8.27$ and showing the role of: (a)– V_b for $Q_o = \text{constant}$; (b)– Q_o for $V_b = 10^{-3}$ V; (c)–the family of barrier shapes Schmidlin [65] for different impact parameters— r' and $\Phi_0 = 1.8$ eV showing the strong local effect of positive ion at depth $a = 0.2$ (both r' and a' are in s_0 units)

Fig. 8 presents some possible shapes of the barrier indicating that its mean value can significantly differ from Φ_0 even by one order and therefore ignoring any component of Eq. (4.12), except q_{mi} , can cause an error in current density evaluation even by many orders.

So far, a convincing and reliable way of computing δE energy is unavailable [16, 68, 69]. For our purposes we have taken the expression formulated by Hill [68]

$$2\delta E = \frac{e^2}{8\pi K K_0 d_0} \left[1 + \left(\frac{d_0}{s_0} + 1 \right)^{-1} \right]. \quad (4.16)$$

This is a weak function of s_0 for $s_0 \ll d_0$ and it is very similar to the function used by Pike and Seager [16].

4.3. BARRIER STRAIN

We now turn to the evaluation of strain ε_g of the glass in the barrier region. Since the thickness of the substrate is much greater than the thickness of the resistive film and the elastic modulus of the alumina ceramic is at least a few times greater than the modulus of the film E_F [25] and, thus, an average strain of the film is practically equal to the strain of the substrate surface (ε). To evaluate ε_g , we based on the criterion of constant stresses in media forming the resistive film. At steady state conditions the stresses induced in all the media should be equal to the average stress in the film. Using Hook's law. [70] in a simple form we have

$$\left. \begin{aligned} \sigma_F &= \varepsilon E_F \\ \sigma_g &= \varepsilon_g E_g \end{aligned} \right\} \sigma_F = \sigma_g \quad \varepsilon_g = \frac{\varepsilon E_F}{E_g}, \quad (4.17)$$

where σ_F and σ_g are stresses in the film and the glass respectively, E_g is elastic modulus of the glass.

The elastic modulus of a porous film can be calculated from [71, 72]

$$E_F = E_{F0} \exp[-(h'_1 v_p + h'_2 v_p^2)], \quad (4.18)$$

where E_{F0} is elastic modulus of nonporous film h'_1 and h'_2 are positive constants, usually not exceeding 10, and v_p is pore volume fraction in the film. The research done by the author by an analysis of polished film surface photos [25, 28, 73] shows that a dependence of v_p on volume fraction of bismuth ruthenate v_c in TFRs is given by

$$v_p \cong b v_c^2, \quad (4.19)$$

where b is constant ranging from 0.5 to 1.2 in dependence on the firing process, v_c is ruthenate volume fraction evaluated with the assumption that $v_p = 0$.

A combination of (4.18) and (4.19) gives

$$E_F = E_{F0} \exp[-(h_1 v_c^2 + h_2 v_c^4)], \quad (4.20)$$

where $h_1 = h'_1 b$ and $h_2 = h'_2 b$.

At present, there are many models describing the elastic modulus of two phase heterogeneous materials having a structure similar to that of a resistive film but almost all of them, independently from their complexity, give, qualitatively identical results. For further

consideration we chose Counto's model [40, 74] used by the author earlier for ruthenate TFRs [28]

$$E_{F0} = E_g [1 - v_c^{0.5} + (v_c^{-0.5} - 1 + e_{gc})^{-1}]^{-1}, \quad (4.21)$$

where $e_{gc} = E_c/E_g$, E_c is elastic modulus of the conductive medium. Putting (4.21) into (4.20) and using (4.17) we eventually obtain

$$\varepsilon_g = \varepsilon [1 - v_c^{0.5} + (v_c^{-0.5} - 1 + e_{gc})^{-1}]^{-1} \exp[-(h_1 v_c^2 + h_2 v_c^4)]. \quad (4.22)$$

The volume fraction of the conductive medium v_c can be calculated basing on a paste composition as [25]

$$v_c = \left[1 + \frac{(100 - m_c) \delta_c}{m_c \delta_g} \right]^{-1}, \quad (4.23)$$

where δ_c and δ_g are densities of the conductive medium and glass respectively, m_c is amount, in weight percentage of the conductive medium in a two-component paste.

Eqs. (4.22) and (4.23) show that ε_g depends on m_c and ε only because e_{gc} , h_1 , h_2 , δ_c and δ_g are material constants.

4.4. INFLUENCE OF CONTACTS

Since, we have assumed that the metal ions diffuse from a constant source, then in a long resistive film the ion distribution near the contacts is given by the erfc type function [75],

$$Q_0(l) = N_0 \operatorname{erfc} L_D(l) \quad \text{for} \quad L_D(x_F) \gtrsim 8, \quad (4.24)$$

where $L_D(l) = 0.5 l (Dt_p)^{-0.5}$, l is a distance from the contact (see Fig. 6 — left hand contact), D is effective diffusion coefficient of ions in the film at the peak firing temperature (T_p), t_p is time of diffusion.

The unidirectional effective diffusion coefficient in the film is proportional to the fraction of the film's cross-section area occupied by molten glass which is equal to the volume fraction of glass v_g in the film [73]. Using Eq. (4.19) this fraction becomes

$$v_g = 1 - v_c - v_p \simeq (1 - v_c)/(1 + bv_c^2). \quad (4.25)$$

Since the diffusion inside the film is a three dimensional process, the coefficient D may be obtained from

$$D \simeq D_0 [(1 - v_c)/(1 + bv_c^2)]^3 \quad (4.26)$$

where D_0 is diffusion coefficient of the ions in molten glass.

In the case of a short resistors i.e. when $L_D(x_F) \lesssim 8$ the ion distribution has a more complex form [75]

$$Q_0(l) = N_0 \left\{ 1 - \frac{2}{\pi} \sum_{m=1}^{\infty} \left[\frac{1 - (-1)^m}{m} \right] \exp \left[- \left(\frac{\pi m}{x_F} \right)^2 Dt_p \right] \sin \left(\frac{m\pi l}{x_F} \right) \right\}. \quad (4.27)$$

This function converges very rapidly and in many cases a satisfactory accuracy is achieved even with the first term of the sum. Since $s_0 \ll d_0 \ll x_F$ the ion concentration in the i -th barrier may be evaluated by use of (4.27):

$$Q_{oi} = Q_0(l_i). \quad (4.28)$$

Eqs. (4.26)–(4.28) indicate that Q_{oi} is dependent on m_c and l_i only when x_F , D_0 , t_p , N_0 and b are fixed.

4.5. SEQUENCE AND CONDITIONS OF CALCULATIONS

In order to calculate the longitudinal strain gauge G_{LC} (Eq. 4.6) and coefficient r_{sc} (Eq. 4.8) by means of the above presented model the following sequence of calculations were used. Knowing the paste composition (i.e. m_c), the properties of component materia (i.e. d_0 , e_{gc} , δ_c , δ_g , $D_0(T_p)$, N_0), some properties of resistor structure (i.e. b , h'_1 , h'_2 , x_F) and t_p as the time of the resistor rests at the peak temperature during firing we can calculate v_c (Eq. 4.23), ε_g (Eq. 4.22) and D (Eq. 4.26). Having evaluated n_c and l_i (Eq. 4.3) we can obtain Q_{oi} by means of Eqs. (4.27) and (4.28) and next after assuming the Φ_0 , s_0 and V_b values, the function — $\Phi_i(x)$. The solution of the equation $\Phi_i(x) = 0$ give us x_{1i} and x_{2i} . Form Eqs. (4.16) we receive δE and with the help of Eqs. (4.11)–(4.9) value of J_i . Numerical computation of J_i repeated n_{ck} times gives G_{LC} (Eqs. 4.4 and 4.2) and r_{sc} (Eqs. 4.7 and 4.8). We took $V_b = 10^{-3} \text{ V} \ll \Phi_i/e$ and thus [60] we could use Eq. (4.4) for evaluation of f_c and next G_{LC} .

While fitting the model to the experimental data (this refers mainly to s_0 and Φ_0 values, the other earlier estimated parameters e.g. N_0 and D_0 , being difficult for exact evaluation have been slightly corrected) we minimized the objective function f_0 defined as

$$f_0 = \sum_{p_m} (G_{LC} - \bar{G}_L)^2 + \sum_{p_m} (r_{sc} - \bar{r}_s)^2, \quad (4.31)$$

where $\bar{G}_L$ and $\bar{r}_s$ are measured mean values of the longitudinal gauge factor and sheet resistivity ratio respectively, p_m is number of test resistor series.

5. EXPERIMENTAL PROCEDURE

For testing our model we have selected the prototype paste serial PS, elaborated at Technical University of Rzeszów, based on bismuth ruthenate and for comparison a commercial paste 8049 (a Dupont product) which is also based on bismuth ruthenate [76].

We have synthesized bismuth ruthenate by a solid state reaction method using RuO_2 and Bi_2O_3 powders. The obtained ruthenate powder had an average aggregate size $\sim 1.6 \mu\text{m}$ (determined by quantitative microscopy) and the surface area of $4.4 \text{ m}^2/\text{g}$ measured by the B.E.T. method using N_2 adsorption. This corresponds to the particle size (d_0) equal to about 150 nm. Thick film resistor paste compositions were made by mixing $\text{Bi}_2\text{Ru}_2\text{O}_7$ and glass powders together with vehicle of 3.5% ethyl cellulose and terpineol in an agate milling devices. The glass was a lead borosilicate composed of 65 wt% PbO — 25 wt% SiO_2 — 10 wt% B_2O_3 which had been milled to a typical particale size ranging

from 0.5 to 5 μm with the mean value of about 1.7 μm . We made five pastes with the content of bismuth ruthenate m_c varying from 20 to 60 wt% with 10 wt% step.

The test pattern for measuring G_L contained resistors of 0.5, 1, 3 and 6 mm in length, parameter x_a (see Fig. 2) was equal to 21, 20.5, 18.5 and 15.5 mm respectively and distance $L_0 = 46$ mm and for G_T evaluation — four identical resistors arranged perpendicularly. The resistors were printed^{a)} with selected pastes through 200 mesh screen on 2×0.6 in. alumina substrates (Kyocera product), 0.025 in. thick with prefired Pd/Ag (Dupont 6120) terminations. The resistors were dried at 150°C for 15 min before being fired in a 6-zone belt furnace with a typical resistor profile for ~ 700 s period at a peak of 850°C and the total furnace cycle time of 1 h.

The typical fired resistor resistance thicknesses were 12–18 μm and the batch quantity was 16. The resistor resistance was measured with YEW 2803 ohmmeter with $\pm 0.02\%$ accuracy. The strain gauge factors were measured by three point bending method (see Chapt. 2) with the use of a special fixture. The deflection of the substrate was measured with ± 0.5 percent accuracy with a MDNF-1a digital micrometer and resolution of 0.1 μm . Taking into consideration the errors resulting from the evaluation of the resistor position x_a , the distance L_0 and the substrate thickness (see Eq. (2.5) we estimated that the measured gauge factors are accurate within at least $\pm 2.5\%$, and thus the accuracy of their average values is about $\pm 0.6\%$.

6. RESULTS

The dependences of the longitudinal gauge factor (G_L) and the coefficient r_s (describing sheet resistivity changes related to R_s of a 0.5 mm long resistor) on the resistor length x_F and the content of bismuth ruthenate in the paste m_c as a parameter are shown in Figures 9^{b)} and 10. The solid curves in Fig. 9 and 10 are fits of the model to the experimental data with the objective function $f_0 = 1.589$ and following the parameters of the model: $b = 1$, $K = 8.27$, $d_0 = 150$ nm, $t_p = 700$ s, $e_{gc} = 2.1$, $D_0 = 1.2 \cdot 10^{-6} \text{ cm}^2\text{s}^{-1}$, $N_0 = 4e \cdot 10^{20} \text{ cm}^{-3}$, $h_1 = 6$, $h_2 = 8$, $s_0 = 1.096$ nm, $\Phi_0 = 1.12$ eV. Since our model displays some obvious imperfections, the computations were time-consuming, we decided to terminate the numerical fitting process when the difference between succeeding f_0 values was below 0.1. The values of b , K , d_0 and t_p were measured and fixed as model parameters. The values of D_0 , N_0 , h_1 , h_2 , e_{gc} were estimated with an accuracy of 10–40% and their starting values were $D_0 = 1.5 \cdot 10^{-6} \text{ cm}^2\text{s}^{-1}$, $N_0 = 5e \cdot 10^{20} \text{ cm}^{-3}$, $h_1 = 6$, $h_2 = 9$, $e_{gc} = 2.0$. The essential fitting process concerned the selection of the unit barrier width — s_0 and height — Φ_0 . Basing on Chapt. 3 we assumed the upper limits of these parameters as $s_0 \leq 2.5$ nm and $\Phi_0 \leq 3$ eV. Their starting values were $s_0 = 0.8$ nm, $\Phi_0 = 2$ eV. The estimations of D_0 and N_0 were made basing on measurements of silver content and potential distribution near the termination of TFRs with low m_c values (similarly as in

^{a)} The test resistor were prepared during the author's stay at E.S.A.T. Laboratories of Department Elektrotechniek, Katholieke Universiteit Leuven, Belgium.

^{b)} The resistors with $m_c = 60$ wt% and some with $m_c = 50$ wt% showed an irreversible change in resistance during G_L measurement appeared here in the standard deviation (σ) increase.

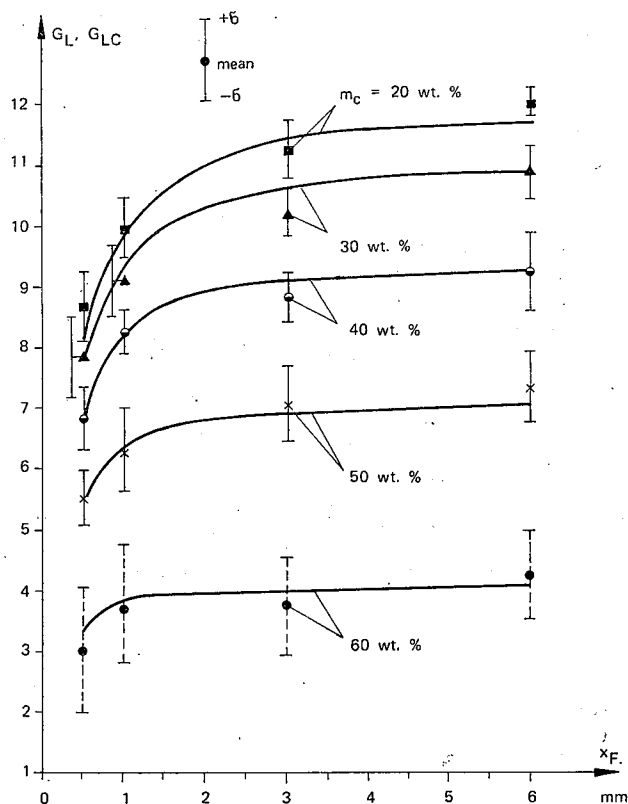


Fig. 9. Fits of the model (G_{LC} — solid curves) to the measured longitudinal strain gauge factors (G_L) vs resistor length (x_F) with bismuth ruthenate amount (m_c) as parameter

Cattaneo et al. [29]). The values of h_1 and h_2 were evaluated by means of b and Eq. (4.20) and by the microhardness of the resistive film with m_c ranging from 0 to 95 wt% and single $\text{Bi}_2\text{Ru}_2\text{O}_7$ crystal measurement.

The measured elastic modulus of the glass was in good agreement with the one calculated by an additive method.

The dependence of the sheet resistivity (R_s) and the longitudinal gauge factors, both measured (G_L) and calculated (G_{LC}), on the volume fraction of $\text{Bi}_2\text{Ru}_2\text{O}_7$ (v_c) for 6 mm long resistors is shown in Fig. 11. The shapes of these dependences are typical for ruthenate TFRs [8, 27].

The gauge factors were practically independent from the strain for $|\varepsilon| < 10^{-3}$ and the representative dependences of $G_L(\varepsilon)$ and $G_{LC}(\varepsilon)$ for 3 mm long resistors with $m_c = 30$ wt% are presented in Fig. 12.

Fig. 13 shows measured $G_L(x_F)$, $G_T(x_F)$ and $r_s(x_F)$ and theoretical $G_{LC}(x_F)$ and $r_{SC}(x_F)$ characteristics of 8049 resistors. The objective function equals 1.570 with the following model parameters: $b = 1$, $K = 8.27$, $d_0 = 100$ nm, $t_p = 700$ s, $m_c = 16$ wt%, $h_1 = 6$, $h_2 = 8$, $D_0 = 1.2 \cdot 10^{-6} \text{ cm}^2 \text{ s}^{-1}$, $N_0 = 4e \cdot 10^{20} \text{ cm}^{-3}$, $e_{gc} = 2.4$, $s_0 = 1.247$ nm, $\Phi_0 =$

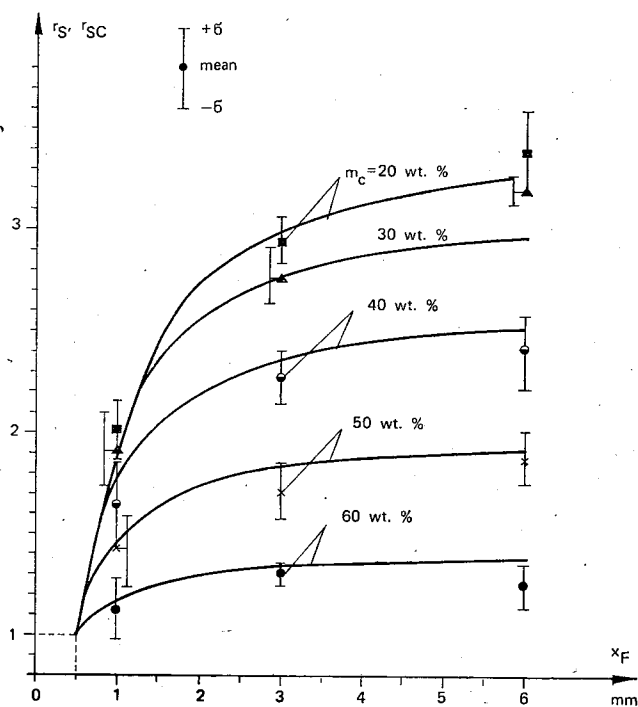


Fig. 10. Fits of the model (r_{SC} -solid curves) to the measured sheet resistivity ratios $r_s = R_s(x_F)/R_s(0.5)$ vs resistor length (x_F) with bismuth ruthenate amount (m_c) as parameters

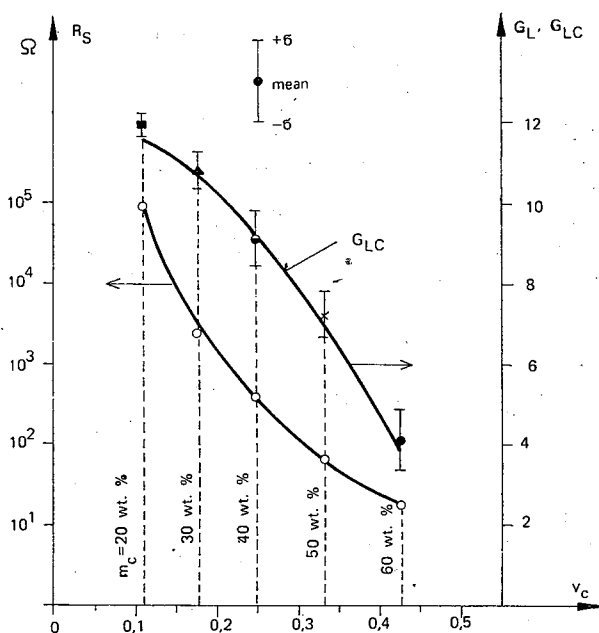


Fig. 11. Dependences of sheet resistivity (R_s) and longitudinal gauge factors — measured (G_L) and theoretical (G_{LC}) — on volume fraction of bismuth ruthenate (v_c)

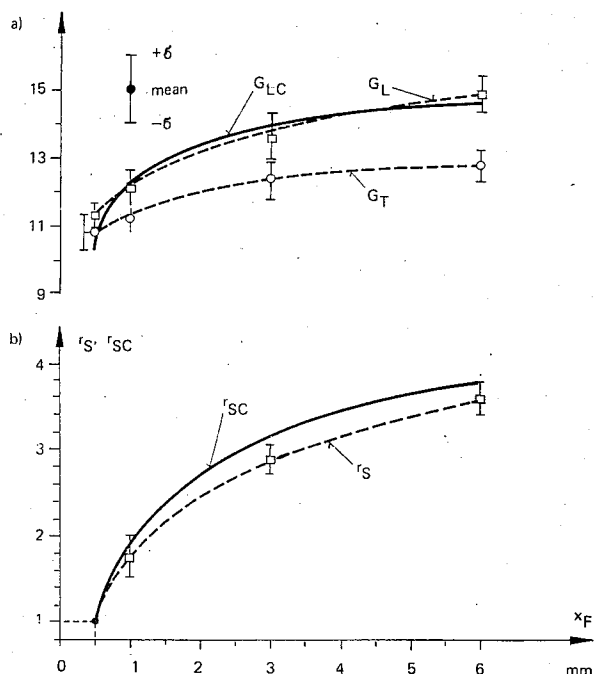


Fig. 12. Fit of the model (G_{LC} -solid curve) to the measured longitudinal strain gauge factor (G_L) vs resistor strain (ϵ) for resistors 3 mm long and with $m_c = 30$ wt%

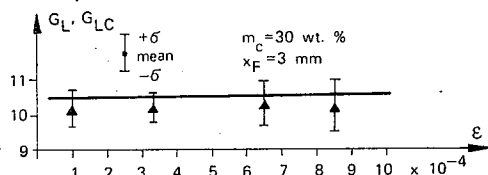


Fig. 13. (a) — longitudinal (G_L) and transverse (G_T) strain gauge factors and (b) — sheet resistivity ratio (r_s) as a function of resistor length for 8049 resistors. Solid lines are model fits to experimental data

= 1.195 eV. The starting values of fitted parameters were $e_{gc} = 2.0$, $\Phi_0 = 1$ eV, $s_0 = 1$ nm. The other parameters values were assumed and fixed during fitting process.

The shapes of the $G_T(x_F)$ characteristics of prototype PS resistors were similar to those shown in Fig. 17a for 8049 resistors and generally the differences $G_L - G_T$ increased with an increase in x_F from about 0.5–1 for $x_F = 0.5$ mm to about 2–3 for $x_F = 6$ mm.

We have found that the model is insensitive to diameter of conductive particles changes for $d_0 \gtrsim 5$ nm and three figures accuracy can be obtained with $n_{ck} \gtrsim 32$.

7. CONCLUSIONS

We have shown that the longitudinal strain gauge coefficient of resistance as well as the relative sheet resistivity of ruthenate TFRs can be obtained by considering the tunnelling transport between neighbouring conductive particles with the tunnel barrier height con-

ditioned by metal ion concentration caused by metal migration from terminations of the resistive film. The analysis of the results shows very reasonable figures for the involved physical quantities Φ_0 and s_0 from the model. Similar values of these parameters were obtained in other papers but it should be noted that the fitting of our model to the experimental data was obtained for a single set of parameters for all the investigated TFRs of PS series, i.e. for two different sets of characteristics, unlike the other models based on tunnelling process [37, 38]. Because of the specific TFRs structure properties (see Chapt. 3) resulting mainly, from the initial material properties and the firing process, it would be groundless to assume that TFRs which differ in volume fractions of conductive particles differ also in s_0 and Φ_0 .

Our data have demonstrated that the strain gauge factors of resistance in ruthenate TFRs are distinctly dependent on resistor length when Pd/Ag terminations are used. An increase in G_L when resistor length increases from 0.5 mm to 6 mm is about 30–35% almost independently from the resistor sheet resistivity, but in the case of $G_T(x_F)$ this increase is about 15–20%. At the same time we have observed a significant dependence of gauge factors on the amount of bismuth ruthenate in the film and the shape of $G_L(R_s)$ characteristic was close to the one recently presented [8, 9].

The measured dependences of TFRs sheet resistivity ratio on resistor length (Fig. 10 and 13b) indicate that it increases with resistor length increase even more than 3 times for resistors with low m_c values. These changes in R_s , like the above mentioned of gauge factors, get saturated for $x_F \gtrsim 6$ mm and their level decreases with an increase in m_c . Similar observations of the influence of Pd/Ag terminations on TFRs sheet resistivity were made by Naguib [77] and Stain et al [78]. The results of Fig. 13a agree well with Eq. (2.4) for $x_F \gtrsim 6$ mm (i.e. $G_L - G_T \approx 2.5$). Similar results were obtained for ESL 2900 resistors with Pd/Au terminations [8] used for confining the metal migration effect. All these results show that when the influence of terminations can be neglected the resistive film can be considered as an isotropic material from point of view of its piezoresistive properties.

The independence of G_L and G_T from strains (Fig. 12) and insensitivity of the model to conductive particles diameter changes are in agreement with some literature data confirming lack of correlation between G_L factor and the particle size [8] and the linearity of $r(\varepsilon)$ characteristics [6].

Generally, we have achieved a good agreement between the experiment and the theory though our model has some simplifications. It takes into account a considerable number of technological and structural TFRs properties. They can be either directly determined or measured (e.g. x_F , m_c , d_0 , e_{gc} , t_p) or estimated on the basis of TFRs structure (i.e. D_0 , N_0 , b , $h'_{1,2}$). Since the firing process parameters of TFRs (i.e. mainly t_p and T_p) influence some of the model parameters (e.g. D_0 , N_0 , b) we can expect that having determined these dependences we will be able to involve them into the model with fair result.

For calculation of the transverse gauge factor the model will be developed for the two — dimensional one. This means that we will take into consideration the electrical interaction between conductive paths and deformation of ball-shaped barrier electrodes under strain.

The efforts to utilize the above presented model for describing the temperature cha-

racteristics, time stability and non-linearity of current-voltage characteristics of TFRs make an important trend in the present day research in the field.

ACKNOWLEDGMENTS

The author thanks Prof. B. W. Licznarski from Technical University of Wrocław for many valuable and stimulating discussions especially those concerning tunnelling phenomena and model of resistor structure, Prof. W. Sansen and Dr. B. Puers from E.S.A.T. Laboratories, Katholieke Universiteit Leuven, Belgium, for their help in preparing the test samples and Prof. L. Trybus from Department of Automatics and Informatics, Technical University of Rzeszów, for providing facilities for making the necessary computer analysis.

REFERENCES

1. P. J. Holmes: *Changes in thick-film resistor value due to substrate flexure*. Microelectronics and Reliability, V. 12, 1973, pp. 395—396
2. G. R. Witt: *The electromechanical properties of thin films and the thin film strain gauge*. Thin Solid Films, V. 22, 1974, pp. 133—156
3. C. R. Tellier: *Thin metal film sensors*. Active and Passive Elec. Comp., V. 12, 1985, pp. 9-12
4. R. Dell'Acqua, G. Dell'Orto, A. Simonetta, C. Canali: *Long term stability of thick film resistors under strain*. Int. J. for Hyb. Microelectron., V. 5, 1982, pp. 82—85
5. C. Canali, D. Malavasi, B. Morten, M. Prudenziati, A. Taroni: *Strain sensitivity in thick film resistors*. IEEE Trans. on Comp., Hyb. and Manuf. Technol., CHMT-3, 1980, pp. 421—423
6. J. S. Shah: *Strain sensitivity of thick film resistors*. IEEE Trans. on Comp., Hyb. and Manuf. Technol., CHMT-3, 1980, pp. 554—564
7. B. Morten, L. Pirozzi, M. Prudenziati, A. Taroni: *Strain sensitivity in film and cermet resistors: measured and physical quantities*. J. Phys. D.: Appl. Phys., V. 12, 1979, pp. L51-L54
8. C. Canali, D. Malavasi, B. Morten, M. Prudenziati, A. Taroni: *Piezoresistive effects in thick-film resistors*. J. Appl. Phys., V. 51, 1980, pp. 3282—3288
9. P. F. Carcia, A. Suna, W. D. Childers: *Electrical conduction and strain sensitivity in RuO₂ thick film resistors*. J. Appl. Phys., V. 54, 1983, pp. 6002—6008
10. S. Paszczyński, J. Potencki, B. Rząsa: *Piezoresistive effects in thick-film structures*. Prace Naukowe ITE Politechniki Wrocławskiej, V. 30 1985, (Proc. 5th and 6th Conf. of the ISHM-Poland Chapter, Bystre 1984), pp. 103—114
11. B. Puers, W. Sansen, S. Paszczyński: *Assessment of thick-film fabrication methods for force (pressure) sensors*. Sensors and Actuators, V. 12, 1987, pp. 57—76
12. B. Puers, W. Sansen, S. Paszczyński: *Miniature highly sensitive pressure-force sensors using hybrid technology*. Proc. of the 6th European Hyb. Microelectronics Conf., Bournemouth-England, June 1987, pp. 416—420
13. M. Prudenziati, B. Morten, A. Taroni: *Characterization of thick-film resistor strain gauges on enamel steel*. Sensors and Actuators, V. 2, 1981/82, ss. 17—27
14. B. Rząsa, J. Potencki, S. Paszczyński: *Realisierungsmöglichkeiten von Temperatur- und Druckfühlern in Dickschichttechnik*. Wiss. Beiträge der IH Zwickau, V. 11, no 3, 1985, pp. 66—73
15. S. Paszczyński, B. Puers, W. Sansen: *Some new constructions of force (pressure) sensors based on piezoresistive effects in thick film resistors*. Proc. of 4th Czechoslovak Conf. on Microelectronics and Microsystem, Bratislava, 1986, pp. 143—150

16. G.E. Pike, C.H. Seager: *Electrical properties and conduction mechanisms of Ru based thick-film (cermet) resistors*. J. Appl. Phys. V. 48, 1977, pp. 5152—5169
17. M. Prudenziati: *Electrical transport in thick film (cermet) resistors*. Electrocomp. Sci. Technol., V. 10, 1983, pp. 285—293
18. S. Paszczyński: *On possibilities of synthesis of a useful model of thick film components*. Active and Passive Elec. Comp., V. 12, 1985, pp. 71—89
19. S. Paszczyński: *Piezorezystywne właściwości rezystorów grubowarstwowych*. Rozprawy Elektrotechniczne (in printing)
20. J.M. Darmon: *Further developments on thick film strain gauges applied to pressure transducers*. Presented at Transducer-Tempcon Conf., Harrogate-England, 1984 (unpublished)
21. S.P. Timoshenko, J.N. Goodier: *Theory of elasticity*. Mc Graw-Hill Book Comp., New York 1973, pp. 40—46
22. S. Timoshenko: *Strength of Materials: Part I — Elementary theory and problems*. Van Nostrand, New York 1958, pp. 143—146
23. M.V. Coleman: *Evaluation methods for the examination of thick film materials*. Radio and Electronic Engineering, V. 45, 1975, pp. 121—130
24. M.P. Ansell: *Conduction processes in thick film resistors. Part I and II*. Electrocomp. Sci. Technol., V. 3, 1976, pp. 131—151
25. S. Paszczyński: *Wpływ wybranych czynników technologicznych na właściwości warstw grubych z rutenianem bizmutu*. Doctoral dissertation, Warszawa 1979
26. T. Inokuma, Y. Taketa, M. Haradome: *The microstructure of RuO₂ thick film resistors and the influence of glass particle size on their electrical properties*. IEEE Trans. on Comp. Hyb., and Manufacturing Technol., CHMT-7, 1984, pp. 166—175
27. P.F. Carcia, A. Ferreti, A. Suna: *Particle size effects in thick film resistors*. J. Appl. Phys., V. 53, 1982, pp. 5282—5288
28. S. Paszczyński: *Niektóre aspekty wpływu tlenku kadmowego na właściwości rezystorów grubowarstwowych z rutenianem bizmutu*. Rozprawy Elektrotechniczne, V. 29, 1983, pp. 217—255
29. A. Cattaneo, M. Cocito, F. Forlani, M. Prudenziati: *Influence of the metal migration from screen and -fired terminations on the electrical characteristics of thick-film resistors*. Electrocomp. Sci. Technol., V. 4, 1977, pp. 205—211
30. P. Palanisamy, D.H.R. Sarma, R.W. Vest: *Liquid-phase sintering in thick-film resistor processing*. J. Amer. Cer. Soc., V. 68, 1985, pp. C215-C216
31. T. Nishizawa, M. Senna, H. Kuno: *Preparation electrical resistivity and its temperature dependence of RuO₂ — polystyrene composite powder*. J. Mater. Sci., V. 18, 1983, pp. 1346—1352
32. A. Malliaris, D.T. Turner: *Influence of particle size on the electrical resistivity of compacted mixtures of polymeric and metallic powders*. J. Appl. Phys., V. 42, 1971, pp. 614—618
33. S. Paszczyński: *Anwendung von Widerstandpasten in der Dickschichttechnik-ausgewählte Probleme*. Wiss. Beiträge der IH Zwickau, V. 6, no 4, 1980, pp. 70—74
34. P.F. Carcia, A. Suna: *A comparison of the conduction properties of RuO₂ and pyrochlore thick film resistors*. Proc. of the 33 rd Electron. Comp. Conf., Orlando—USA, 1983, pp. 280—283
35. E. Listkiewicz, A. Kusy: *Computer simulation of thick resistive films as two-component percolation systems: segregation of the conducting component*. Thin Solid Films, V. 130, 1985, pp. 1—15
36. A. Cattaneo, M. Prudenziati: *Effects of refiring processes on electrical and structural properties of thick-film resistors*. Electrocomp. Sci. Technol., V. 6, 1980, pp. 165—171
37. A. Szpytma: *Przewodnictwo elektryczne i szumy 1/f rezystywnych warstw grubych typu RuO₂ + szkliwo*. Doctoral dissertation, Wrocław 1986
38. M. Prudenziati, A. Rizzi, P. Davoli, A. Mattei: *Tunneling in thick-film (cermet) resistors and the minimum of resistance*. Il Nuovo Cimento, V. 2D, 1983, pp. 697—710
39. R.W. Vest: *The effects of substrate composition on thick film circuit reliability*. Final Technical Report, Contract No. N00019-78-C-0236, Purdue University 1979
40. W.D. Kingery, M.K. Bowen, D.R. Uhlman: *Introduction to ceramics*. J. Wiley and Sons, 1976

41. A. Ikegami, K. Shinriki: *Chemical reaction and electrical properties of amorphous RuO₂ — Ag — Glass ternary system*. Proc. European Hyb. Microelectron. Conf., Bad. Hamburg—Germany, 1977, pp. XVII 1—7
42. S. Kornaga: *Wpływ szkliva na wybrane właściwości rezystorów grubowarstwowych z rutenianem bizmutu*. Master thesis, Rzeszów 1983
43. W. R. Vest: *The effects of substrate composition on thick film circuit reliability*. Final Technical Report, Contract No. N00019-77-C-0327, Purdue University, 1978
44. G. Adie, B. Hołodnik, K. Pitt: *SEM/EDX analyses of some interactions between thick film resistors and dielectrics*. Microelectronics J., v. 15, no 2, 1984, pp. 38—43
45. A. Kusy, S. Paszczyński: *Wpływ grubości na rezystancję na kwadrat rezystywnych warstw grubych*. Proc. of the 2nd Microelectron. Thick Film Technol. Symp., Rzeszów 1977, pp. 106—119
46. S. Brunauer, P.H. Emmett, E. Teller: J. Amer. Chem. Soc., V. 60, 1938, p. 309
47. N.C. Halder, R.J. Snyder: *Measurement of the tunneling and hopping parameters in RuO₂ thick films*. Electrocomp. Sci. Technol., V. 11, 1984, pp. 123—136
48. C. Jacoboni, A. Rizzi: *Electrical conductivity of thick-film resistors*. J. Appl. Phys., V. 54, 1983, pp. 5852—5857
49. V. Ambegaokar, S. Cochran, J. Kurkijarvi: *Conduction in random systems*. Phys. Review B., V. 8, 1973, pp. 3682—3688
50. Ping Sheng, B. Abeles, Y. Arie: *Hopping conductivity in granular metals*. Phys. Rev. Letters, V. 31, 1973, pp. 44—47
51. M. Mostefa, G. Olivier: *Quantum size effect for hopping conduction in granular metals*. J. Phys. C: Solid State Phys., V. 18, 1985, pp. 93—99
52. T.M. Chen, S.F. Su, D. Smith: *1/f noise in Ru-based thick-film resistors*. Solid State Electronics, V. 25, 1982, pp. 821—827
53. M. Thielemans, V. Leo, R. Deltur, M. Membod: *Influence of a mechanical modulation on electrical transport in point contact junctions*. J. Appl. Phys. V. 50, 1979, pp. 5841—5846
54. A. Kusy: *An equivalent network for resistance and temperature coefficient of resistance versus temperature and composition of thick films*. J. Appl. Phys., V. 62, 1987, pp. 1324—1334
55. A. Kusy, E. Listkiewicz: *Percolation-network for thick resistive films*. Solid State Electronics, (in printing)
56. A. Kubovy: *A percolation model of the conduction threshold in thick-film resistors: segregated structures*. J. Phys. D.: Appl. Phys., V. 19, 1986, pp. 2171—2183
57. B.W. Licznarski: *Conduction mechanisms identification in thick-film resistors*. Proc. of the 5-th European Hyb. Mic. Conference, Stresa-Italy, 1985, pp. 389—398
58. B.W. Licznarski, H. Pagnia, N. Sotnik: *Conductivity model of cermets treated as critical percolative systems*. Phys. Stat. Sol. (a), V. 104, 1987, pp. 891—902
59. R. Holm: *The electric tunnel effect across thin insulating films in contacts*. J. Appl. Phys., V. 22, 1951, pp. 569—579
60. J.G. Simmons: *Generalized thermal J-V characteristic for the electric tunnel effect*. J. Appl. Phys., V. 35, 1964, pp. 2655—2658
61. R. Stratton: *Volt-current characteristics for tunnelling through insulating films*. J. Phys. Chem. Solids, V. 23, 1962, pp. 1170—1190
62. J.M. Rowell: *Tunnelling anomalies-experiment*, in *Tunneling Phenomena in Solids*. (editors: E. Burstein, S. Lundquist), Plenum Press. New York, 1969, pp. 385—404
63. J.C. Fisher, I. Giaever: *Tunnelling through thin insulating layers*. J. Appl. Phys., V. 52, 1961, pp. 172—177
64. J.G. Simmons: *Image force in Metal-Oxide-Metal tunnel junctions*, in *Tunneling Phenomena in Solids*. (editors: E. Burstein, S. Lundquist), Plenum Press, New York, 1969, pp. 135—148
65. F.W. Schmidlin: *Enhanced tunnelling through dielectric films due to ionic defects*. J. Appl. Phys., V. 37, 1966, pp. 2823—2832

66. M. Shatzke: *Injection into insulators in the presence of space charge*. J. Appl. Phys., V. 49, 1978, pp. 4868—4872
67. D. V. Geppert: *Theoretical Shape of metal-insulator-metal potential barriers*. J. Appl. Phys., V. 34, 1963, pp. 490—493
68. R. M. Hill: *Electrical conduction in discontinuous metal films*. Contemp. Phys., V. 10, 1969, pp. 221—240
69. P. Sheng, B. Abeles: *Voltage-induced tunnelling conduction in granular metals at low temperatures*. Phys. Rev. Lett., V. 28, 1972, pp. 34—37
70. A. G. Guy: *Introduction to materials science*. McGraw-Hill Book Comp., New York, 1971
71. R. Pampuch: *Zarys nauki o materiałach — materiały ceramiczne*. PWN, Warszawa, 1977, pp. 117—128
72. J. C. Wang: *Young's modulus of porous materials: Part 1 — Theoretical derivation of modulus-porosity correlation*. J. Mater. Sci., V. 19, 1984, pp. 801—808
73. R. T. De Hoff, F. N. Rhines: *Quantitative microscopy*. McGraw-Hill Book Comp., 1971, pp. 45—75
74. O. J. Cunto: *Mag. Concrete Res.*, V. 16, 1964, p. 129 (cited from L. Holiday, *Composite materials*. Elsevier Publ. Comp., Amsterdam 1966)
75. S. K. Ghandi: *The theory and practice of microelectronics*. J. Wiley and Sons, New York 1968, pp. 78, 447—465
76. J. M. Wheeler: *Thick film conductors and resistors on dielectris for high reliability applications*. Proc. 5th European Hyb. Microelectron. Conf., Stresa—Italy, 1985, pp. 51—59
77. H. M. Naguib: *The characterization of thick film resistors terminated with Pd/Ag conductor of various compositions*. Proc. of the 1977 International Microelectron. Symp. Baltimore—USA, 1977, pp. 48—59
78. S. J. Stein, C. Huang, A. S. Gelb: *A high voltage high performance thick film resistor system*. Electrocomp. Sci. Technol., V. 4, 1977, pp. 95—104

S. PASZCZYŃSKI

MODEL OPISUJĄCY EFEKTY PIEZOREZYSTYWNE W REZYSTORACH GRUBOWARSTWOWYCH

Streszczenie

W pracy przedstawiono analityczny model zjawiska piezorezystywnego w rezystorach grubowarstwowych. Model struktury rezystora ma postać łańcucha złożonego z jednostkowych barier typu MIM (Metal — Izolator — Metal), w których izolatorem jest cienka warstwa szkliwa a elektrodami ziarna ośrodka przewodzącego. Założono, że mechanizmem decydującym o przepływie prądu przez barierę jest proces tunelowania uwarunkowany obecnością jonów metalu dyfundujących z kontaktów rezystora oraz efektywną zmianą szerokości bariery wywołaną wydłużeniem rezystora. Parametrami modelu są niektóre parametry struktury i procesu technologicznego rezystora takie jak na przykład: długość warstwy rezystywnej (x_F), zawartość ośrodka przewodzącego w paście (m_c), stosunek modułów sprężystości ośrodków tworzących warstwę, czas wytrzymania warstwy rezystywnej w temperaturze szczyłowej. Niemierzalnymi, dopasowywanymi parametrami modelu są różnica prac wyjścia elektronu obu ośrodków (Φ_0) i szerokość bariery (s_0). Model dobrze aproksymuje dwie rodziny otrzymanych eksperymentalnie charakterystyk typu $G_L(x_F, m_c)$ i $r_s(x_F, m_c)$ gdzie G_L jest podłużną czułością odkształceniową rezystancji a r_s względną rezystancją na kwadrat rezystorów rutenianowych, do wykonania których wykorzystano prototypową serię past PS o zawartości $\text{Bi}_2\text{Ru}_2\text{O}_7$ zmienianej w zakresie 20—60% wag. Do weryfikacji modelu posłużyły również wyniki badań charakterystyk $G_L(x_F)$ oraz $r_s(x_F)$ testowych rezystorów rutenianowych wykonanych z pasty produkcji firmy Dupont typu HS8049. Do wykonania kontaktów zastosowano pastę typu Pd-Ag (DP6120). Optymalne wartości parametrów dopasowywanych Φ_0 i s_0 są wiarogodne i zbliżone do uzyskiwanych w innych, dotyczących przewodnictwa podobnych materiałów, pracach.

S. PASZCZYŃSKI

MODÈLE DÉCRIVANT LES EFFETS PIÉZORÉSISTIFS DANS LES RÉSISTANCES À COUCHES ÉPAISSES

Résumé

Dans le travail on a présenté le modèle analytique du phénomène piézorésistif dans les résistances à couches épaisses. Le modèle de la structure de la résistance a la forme d'une chaîne composée de barrières isolées du type MIM (Métal — Isolateur — Métal), dans lesquelles l'isolation est constituée d'une mince couche de glaçure et les électrodes — de grains du milieu conducteur. On a admis que le mécanisme, décidant du parcours du courant par la barrière est l'effet tunnel, dépendant de la présence des ions de métal se diffusant des contacts de la résistance et du changement effectif de la longueur de la barrière, dû à l'allongement de la résistance. Les paramètres du modèle sont définis par certains paramètres de structure et par ceux du processus technologique de la résistance, tels que: longueur de la couche résistive (x_F), teneur du milieu conducteur dans la pâte (m_c), rapport entre les modules d'élasticité des milieux, dont est composée la couche, durée de la résistance de la couche résistive en température maximale. La différence des travaux de la sortie de l'électron de deux milieux (Φ_0) et la largeur de la barrière (s_0) sont les paramètres non mesurables et ajustés. Le modèle permet d'obtenir une approximation de deux familles de caractéristiques, obtenues expérimentalement du type $G_L(x_F, m_c)$ et $r_s(x_F, m_c)$, où G est la sensibilité longitudinale de déformation de la résistance et r_s la résistance relative sur le carré des résistances de ruthénate à l'exécution, desquelles on a employé une série prototype des pâtes PS à teneur de $\text{Bi}_2\text{Ru}_2\text{O}_7$ variant dans l'étendue de 20—60% de poids. Pour la vérification du modèle on s'est servi aussi de résultats d'études des caractéristiques $G_L(x_F)$ et $r_s(x_F)$ des résistances test de ruthénate, exécutées avec la pâte produite par la firme Dupont du type HS8049. Les contacts ont été faits avec la pâte du type Pd—Ag(DP6120). Les valeurs optimales des paramètres ajustés Φ_0 et s_0 s'avèrent être voisines et rapprochées à celles obtenues dans d'autres essais, se rapportant au problème de la conductibilité des matériaux similaires.

S. PASZCZYŃSKI

MODELL FÜR DIE BESCHREIBUNG PIEZORESISTIVER EFFEKTE IN DICKSCHICHTWIDERSTÄNDEN

Zusammenfassung

Im Beitrag wird ein analytisches Modell einer piezoresistiver Erscheinung in Dickschichtwiderständen dargestellt. Das Modell der Widerstandsstruktur hat die Gestalt einer Kette, die aus Einheitsbarrieren der Art MIM (Metall-Isolator-Metall) besteht, wobei die Isolierung durch eine dünne Glasurschicht und die Elektroden durch Körner einer Leitungsschicht abgebildet werden. Es wurde angenommen, daß der Stromfluß durch die Barriere durch den Tunnelmechanismus bedingt wird. Dieser Mechanismus basiert auf der Anwesenheit von Metallionen, die von den Kontaktstellen des Widerstandes defundieren, und auf der effektiven Änderung der Barrierenbreite, die zufolge der Ausdehnung des Widerstandes entsteht. Die Modellparameter werden mittels mancher Strukturparameter und Parameter des technologischen Prozesses des Widerstandes widerspiegelt, wie zum Beispiel: Länge der Widerstandsschicht (x_F), Gehalt des Leitmediums in der Paste (m_c), Verhältnis der die Schicht bildenden Absolutwerte der Längeelastizität der Medien, Zeit der Erhaltung der Widerstandsschicht in der Spitzentemperatur. Die nicht meßbaren Parameter des Modells, die erst angepaßt werden müssen, sind Differenz der Ausgangsarbeiten des Elektrones der beiden Medien (Φ_0) und Barrierenbreite (s_0). Das Modell approximiert zufriedenstellend zwei Sätze der experimentell ermittelten Kennlinien der Art $G_L(x_F, m_c)$ und $r_s(x_F, m_c)$, wobei G_L die Längsverformungsempfindlichkeit des Widerstandes und r_s der relative Widerstand pro Quadrat der Ruthenian-Widerstandes sind. Zur Herstellung dieser Widerstände wurde eine Prototypserie der PS-Paste mit einem $\text{Bi}_2\text{Ru}_2\text{O}_7$ -Gehalt angewandt, der im Gewichtsbereich von 20 bis 60% verändert wird. Zur Überprüfung des Modells dienten auch die Untersuchungsergebnisse der Kennlinien $G_L(x_F)$ und $r_s(x_F)$, als auch der Rutheniantestwiderstände, die aus der Paste HS8049 der Firma Dupont hergestellt wurden. Zur Ausführung der Kontaktstellen wurde die Paste Pd-Ag (DP6120) angewandt. Die optimalen Werte

der angepaßten Parameter Φ_0 und s_0 sind glaubwürdig und korrespondieren mit ähnlichen Werten, die in der Literatur zitiert werden, und die sich mit der Problematik der Leitfähigkeit derartiger Werkstoffe befaßt.

С. ПАЦИНСКИ

МОДЕЛЬ ОПИСЫВАЮЩАЯ ПЬЕЗОРЕЗИСТИВНЫЕ ЭФФЕКТЫ В ТОЛСТОПЛЕНОЧНЫХ РЕЗИСТОРАХ

Резюме

Представлена аналитическая модель пьезорезистивного эффекта в толсто пленочных резисторах. Модель структуры резистора имеет форму цепи состоящей из барьеров типа МИМ (Металл-Изолятор-Металл) в которых изолятором является тонкий слой стекла а электродами — проводящими зернами. Принято что механизм решающий о протекании тока через барьер является процессом туннелирования обусловленным присутствием ионов металла диффузирующих из контактов резистора и эффективным изменением ширины барьера будущим результатом деформации резистора. Параметрами модели являются некоторые параметры структуры и технологического процесса резистора такие как например: длина резистивной плёнки (x_F), содержание проводящей среды в пасте (m_c), пропорция модулей упругости проводящей среды и стекла, время выдерживания резистивной пленки в максимальной температуре. Неизмеряемыми, подгоняемыми параметрами модели являются разница работ выхода электрона из проводящей среды и стекла (Φ_0) и ширина барьера (s_0). Модель хорошо аппроксимирует две группы получаемых экспериментально характеристик типа $G_L(x_F, m_c)$ и $r_s(x_F, m_c)$ где G_L — продольная чувствительность к деформационным сопротивлениям а r_s удельное плёночное сопротивление рутенатовых резисторов при изготовлении которых использована прототипная серия паст PS содержащих $\text{Bi}_2\text{Ru}_2\text{O}_7$ с допискам в переделе 20—60 процент по весу. Для верификации модели использованы также результаты испытаний характеристик $G_L(x_F)$ и $r_s(x_F)$ рутенатовых резисторов изготовленных из пасты фирмы Дипонт типа HS8049. В качестве контактов применены плёнки типа Pd-Ag(DP6120). Оптимальные величины параметров подгоняемых Φ_0 и s_0 достоверны и сравнимы с цитированными в литературе описывающей проводимость похожих материалов.

Modelowanie termicznego sprzężenia zwrotnego w monolitycznych układach scalonych przy pomocy programu SPICE-2G

KRZYSZTOF ZAMŁYŃSKI

Instytut Podstaw Elektroniki, Politechnika Warszawska

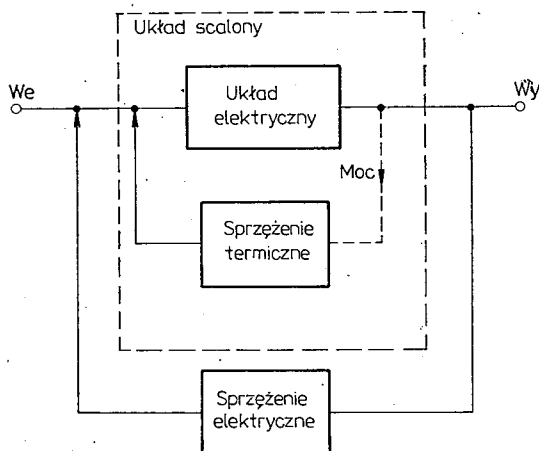
Otrzymano 1988.11.26

Autoryzowano do druku 1989.01.11

W artykule przedstawiono metodę modelowania zjawiska termicznego sprzężenia zwrotnego przy pomocy uniwersalnego programu SPICE-2G. Metodę tę zilustrowano na przykładzie wyznaczenia statycznej przejściowej charakterystyki monolitycznego wzmacniacza mocy UL1498.

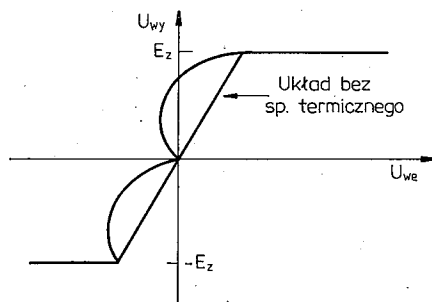
1. WSTĘP

Efekt termicznego sprzężenia zwrotnego jest zjawiskiem typowym dla układów scalonych. Występuje wtedy, gdy moc wydzielana w niektórych elementach układu (np. tranzystory stopni wyjściowych), a tym samym temperatura wszystkich elementów układu zależą od wartości sygnału wyjściowego. Pojawia się wówczas napięcie niezrównoważenia o wartości także zależnej od sygnału wyjściowego (rys. 1.). W przypadku wzmacniaczy



Rys. 1. Sprężenia zwrotne w układzie scalonym

prądu stałego (np. operacyjnych, mocy, komparatorów) efekt termicznego sprzężenia zwrotnego powoduje charakterystyczne odkształcenie statycznej charakterystyki przejściowej $U_{wy} = f(U_{we})$ układu. Przykładową charakterystykę przejściową przedstawiono na rysunku 2. Kształt charakterystyki ze sprzężeniem termicznym wynika z zależności wy-



Rys. 2. Wpływ sprzężenia termicznego na stałoprądową charakterystykę przejściową wzmacniacza

dzielonej w układzie mocy od napięcia wyjściowego i jej wpływu na temperaturę stopni wzmacniających. Charakterystyka przedstawiona na rys. 2 jest typową charakterystyką dla wzmacniacza prądu stałego przy założeniach, że dla każdego punktu rozkład temperatury jest ustalony oraz, że moc wydzielana się tylko w tranzystorach wyjściowych pracujących w klasie B. Dla takiego układu moc wydzielana, a tym samym temperatury poszczególnych elementów układu są zależne od kwadratu napięcia wyjściowego.

Pojawienie się napięcia niezrównoważenia spowodowane jest przez zmianę średniej temperatury układu (czyli zmianę temperatury pracy całego układu) oraz przez różnice temperatur poszczególnych elementów na płytce krzemowej układu scalonego. W zależności od konstrukcji układu (elektrycznej — schematu elektrycznego i geometrycznej — rozmieszczenie elementów na płytce krzemowej) jedno z tych zjawisk może być dominujące. Na przykład wpływ zmiany średniej temperatury pracy układu jest istotny dla układów o źle skompensowanych termicznie punktach pracy, lub dla takich, dla których wzmocnienie zależy od temperatury. Wpływ gradientów temperatury jest istotny w układach termicznie zrównoważonych, w których elementy kompensujące nie leżą na izotermach pola termicznego wytworzonego przez elementy grzejące.

Zjawisko termicznego sprzężenia zwrotnego należy więc uwzględniać w fazie projektowania układu scalonego. W dobrze zaprojektowanym układzie (takimi są praktycznie wszystkie seryjnie produkowane układy) zjawisko to nie powinno mieć wpływu na pracę układu. Znana jest jednak niestabilność wzmacniacza operacyjnego $\mu A741$ przy braku pętli sprzężenia elektrycznego spowodowana właśnie efektami termicznymi [3].

2. ETAPY MODELOWANIA

SPICE-2G jest uniwersalnym programem przeznaczonym do symulacji obwodów elektrycznych. Pozwala on przeprowadzać nieliniową analizę stałoprądową, nieliniową analizę stanów nieustalonych i liniową analizę częstotliwościową. Obwody mogą zawierać: rezystancje, pojemności, indukcyjności, indukcyjności wzajemne, niezależne źródła na-

pięciowe i prądowe, cztery typy źródeł sterowanych (nieliniową transmitancją opisaną wielomianami) i następujące elementy półprzewodnikowe: diody, tranzystory bipolarne, tranzystory polowe-złączowe i MOS.

Od temperatury uzależnione są odpowiednie parametry modeli elementów półprzewodnikowych i rezystancje, z tym, że temperatura traktowana jest jako temperatura otoczenia i wprowadzanie innej wartości temperatury (przy pomocy kart lub instrukcji sterujących) zmienia jednocześnie odpowiednie parametry wszystkich elementów. Ta właściwość programu nie tylko komplikuje modelowanie układów ze sprzężeniem termicznym, ale narzuca konieczność modelowania wielkości termicznych poprzez wielkości elektryczne.

Przy modelowaniu właściwości termicznych schematem elektrycznym przyjęto następującą analogię elektrotermiczną:

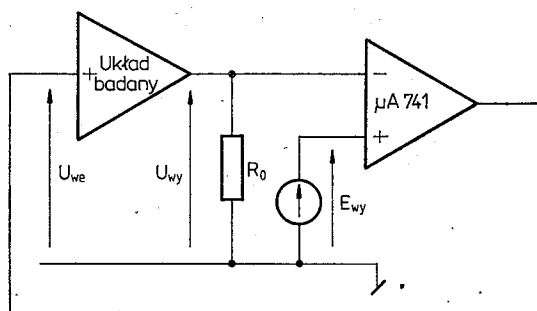
napięcie U [V] — T [K] różnica temperatur

natężenie prądu I [A] — P [W] moc grzejnika

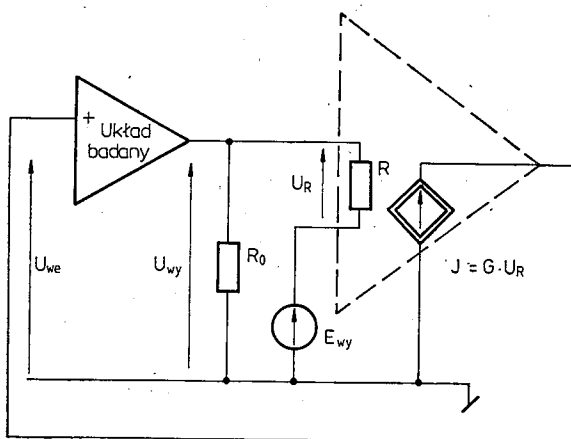
rezystancja elektryczna R [Ω] — R_{TH} [K/W] rezystancja termiczna

Prawo Ohma $U = R \times I$ — $T = R_{TH} \times P$ prawo Fouriera

Dla wzmacniaczy z termicznym sprzężeniem zwrotnym zależność U_{wy} od U_{we} jest



Rys. 3. Schemat układu do pomiaru charakterystyki przejściowej



Rys. 4. Komputerowy model układu z rys. 3

najczęściej niejednoznaczna — rys. 2. W celu umożliwienia doświadczalnego wyznaczenia pełnej charakterystyki należy wyznaczyć zależność odwrotną, to znaczy $U_{we} = f(U_{wy})$. Możliwe jest to w układzie z aktywną pętlą sprzężenia zwrotnego przedstawionym na rys. 3. W układzie tym napięcie wyjściowe U_{wy} ustalane jest za pomocą napięcia E_{wy} , gdyż $U_{wy} = E_{wy}$.

Komputerowy model tego układu przedstawiono na rys. 4. Wzmacniacz operacyjny został zamodelowany jako układ składający się z dużej rezystancji wejściowej — R (rzędu $M\Omega$) i prądowego źródła o dużej transkonduktancji wzajemnej ($G = 0.1\text{--}1S$) sterowanego z rezystancji wejściowej. Ten sposób modelowania pozwolił na uzyskanie zbieżności przy modelowaniu komputerowym przy pomocy programu SPICE-2G.

Analiza zjawiska sprzężenia termicznego powinna uwzględniać następujące procesy fizyczne:

- wydzielenie się mocy w elementach układu
- wzrost temperatury płytki krzemowej, przy czym w ogólnym przypadku należy uwzględnić rozkład temperatury na płycie
- zmianę pod wpływem temperatury elektrycznych parametrów elementów.

Uwzględnienie w procesie modelowania mocy wydzielonych we wszystkich elementach jest niecelowe — w praktyce tylko w niektórych elementach wydziela się dużą moc. Również parametry tylko niektórych elementów zasadniczo wpływają na działanie całego układu. Uzasadniony jest więc wybór elementów — nazywanych dalej elementami termicznie czynnymi — które powinny być uwzględnione przy modelowaniu sprzężenia termicznego.

Uwzględniając opisane procesy fizyczne, założenia i możliwości programu SPICE-2G modelowanie właściwości statycznych układu z termicznym sprzężeniem zwrotnym można podzielić na następujące etapy:

1. Wyznaczenie elementów termicznie czynnych
 - 1.1 Wyznaczenie elementów, w których wydzielają się duże moce
 - 1.2 Wyznaczenie elementów wrażliwych termicznie
2. Bezpośrednie wyznaczenie statycznego rozkładu temperatury na płycie krzemowej — czyli wyznaczenie ilościowych zależności między mocami wydzielanymi w elementach grzejących a temperaturą elementów.
3. Modelowanie elektryczne efektów wydzielania się ciepła w układzie
 - 3.1 Modelowanie transmisji ciepła — od grzejników do elementów wrażliwych termicznie
 - 3.2 Modelowanie zmian charakterystyk elementów wrażliwych termicznie od temperatury.
4. Dołączenie modelu termicznego do modelu elektrycznego.

Wyznaczenie charakterystyki przejściowej całego układu.

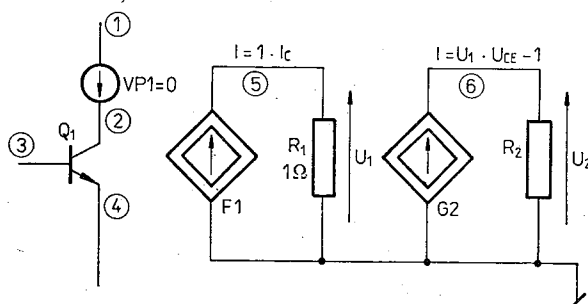
Punktem wyjściowym do analizy wpływu sprzężenia termicznego jest model układu bez tego efektu. Należy znać parametry elementów, punkty pracy tranzystorów oraz charakterystykę przejściową układu.

2.1. WYZNACZENIE ELEMENTÓW TERMICZNIE CZYNNYCH

Elementy grzejące wybiera się przez określenie wielkości mocy w nich wydzielonej. Elementy, w których wydziela się moc niezależna od sygnału zmieniają tylko temperaturę

pracy układu. Elementy, w których moc zmienia się w zależności od sygnału wpływają na kształt charakterystyki przejściowej.

Przykładowy układ próbkowania mocy w tranzystorze npn przedstawiono na rys. 5.



Rys. 5. Układ próbkowania mocy w tranzystorze npn

Wynika on z właściwości programu SPICE-2G, w którym źródło sterowane prądowo musi być sterowane prądem płynącym przez niezależne źródło napięciowe oraz z faktu innego oznaczania źródła sterowanego prądem i źródła sterowanego napięciem. W języku SPICE-2G zapis układu z rys. 5 jest następujący:

```
Q1  2 3 4 MODNPN
VP1 1 2 DC 0.
R1   5 0 1.
F1   0 5 VP1 1.
R2   0 6 1.
G2   0 6 POLY(2) 5 0 4 2 0. 0. 0. 0. 1.
```

Napięcie na rezystorze R1 jest liczbowo równe prądowi płynącemu przez tranzystor Q1. Wydajność źródła prądowego G2 jest równa iloczynowi prądu tranzystora i napięcia U_{CE} tranzystora czyli równa mocy wydzielonej w tranzystorze. W przypadku $R2 = 1\Omega$ napięcie na tym rezystorze jest liczbowo równe mocy wydzielonej w tranzystorze. W przypadku, gdy R2 jest liczbowo równe rezystancji termicznej tranzystor Q1 — otoczenie — napięcie na R2 jest liczbowo równe przyrostowi temperatury (w stosunku do otoczenia) tranzystora Q1.

Kryterium wyboru elementów wrażliwych termicznie powinna być wrażliwość bezwzględna napięcia wejściowego U_{WE} (układ z aktywną pętlą sprzężenia zwrotnego — rys. 3) na zmianę parametru danego elementu spowodowane mocą P_{str} wydzieloną w elementach grzejących czyli

$$\frac{\delta U_{WE}}{\delta P_{str}} = \frac{\delta U_{WE}}{\delta par} \cdot \frac{\delta par}{\delta P_{str}}. \quad (1)$$

Takie kryterium jest niemożliwe do zastosowania przed zamodelowaniem, gdyż jest to inaczej sformułowany cel całego modelowania. Możliwe jest natomiast wyznaczenie wrażliwości napięcia wejściowego na zmianę określonego parametru uzależnionego od temperatury T :

$$\frac{\delta U_{we}}{\delta T} = \frac{\delta U_{we}}{\delta par} \cdot \frac{\delta par}{\delta T}. \quad (2)$$

Przy pomocy programu SPICE-2G można wyznaczyć pierwszy czynnik wzoru (2). Drugi czynnik należy wyznaczyć, dla konkretnego elementu, z analitycznej zależności opisującej zmianę odpowiedniego parametru od temperatury. Wiąże się więc on z opisem termicznym właściwości elementów. Biorąc pod uwagę klasę analogowych układów scalonych, przy modelowaniu statycznej charakterystyki przejściowej należy uwzględnić dwa rodzaje elementów: tranzystory bipolarne i rezystory.

Dla rezystorów przyjmuje się, że rezystywność zależy liniowo od temperatury, stąd

$$R = R_0(1 + \alpha \Delta T) \quad (3)$$

czyli

$$\frac{\delta R}{\delta T} = R_0 \cdot \alpha. \quad (4)$$

Współczynnik α zależy od rodzaju rezystora i od gęstości domieszek. Dla rezystorów bazowych wartość współczynnika wynosi od 1×10^{-3} do 2.5×10^{-3} [1/K].

Metoda wyboru tranzystorów wrażliwych termicznie wynika z możliwości modelowania zmian parametrów tranzystora pod wpływem temperatury. W modelu tranzystora istniejącym w programie SPICE-2G wraz z temperaturą zmieniają się następujące parametry stałoprądowe:

1. U_T — potencjał elektrotermiczny, automatycznie podstawiany przez program jako:

$$U_T = \frac{kT}{q}, \quad (5)$$

gdzie $k = 1.3806226 \cdot 10^{-23}$ [J/K] — stała Boltzmanna

$q = 1.6021918 \cdot 10^{-19}$ [C] — ładunek elektronu

2. I_s — prąd nasycenia:

$$I_s(T) = I_s(T_0) \frac{T^3}{T_0^3} \exp \left\{ \frac{E_g \cdot q}{T} \left\{ \frac{1}{T_0} - \frac{1}{T} \right\} \right\} \quad (6)$$

gdzie T_0 [K] — temperatura odniesienia

$E_g = 1.11$ [V] przerwa energetyczna dla krzemu

Ponadto w modelu Gummela-Poona wraz z temperaturą zmieniają się współczynniki nieidealnej składowej prądu bazy C_2 i C_4 :

3.

$$C_2(T) = \frac{C_2(T_0)}{\left\{ \frac{I_s(T)}{I_s(T_0)} \right\} \left(1 - \frac{1}{N_E} \right)} \quad (7)$$

4.

$$C_4(T) = \frac{C_4(T_0)}{\left\{ \frac{I_s(T)}{I_s(T_0)} \right\} \left(1 - \frac{1}{N_C} \right)}, \quad (8)$$

gdzie N_E i N_C są odpowiednio rzeczywistymi współczynnikami emisji B-E i B-C

zaś $\frac{I_s(T)}{I_s(T_0)}$ można wyznaczyć z równania (6).

W programie SPICE-2G nie jest modelowana zmiana współczynnika wzmocnienia prądowego β_0 wraz z temperaturą.

Ponieważ dla tranzystorów w układzie scalonym praktycznie niemożliwe jest określenie parametrów Gummela-Poona, przyjęto do opisu tranzystora model Ebersa-Molla. Jak podano wcześniej, w programie niemożliwa jest niezależna zmiana temperatury poszczególnych tranzystorów. Niemożliwa też jest automatyczna zmiana parametrów U_T i I_s .

Przyjęto więc, że o własnościach termicznych decyduje (łatwe do modelowania) przesunięcie charakterystyki wejściowej tranzystora, będące efektem zmian parametrów U_T i I_s od temperatury.

Dla tranzystorów wzór (2) jest następujący:

$$\frac{\delta U_{WE}}{\delta T} = \gamma \frac{\delta U_{WE}}{\delta U_{BE}} \quad (9)$$

gdzie

$$\gamma \stackrel{\text{df}}{=} \frac{\delta U_{BE}}{\delta T} \quad (10)$$

Wartość współczynnika γ można łatwo wyznaczyć z modelu Ebersa-Molla. Przyjmując, że tranzystor znajduje się w stanie aktywnym oraz, że zmiany punktu pracy są pomijalnie małe (dla tranzystorów wyjściowych — będących zazwyczaj tranzystorami termicznie wrażliwymi — jest to spełnione), otrzymujemy:

$$\gamma = \left. \frac{\delta U_{BE}}{\delta T} \right|_{I_E = \text{const}} = \frac{K}{q} \ln \left\{ \frac{I_E}{I_s(T_0)} \frac{T_0^3}{T^3} - 3 \right\} - \frac{E_g}{T_0} \quad (11)$$

Ponieważ w programie SPICE-2G wielkość U_{BE} nie występuje jako parametr tranzystora, w celu wyznaczenia czynnika $\frac{\delta U_{WE}}{\delta U_{BE}}$ ze wzoru (9) należy w emiterze tranzystorów wstawić niezależne źródło $VE = 0$ i wyznaczyć zależności $\frac{\delta U_{WE}}{\delta VE}$.

2.2. WYZNACZENIE ROZKŁADU TEMPERATURY

Wyznaczenie rozkładu temperatury na całej strukturze krzemowej nie jest konieczne dla metody modelowania będącej tematem artykułu. Dlatego, poniżej ograniczono się do kilku ogólnych uwag na temat wyznaczania rozkładu temperatury.

Rozkład temperatury zależy od rozmieszczenia elementów grzejących i od konstrukcji obudowy układu scalonego. Można rozróżnić dwie zasadniczo różne konstrukcje obudowy: obudowa małej mocy i obudowa dużej mocy, czyli z dobrze przewodzącym ciepło radiatorem.

Metody wyznaczania rozkładu temperatury można podzielić na metody eksperymentalne i metody obliczeniowe. Metody eksperymentalne są mało przydatne z uwagi na ich raczej jakościowy charakter (detekcja promieniowania podczerwonego, zastosowanie ciekłych kryształów) lub niemożliwość dostępu do końcówek elementów w większych układach scalonych w celu użycia ich jako czujników pomiaru temperatury.

Metody obliczeniowe polegają na rozwiązaniu (analitycznym, numerycznym lub przez modelowanie) równania przewodnictwa cieplnego. Metoda rozwiązania analitycznego jest bardzo przydatna, ale może być stosowana tylko dla układów w obudowach dużej mocy. Dla tego typu obudowy dopuszczalne są znaczne uproszczenia warunków brzegowych umożliwiających rozwiązanie analityczne.

Metody numeryczne oraz modelowania elektrotermicznego polegają na podziale obszaru, w którym szukana jest temperatura, na mniejsze podobszary — czyli na zastąpieniu równania różniczkowego — różnicowym.

W metodzie pierwszej równanie różnicowe rozwiązywane jest metodami numerycznymi np: metodą siatek, lub metodami probabilistycznymi.

W metodzie modelowania elektrotermicznego korzysta się z analogii elektrotermicznej. Poszczególne dyskretne podobszary zastępuje się przez rezystory o wartości zależnej od rozmiarów podobszarów i rezystywności termicznej materiału. Źródła mocy modelowane są przez źródła prądowe, zaś napięcie poszczególnych węzłów równoważne jest temperaturze odpowiednich punktów obszaru. Takie postępowanie nie tylko pozwala na uzyskanie informacji o rozkładzie temperatury, ale jest jednocześnie konstruowaniem elektrycznego modelu właściwości termicznych struktury. Jednak praktyczne wykorzystanie tak skonstruowanego modelu jest niemożliwe ze względu na zbyt dużą liczbę węzłów (zwłaszcza przy modelowaniu w trzech wymiarach) wnoszonych przez utworzoną sieć.

Proponowana dalej metoda modelowania właściwości termicznych obudowy rezygnuje z symulacji rozkładu temperatury na całej płytce i dlatego pozwala na uzyskanie elektrycznego modelu o minimalnej liczbie węzłów.

2.3. ELEKTRYCZNY MODEL TERMICZNY UKŁADU WRAZ Z OBUDOWĄ

Należy zsyntetyzować rezystorową sieć o możliwie małej liczbie węzłów i elementów, w której relacje między napięciami a prądami wymuszonymi w węzłach będą odpowiadały relacjom między temperaturami a mocami czujników i grzejników.

Można zrealizować dwa rodzaje sieci: sieć związaną z opisem wielowrotnikowym i sieć związaną z opisem dwójnikowym.

2.3.1. Sieć wielowrotnikowa

Przy wyznaczaniu struktury wielowrotnikowej należy przyjąć, że wszystkie elementy termiczne czynne (w liczbie N) są jednocześnie czujnikami i grzejnikami. Równanie modelujące rozptyw ciepła jest następujące:

$$[T_{ij}]_{N \times N} = [\alpha_{ij}]_{N \times N} \cdot [P_j]_{N \times N}, \quad (12)$$

gdzie T_{ij} — przyrost temperatury w punkcie i w wyniku mocy wydzielanej w punkcie j

α_{ij} — elementy macierzy rezystancji termicznych

P_j — wartości mocy wydzielonych w punktach j

Przyjmując w równaniu (12) $[P_j]_{N \times N} = [1]$ otrzymujemy

$$[T_{ij}] = [\alpha_{ij}], \quad (13)$$

co oznacza, że współczynniki α_{ij} są równe przyrostom temperatury w punkcie i wynikłym z wydzielania się mocy jednostkowej w punkcie j .

W celu wyznaczenia współczynników α_{ij} należy przy pomocy dowolnej metody wyznaczania rozkładu temperatury (w przykładzie podanym na końcu artykułu posłużono się analitycznym rozwiązaniem równania przewodnictwa) wyznaczyć wartości przyrostów temperatur dla wszystkich możliwych kombinacji położenia źródeł i czujników dla wartości mocy w źródłach równej 1 W.

W ogólnym przypadku macierz $[\alpha_{ij}]$ nie jest symetryczna. Wynika to z modelowania struktury o stałych rozłożonych (źródła mają różne powierzchnie, zaś temperatury określa się w punkcie) przez strukturę dyskretną. W celu otrzymania rezystancyjnej sieci pasywnej i odwracalnej należy macierz $[\alpha_{ij}]$ zastąpić macierzą symetryczną. Najprostszym sposobem praktycznej symetryzacji jest takie ułożenie elementów czynnych (ponumerowanie źródeł), aby wyznaczyć transmisje ciepła od elementów, które w rzeczywistości są źródłami ciepła do rzeczywistych czujników i wyznaczenia jednej połówki macierzy $[\alpha_{ij}]$ (wraz z elementami α_{ij} dla $i = j$) i uzupełnienie całej macierzy przez symetryczne przeniesienie pierwszej połówki*). Dodatkową zaletą takiej symetryzacji jest znaczne zmniejszenie liczby potrzebnych do wyznaczenia współczynników α_{ij} .

W analogii elektrotermicznej równanie (12) można zapisać następująco

$$[U_{ij}] = [\beta_{ij}] \cdot [J_j]. \quad (14)$$

Dla $[J_j] = [1]$ można wykazać, że

$$[G_w] = [\beta_{ij}]^{-1}. \quad (15)$$

gdzie G_w jest macierzą przewodności węzłowych sieci elektrycznej.

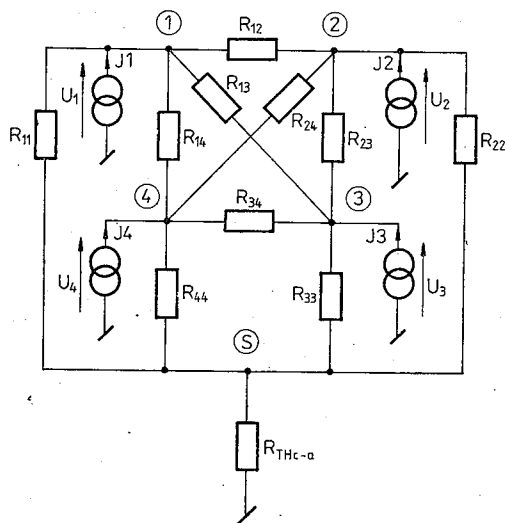
Wyznaczenie wartości poszczególnych rezystorów sieci elektrycznej z macierzy przewodności węzłowych jest proste:

$$R_{ij} = -\frac{1}{G_{ij}} \quad \text{dla } i \neq j$$

$$R_{ij} = \left(\sum_{k=1}^N G_{ik} \right)^{-1} \quad \text{dla } i = j \quad (16)$$

Przykład sieci wielowrotnikowej dla $N = 4$ przedstawiono na rys. 6. Na rysunku tym przedstawiono również sposób modelowania odpływu ciepła do otoczenia — czyli rezystancji termicznej obudowa — otoczenie RTH_{c-a} . Nie jest konieczne elektryczne modelowanie temperatury otoczenia, gdyż w programie SPICE-2G jest ona uwzględniana bezpośrednio. W modelu na rys. 6 wydajności źródeł prądowych są równoważne mocy wydzielanej w elementach, napięcia poszczególnych węzłów (względem masy) odpowiadają przyrostom temperatury względem temperatury otoczenia.

*) Błędy wynikające z symetryzacji są pomijalnie małe, gdyż praktycznie wszystkie elementy termicznie wrażliwe są w dużej odległości od źródeł ciepła, a wtedy wartość temperatury niewiele zależy od rozmiarów źródeł.



Rys. 6. Przykład sieci wielowrotnikowej

2.3.2. Sieć dwódnikowa

Przy wyznaczaniu struktury dwódnikowej można od razu podzielić elementy termicznie czynne na źródła i czujniki (ich liczby odpowiednio — J , I), przy czym ogólnie

$$N \leq I + J \leq 2N. \quad (17)$$

Transmisję ciepła można opisać następującymi równaniami:

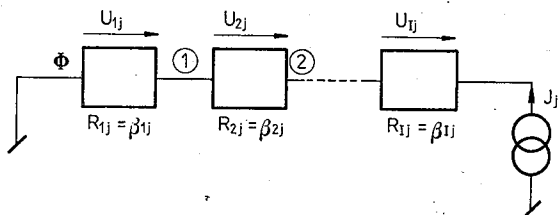
$$\frac{T_{1j}}{\alpha_{1j}} = \frac{T_{2j}}{\alpha_{2j}} = \dots = \frac{T_{Ij}}{\alpha_{Ij}} = P_j \quad j = 1 \div J \quad (18)$$

Liczba równań jest równa liczbie źródeł mocy — J .

W analogii elektrotermicznej zapis równań (18) jest następujący:

$$\frac{U_{1j}}{\beta_{1j}} = \frac{U_{2j}}{\beta_{2j}} = \dots = \frac{U_{Ij}}{\beta_{Ij}} = J_j \quad j = 1 \div J. \quad (19)$$

Obwód modelujący jedno z równań (19) przedstawiono na rys. 7. Przyrost temperatury i -tego elementu można wyznaczyć z zależności:

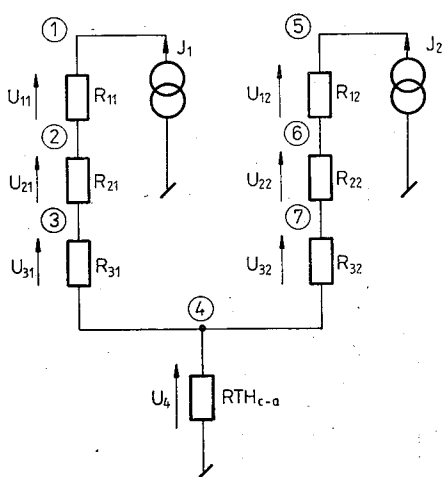


Rys. 7. Układ modelujący równanie rozptyłu ciepła

$$T_i = \sum_{j=1}^J T_{ij}, \quad (20)$$

W analogii elektrotermicznej:

$$U_i = \sum_{j=1}^J U_{ij}. \quad (21)$$



Rys. 8. Przykład sieci dwójnikowej

Na rys. 8 przedstawiono przykład modelu dwójnikowego dla dwóch grzejników i trzech czujników ($J = 2$, $I = 3$). Uwzględniono rezystancję termiczną obudowa-otoczenie.

2.3.3. Porównanie obu metod modelowania elektrotermicznego

Przy porównywaniu obu modeli sieci termicznej należy brać pod uwagę liczbę potrzebnych współczynników α_{ij} , łatwość posługiwania się modelem, liczbę wprowadzanych przez model węzłów i elementów oraz czas obliczeń całego układu.

Niewątpliwą wadą układu wielowrotnikowego jest fakt, że do jego syntezy potrzebna jest znajomość całej macierzy $[\alpha_{ij}]$. Nawet w przypadku symetryzacji macierzy przez wyznaczenie jej części czas obliczeń współczynników jest duży. Pominięcie nieistotnych kierunków transmisji ciepła (np. między tranzystorami czujnikami) jest możliwe po wyznaczeniu wszystkich rezystorów zastępczej sieci termicznej.

W przypadku opisu dwójnikowego można na samym początku procesu syntezy sieci pominąć mało istotne kierunki transmisji ciepła — stąd opis ten może być przydatny dla bardzo prostych układów, dla których wartości współczynników α_{ij} można uzyskać eksperymentalnie.

W opisie wielowrotnikowym napięcia danego węzła modelują całkowitą temperaturę w punkcie związanym z tym węzłem. Tej własności nie ma sieć dwójnikowa: aby uzyskać informację o całkowitej temperaturze należy zsumować napięcia na odpowiednich re-

zystorach sieci. W przypadku dużej liczby grzejników komplikuje to zapis i ewentualne eksperymenty komputerowe. Drugą dodatnią cechą sieci wielowrotnikowej jest fakt, że jej opis i opis struktury rezystancyjnej nie ulega zmianom, gdy w ramach N wybranych elementów termicznie czynnych nie wszystkie będą traktowane jako elementy grzejące lub czujniki.

Liczbę elementów i węzłów wprowadzanych przez model dla N elementów termicznie czynnych, w tym J grzejników i I czujników spełniających zależność (17) przedstawiono w tabeli poniżej.

Źródła prądowe sterowane i źródła napięciowe niezależne związane są z próbkowaniem mocy wydzielanej w grzejnikach. Źródła napięciowe sterowane związane są z modelowaniem własności termicznych tranzystorów — szczegóły podane są w rozdziale następnym.

Porównanie właściwości sieci wielowrotnikowej i dwójnikowej

ELEMENTY	sieć wielowrotnikowa		sieć dwójnikowa	
	$I, J \leq N$	$I = J = N$	$I, J \leq N$	$I = J = N$
rezystory	$\frac{N*(N+1)}{2} + J + 1$	$\frac{N*(N+3)}{2} + 1$	$J*(I+1) + 1$	$N*(N+1) + 1$
źródła prądowe sterowane	$2*J$	$2 \times N$	$2*J$	$2*N$
źródła napięciowe sterowane	I	N	I	N
źródła nap. niezależne	J	N	J	N
węzły	$N + 2*J + I + 1$	$4*N + 1$	$I*JI + 2*J + 1$	$N*(N+3) + 1$

Czas wyznaczenia charakterystyki przejściowej całego układu zależy przede wszystkim od liczby węzłów układu, będzie więc większy w przypadku stosowania struktury dwójnikowej.

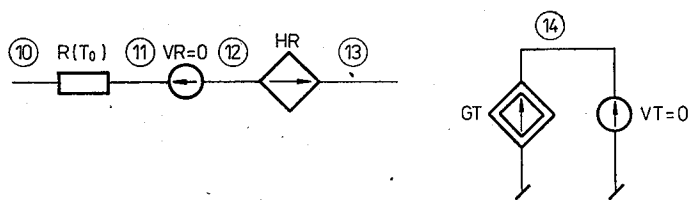
Z powyższych rozważań wynika, że oprócz przypadku, gdy tylko opis dwójnikowy jest możliwy (np. niemożność wyznaczenia wszystkich współczynników α_{ij}) sieć dwójnikową opłaca się stosować w przypadku małej liczby (1–3) źródeł mocy w stosunku do liczby czujników.

3. POŁĄCZENIE MODELU ELEKTROTERMICZNEGO Z ELEKTRYCZNYM

Właściwości termiczne i założenia związane z modelowaniem podstawowych elementów analogowych układów scalonych: rezystora i tranzystora bipolarnego zostały przedstawione w punkcie 2.1.

Model rezystora przedstawiono na rys. 9. Wynika on z zależności (3), którą można przedstawić inaczej:

$$R(T) = R(T_0) + \alpha R(T_0) \Delta T \quad (22)$$



Rys. 9. Model rezystora uwzględniający wpływ temperatury

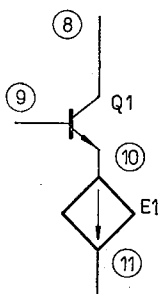
Drugi składnik zależności (22) modelowany jest przez źródło HR zależne od iloczynu prądu płynącego przez to źródło i prądu proporcjonalnego do przyrostu temperatury. Źródło GT przekształca napięcie między węzłami 1—0 (będące równoważne temperaturze) na prąd płynący przez źródło VT. Skomplikowany zapis wynika z konieczności próbkiowania tych prądów przy pomocy źródeł napięciowych (VR i VT). Przykładowy zapis w języku SPICE-2G dla $R(T_0) = 1 \text{ k}\Omega$, $\alpha = 0.0025$ (tzn. $\alpha R(T_0) = 2.5$) dla sieci wielowrotnikowej jest następujący:

```

GT  0  14  1  0  1
VT  14  0  .DC  0
RTO 10 11 1K
VR  11 12  .DC  0
HR  13 12  POLY(2) VR VT 0 0 0 0 2.5

```

Dla tranzystorów bipolarnych przyjęto, że temperaturowe zmiany U_T i I_s powodują przesuwanie się charakterystyki wejściowej tranzystora. Zmianę tę można modelować przez uzależnienie włączonych szeregowo w emiter tranzystorów źródeł napięciowych od napięcia modelującego przyrosty temperatury. Współczynnikiem zależności jest wartość γ (wzór 11). Przykład układu modelującego właściwości tranzystora npn przedstawiono na rys. 10. W przypadku wielowrotnikowego modelu sieci termicznej napięciowe



Rys. 10. Model tranzystora uwzględniający wpływ temperatury

źródło sterowane E1 uzależnione jest od jednego napięcia. Zapis w języku SPICE-2G dla sieci z rys. 7 jest następujący (przyjęto $\gamma = 2 \text{ mV/K}$)

```

E1 10 11 1 0 -2M

```

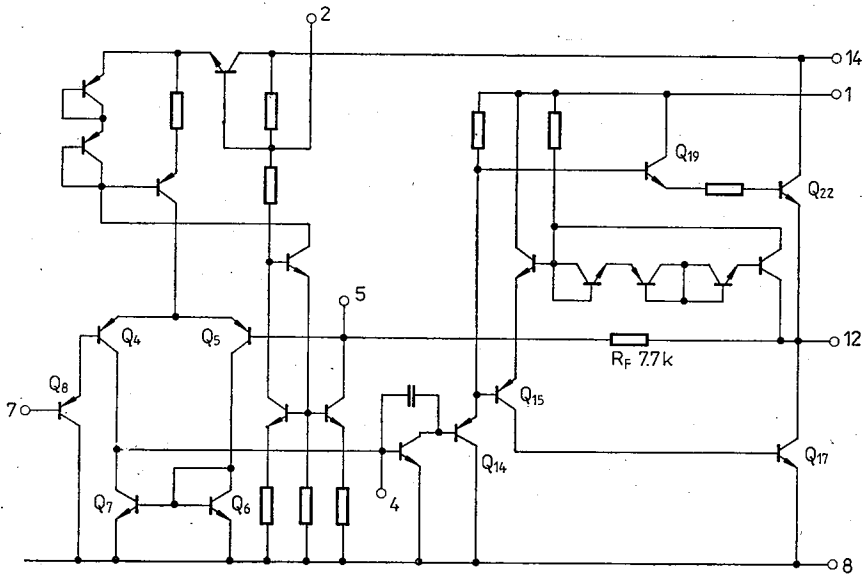
W przypadku dwójnikowego modelu sieci termicznej źródło E_1 sterowane jest $J+1$ napięciami (liczba źródeł mocy plus napięcie na rezystancji $R_{TH_{c-a}}$). Przykładowy zapis w języku SPICE-2G dla sieci z rys. 8 jest następujący

E1 10 11 POLY(3) 1 2 5 6 4 0 0 -2M -2M -2M

4. PRZYKŁAD MODELOWANIA

W celu zilustrowania przedstawionej metody modelowania wpływu termicznego sprzężenia zwrotnego w układach scalonych wyznaczono, w opisany powyżej sposób stałoprądową charakterystykę przejściową scalonego wzmacniacza mocy UL1498.

Schemat układu UL1498 przedstawiono na rys. 11. Ponieważ układ jest dość skompliko-



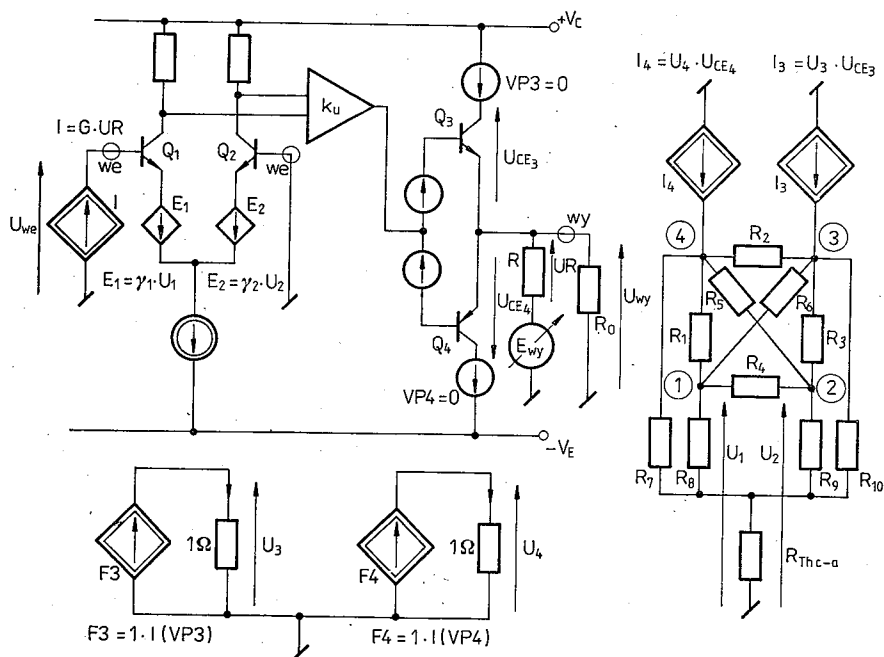
Rys. 11. Schemat układu wzmacniacza mocy UL1498

wany na rys. 12 przedstawiono ideę modelowania na przykładzie uproszczonego wzmacniacza operacyjnego. Składa się on z różnicowego stopnia wejściowego Q_1 Q_2 , stopnia wzmacniającego K_u niezależnego od temperatury i stopnia wyjściowego Q_3 i Q_4 pracującego w klasie B.

Elementy R , E_{wy} , oraz $J = G \times U_R$ realizują układ aktywnego sprzężenia zwrotnego umożliwiającego wyznaczenie charakterystyki $U_{we} = f(U_{wy})$. Elementy $R_1 - R_{10}$ oraz $R_{TH_{c-a}}$ tworzą wielowrotnikowy model termicznych właściwości obudowy układu. Węzły 1–4 związane są z tranzystorami $Q_1 - Q_4$. Napięcia U_1 i U_2 reprezentują temperaturę tranzystorów termicznie wrażliwych Q_1 i Q_2 . Sterują one napięciowe źródła sterowane E_1 i E_2 modelujące właściwości termiczne tranzystorów wejściowych. Źródła prądowe sterowane J_3 i J_4 reprezentują moc wydzieloną w tranzystorach Q_3 i Q_4 . Źródła

F_3 i F_4 związane są z próbkowaniem prądów kolektorowych Q_3 i Q_4 . Napięcia U_3 i U_4 są liczbowo równe tym prądom.

Układ UL1498 z racji swojego przeznaczenia — wzmacniacz sygnałów o częstotliwościach akustycznych — ma wewnętrzne sprzężenie zwrotne dla składowej stałej oraz zasilany jest napięciem niesymetrycznym.



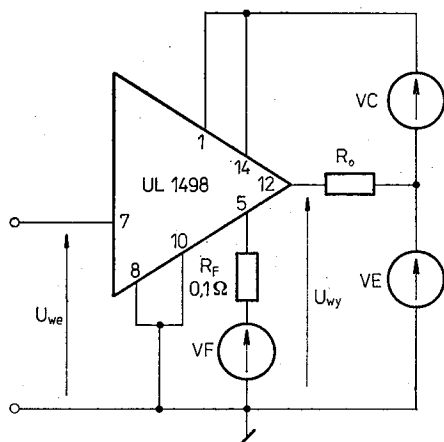
Rys. 12. Idea komputerowego modelowania sprzężenia termicznego

W celu wyznaczenia stałoprądowej charakterystyki przejściowej układu należy: zlikwidować wewnętrzne sprzężenie zwrotne dla składowej stałej, zlikwidować układ „bootstrap” oraz zasiląć układ napięciem symetrycznym. Likwidację sprzężenia zwrotnego zrealizowano przez wymuszenie na piątej końcówce układu stabilizowanego napięcia o wartości wyznaczonej uprzednio w typowym układzie aplikacyjnym. Oszacowana rezystancja tego źródła, wraz z uwzględnieniem rezystancji metalizacji wynosiła 0.1Ω i została uwzględniona przy obliczeniach komputerowych.

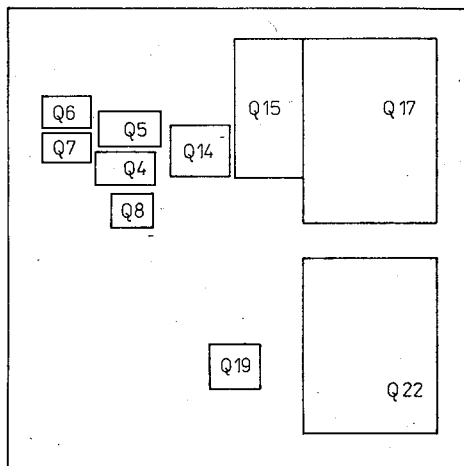
Uproszczony układ pomiarowy (pominięto pojemności kompensacji częstotliwościowej) dostosowany do wyznaczenia stałoprądowej charakterystyki przejściowej przedstawiono na rys. 13.

Wpływ rezystorów na charakterystykę przejściową okazał się dużo mniejszy niż wpływ tranzystorów. Jako elementy termicznie czynne wybrano więc tylko tranzystory: pięć tranzystorów podgrzewających — Q_{22} , Q_{17} , Q_{15} , Q_{19} i Q_{14} oraz pięć tranzystorów czujników — Q_4 , Q_5 , Q_6 , Q_7 i Q_8 . Współczynniki γ dla tranzystorów czujników wyznaczono ze wzoru (11). Położenie tranzystorów termicznie czynnych na płytce krzemowej przedstawiono na rys. 14.

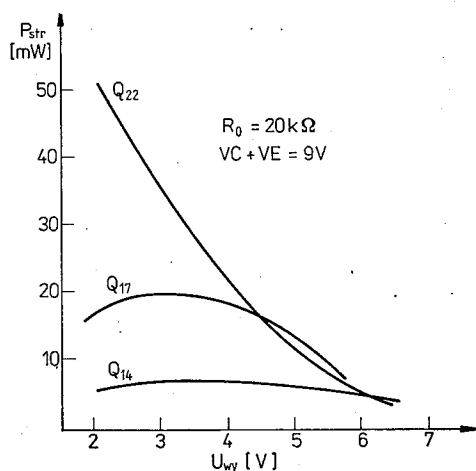
Zależności mocy strat w tranzystorach grzejnikach od napięcia przedstawiono na rysunkach 15 i 16. Przyjęto dwie wartości rezystancji obciążenia $R_o = 20 \text{ k}\Omega$ (praktyczny brak obciążenia) i $R_o = 11 \Omega$.



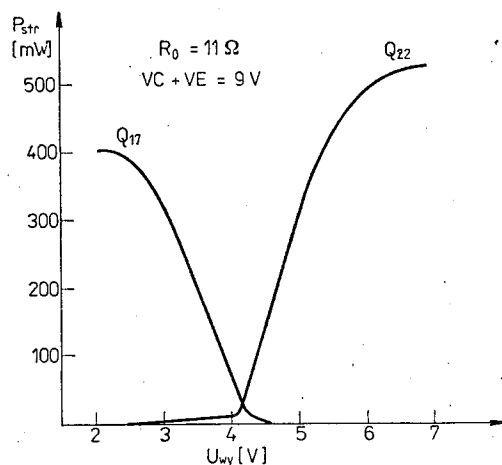
Rys. 13. Schemat układu do pomiaru charakterystyki przejściowej układu UL1498



Rys. 14. Rozmieszczenie tranzystorów termicznie czynnych na płytce krzemowej układu UL1498



Rys. 15. Zależność mocy traconej od napięcia wyjściowego. $R_o = 20 \text{ k}\Omega$



Rys. 16. Zależność mocy traconej od napięcia wyjściowego. $R_o = 11 \Omega$

Rozkład temperatury wyznaczono na podstawie analitycznego rozwiązania układu równań przewodnictwa cieplnego. W tabeli poniżej przedstawiono wybrany fragment macierzy współczynników α_{ij} — przed symetryzacją. Uwzględniono najważniejsze czynniki i grzejniki. Cała macierz zawiera 10×10 elementów.

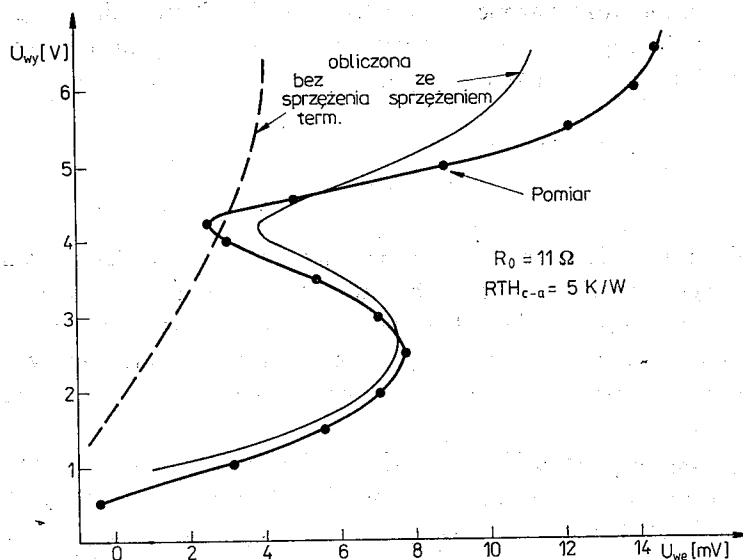
Przykład macierzy współczynników α_{ij}

α_{ij}	Q_{22}	Q_{17}	Q_5	Q_4	Q_8
Q_{22}	14.52	2.70	1.55	1.45	1.94
Q_{17}	2.70	14.38	2.06	2.27	1.96
Q_5	1.50	2.01	45.92	13.93	13.69
Q_4	1.41	2.31	13.82	45.10	7.04
Q_8	1.69	1.91	14.40	7.18	55.80

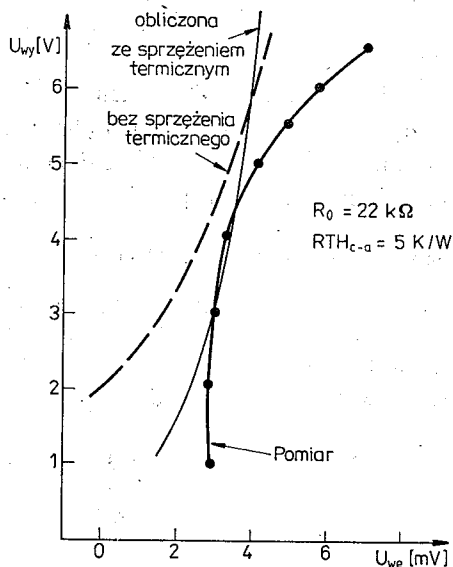
Układ elektryczny wzmacniacza zawiera 42 elementy i 32 węzły. Dane dotyczące liczby elementów, węzłów, zajętości pamięci i czasu obliczania 59 punktów charakterystyki przejściowej przy pomocy programu PC-USPICE Unit Software & Consulting, Inc. Tempe, Arizona na komputerze IBM-XT przedstawiono w tabeli poniżej.

Porównanie czasu obliczeń i zajętości pamięci

	UKŁAD ELEKTRYCZNY		
	bez modelowania sprężenia termicznego	z siecią termiczną	
		wielowrotnikową	dwójnikową
elementy	42	123	93
węzły	32	58	73
pamięć	10 272 B	16 368 B	17 164 B
czas	171 s	480 s	567 s

Rys. 17. Wyniki pomiarów i symulacji komputerowej. $R_0 = 11 \Omega$

Wyznaczono charakterystyki przejściowe ze sprzężeniem termicznym dla rezystancji termicznej radiatora $R_{TH_{c-a}} = 5 \text{ K/W}$ i dla dwóch rezystancji obciążenia $R_0 = 11 \Omega$ oraz $R_0 = 20 \text{ k}\Omega$. Wyniki pomiarów i symulacji komputerowej przedstawiono na rysunkach 17, 18. Porównując charakterystyki zmierzone z obliczonymi widać niedokładności



Rys. 18. Wyniki pomiarów i symulacji komputerowej. $R_0 = 20 \text{ k}\Omega$

w modelu elektrycznym (malenie wzmocnienia tranzystora wyjściowego Q_{22} przy dużym sygnale dodatnim). Bardzo znamienny jest fakt, że dla rezystancji obciążenia $20 \text{ k}\Omega$, a więc dla przypadku, gdy moc wydzielona w układzie jest najmniejsza, można również obserwować wpływ sprzężenia termicznego.

PODSUMOWANIE

Przedstawiona metoda umożliwia wyznaczenie właściwości statycznych dowolnych układów scalonych z uwzględnieniem wpływów termicznych. Zaletą tej metody jest możliwość niezależnego wyznaczenia rozkładu temperatury — dowolną metodą dostosowaną do typu obudowy układu scalonego. Dzięki temu wielkość układu elektrycznego modelującego właściwości termiczne obudowy jest uzależniona tylko od elektrycznych właściwości badanego układu.

Metoda syntezy sieci modelującej właściwości termiczne obudowy jest oczywiście możliwa do zastosowania z każdym uniwersalnym programem analizującym układy elektryczne (np. NAP2). W artykule przedstawiono zastosowanie jej wraz z programem SPICE-2G, gdyż przy pomocy tego programu łatwo jest symulować działanie tak dużego układu jak układ scalony UL1498.

Z przeprowadzonych pomiarów, potwierdzonych symulacją komputerową wynika, że w dużych układach wzmacniaczy prądu stałego należy liczyć się z występowaniem termicznego sprzężenia zwrotnego. Wynika stąd oczywisty wniosek, że poprawne modelowanie musi to sprzężenie uwzględniać. Przedstawiony w artykule tok postępowania należy traktować jako pierwszy krok w dokładniejszym modelowaniu analogowych właściwości układów scalonych.

BIBLIOGRAFIA

1. A. Auer: *Modelowanie analogowe procesów o stałych rozłożonych*. PWN, Warszawa 1976
2. R. Castello, P. Antognetti: *Integrated-circuit thermal modeling*. IEEE, Journal of Solid-State Circuits vol. SC-13, no 3, June 1978
3. K. Fukahori: *Computer simulation of monolithic circuit performance in the presence of electro-thermal interaction*. Memor. no UCB/ERC 77/13, University of California, Berkeley 1977
4. L. W. Nagel: *SPICE 2: A computer program to simulate semiconductor circuits*. Memor. no ERL-M520, 1975
5. K. Zamyński: *Modelowanie i analiza elektrotermiczna monolitycznych układów scalonych*. Rozprawa doktorska. Politechnika Warszawska, Warszawa 1982

Praca została wykonana w ramach CPBP 02.14.

K. ZAMLYŃSKI

MODELLING OF THERMAL FEEDBACK IN MONOLITHIC INTEGRATED SYSTEMS BY MEANS OF THE PROGRAM SPICE-2G

Summary

A modelling method of the thermal feedback phenomenon by means of the general purpose program SPICE-2G, is represented. The method is illustrated by an example of the determination of the transitory characteristics of a monolithic power amplifier type UL 1498.

K. ZAMLYŃSKI

MODELAGE DE LA RÉTROACTION THERMIQUE DANS LES CIRCUITS INTÉGRÉS MONOLITHIQUES À L'AIDE DU PROGRAMME SPICE-2G

Résumé

Dans l'article on a présenté la méthode du modelage du phénomène thermique de la rétroaction à l'aide du programme universel SPICE-2G. Cette méthode a été illustrée sur l'exemple de la détermination de la tangente de la caractéristique mutuelle de l'amplificateur monolithique de puissance UL.

K. ZAMLYŃSKI

THERMISCHE MODELLIERUNG DER RÜCKKOPPLUNG MITTELS PROGRAMM SPICE-2G BEI INTEGRIERTEN MONOLITHSYSTEMEN

Zusammenfassung

Im Artikel wurde eine Modelliermethode für das thermische Phänomen der Rückkopplung mit Hilfe eines Universalprogramms SPICE-2G geschildert. Diese Methode wurde am Beispiel der Bestimmung der statischen Teilcharakteristik eines Monolithleistungsverstärkers erläutert.

K. ЗАМЛЫНСКИ

МОДЕЛИРОВАНИЕ ТЕМПЕРАТУРНОЙ ОБРАТНОЙ СВЯЗИ В МОНОЛИТНЫХ
ИНТЕГРАЛЬНЫХ СХЕМАХ ПРИ ПОМОЩИ ПРОГРАММЫ SPICE-2G

Резюме

Представлен метод моделирования явления температурной обратной связи при помощи универсальной программы SPICE-2G. Метод иллюстрирован примером расчета статической переходной характеристики монолитного усилителя мощности UL 1498.

Właściwości przetworników wiropądowych do nieniszczących badań prętów ferromagnetycznych

ANNA LEWIŃSKA-ROMICKA, STANISŁAW PISKORSKI

*Centrum Uczelniano-Przemysłowe Metrologii i Systemów Pomiarowych
Politechnika Warszawska*

Otrzymano 1988.03.25

Autoryzowano do druku 1988.11.25

Opisano metodykę i wyniki badań przetworników wiropądowych, opracowanych do wykrywania pęknięć prętów wykonanych z materiałów ferromagnetycznych. Przedstawiono sposób określania częstotliwości pracy przetworników.

1. WSTĘP

Wzrastające wymagania dotyczące jakości półfabrykatów oraz wyrobów gotowych powodują, że nieniszczące badania materiałów i produktów stają się coraz częściej składową procesy wytwarzania. Oznacza to, że produkty już w czasie wytwarzania są badane i mogą być zaliczane do różnych klas jakości na podstawie wyników kontroli.

Badania nieniszczące, wykonane w możliwie najwcześniejszym stadium wytwarzania, pozwalają na zaoszczędzenie kosztów materiałowych oraz zbędnych kosztów dodatkowych, jakie mogą powstawać przy dalszej obróbce elementów, które już są wadliwe.

Metoda badania jest właściwa, o ile daje się scalać z istniejącym procesem wytwarzania, jak również zautomatyzować. Proces produkcji nie może podlegać zakłóceniom w wyniku badań. Oznacza to również między innymi, że metoda badania powinna być bezkontaktowa, a prędkość kontroli — dostatecznie duża. Powyższe wymagania spełnia, między innymi, metoda oparta na wykorzystaniu zjawiska prądów wirowych, która jest przydatna do kontroli obiektów metalowych, tj. przewodzących prąd elektryczny.

Zjawisko prądów wirowych można w najogólniejszy sposób określić w sposób następujący. Jeżeli w zmiennym polu elektromagnetycznym znajduje się obiekt z materiału przewodzącego prąd elektryczny, to na skutek określonego przez prawo Faraday'a zjawiska indukcji elektromagnetycznej powstają w tym obiekcie tzw. prądy wirowe.

Rozkład gęstości prądów wirowych wytwarzanych w obiekcie zależy od następujących czynników:

- a) częstotliwości i rozkładu pola elektromagnetycznego,
- b) kształtu i wymiarów obiektu oraz jego usytuowania względem pola,
- c) konduktywności, przenikalności magnetycznej, a więc i temperatury oraz stopnia

niejednorodności (w tym również anizotropowości) materiału, z którego obiekt został wykonany.

Prądy wirowe mogą być więc źródłem informacji o obiekcie, w którym zostały wzbudzone. Odbiór tej informacji jest możliwy za pośrednictwem wytwarzanego przez prądy wirowe ich własnego, w odróżnieniu od zewnętrznego pola elektromagnetycznego, tzw. pola prądów wirowych. Wzbudzanie prądów wirowych w badanych obiektach, jak też odbiór informacji przenoszonej przez pole magnetyczne tych prądów odbywa się za pośrednictwem przetworników przedstawiających sobą różnego rodzaju cewki indukcyjne.

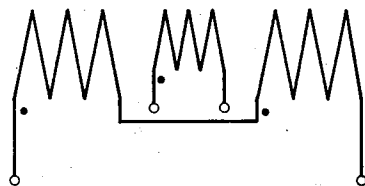
Niniejszy artykuł dotyczy wykrywania nieciągłości w obiektach o kształcie walca, przy zastosowaniu metody prądów wirowych, na tle występujących zakłóceń pomiaru.

Zagadnienie wykrywania nieciągłości przedstawiono na przykładzie kontroli prętów wykonanych z materiałów ferromagnetycznych. Przy badaniu pęknięć obiektów ze względu na to, że rozkład gęstości prądów wirowych zależy m.in. od właściwości fizykochemicznych obiektów istnieją następujące główne czynniki zakłócające pomiar: zmiany składu chemicznego, stanu obiektu wynikającego m.in. z obróbki cieplnej (twardość, struktura wewnętrzna) oraz zmiany rozmiarów.

Przy wykrywaniu pęknięć obiektów metodą prądów wirowych, w celu uzyskania odpowiedniego dużego stosunku sygnałów pochodzących od pęknięć do zakłóceń, wywołanych powyższymi czynnikami, konieczny jest dobór odpowiednich: przetworników pomiarowych, układów pracy przetworników, częstotliwości pracy przetworników, układów elektronicznych opracowywania sygnałów uzyskiwanych z przetworników, monitorów wad. Rezultaty badań nad zagadnieniami defektoskopii wiropądowej, prowadzonych w Centrum Uczelniano-Przemysłowym Metrologii i Systemów Pomiarowych Politechniki Warszawskiej, pozwoliły na skonstruowanie szeregu defektoskopów [2, 19÷24].

2. PRZETWORNIK WIROPĄDOWY

Do wykrywania wad drutów i prętów, wykonanych z materiałów ferromagnetycznych, opracowano przetwornik generacyjny (rys. 1) o nierównomiernym rozkładzie pola magne-



Rys. 1. Przetwornik wiropądowy generacyjny

tycznego wzdłuż osi obiektu. Przetwornik pracuje w układzie porównawczym ze źródłem napięcia o regulowanej amplitudzie i fazie. Konfiguracja oraz układ pracy przetwornika zostały dobrane tak, aby uzyskać dużą czułość wykrywania pęknięć drutów i prętów (p. 4).

Dwie symetryczne sekcje uzwojenia wejściowego przetwornika (rys. 1) połączone są szeregowo, a strumienie magnetyczne wytworzone przez przepływający przez nie prąd dodają się. Pomiedzy sekcjami uzwojenia wejściowego umieszczono uzwojenie wyjściowe [19].

Przy analizie właściwości przetworników przyjęto określać składowe napięcia tych przetworników. Zależności opisujące unormowane składowe napięcia przetwornika wiroprowodowego zawierającego próbkę ferromagnetyczną można przedstawić w następującej postaci [9]:

$$\begin{aligned}\frac{Re(U_r)}{U_0} &= \eta \mu_r Im(\mu_s) \\ \frac{Im(U_r)}{U_0} &\cong \eta \mu_r Re(\mu_s)\end{aligned}\quad (1)$$

przy czym U_r i U_0 przedstawiają odpowiednio napięcie indukowane w uzwojeniu wyjściowym przetwornika zawierającego obiekt i przetwornika pustego. η jest to tzw. współczynnik wypełnienia przetwornika przez obiekt badany; $\eta = (D_0/D_p)^2$, gdzie D_0 — średnica obiektu, D_p — średnia średnica przetwornika, μ_s — jest to tzw. przenikalność skuteczna [9]; jest ona stabilizowana.

Na rys. 2 i 3 przedstawiono unormowane składowe napięcia przetwornika wiroprowodowego, zawierającego ferromagnetyczną próbkę walcową. Częstotliwość pracy przetwornika oznaczono jako f , a f_g jest tzw. częstotliwością graniczną, opisaną przez zależność [9, 26]

$$f_g = \frac{5066}{\gamma \mu_r D_p^2} \text{ [Hz]}, \quad (2)$$

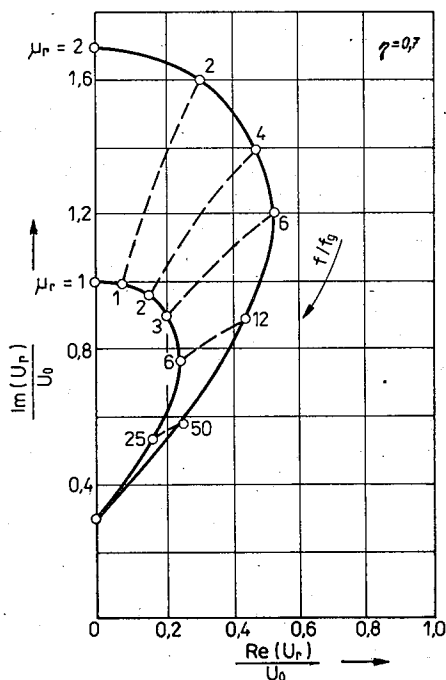
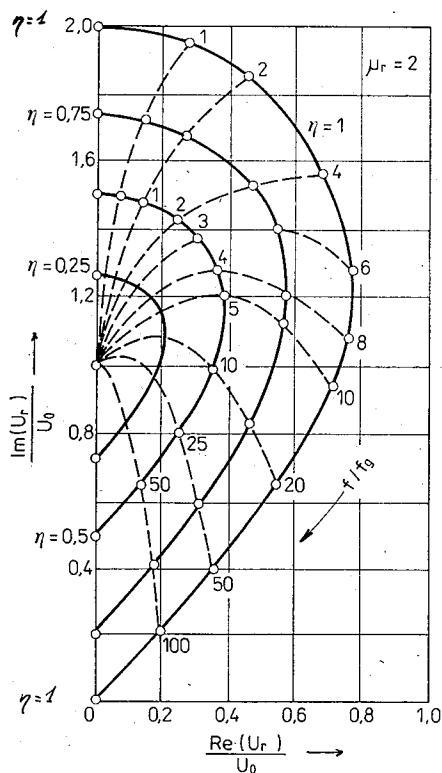
przy czym γ — przewodność właściwa obiektu, w S/m, μ_r — przenikalność magnetyczna względna obiektu, D_p — średnica obiektu, w cm.

Przy zastosowaniu metody prądów wirowych do wykrywania nieciągłości, ważne konsekwencje pomiarowo-techniczne ma możliwość wzajemnie prostopadłego ustawienia w płaszczyźnie zespolonych składowych napięć przetwornika pomiarowego składowej napięcia pochodzącej od zmian średnicy obiektu i składowej pochodzącej od zmian właściwości fizycznych (przewodność elektryczna γ i przenikalność magnetyczna μ_r), gdyż dzięki temu możliwe jest rozdzielenie poszczególnych wpływów.

Korzystając z rys. 2 [5] zbadamy wpływ zmian średnicy ferromagnetycznej próbki na unormowanie urojone i rzeczywiste wartości napięć przetwornika pomiarowego. Aby umożliwić przejrzyste przedstawienie wpływu zmian rozmiarów próbki ferromagnetycznej przyjęto małą wartość przenikalności $\mu_r = 2$. Proste rozważania pozwalają stwierdzić ważną okoliczność: gdy średnica próbki zwiększa się tak, że współczynnik wypełnienia wzrasta od $\eta = 0,75$ do $\eta = 1$, to wartość początkowa unormowanego napięcia wyjściowego przetwornika rośnie przy $f/f_g = 0$ od $Im(U_r)/U_0 = 1,75$ do $Im(U_r)/U_0 = 2$.

Natomiast przy tym samym wzroście średnicy próbki, tj. przy współczynniku wypełnienia zawartym w zakresie od $\eta = 0,75$ do $\eta = 1$, wartość końcowa unormowanego

napięcia wyjściowego przetwornika zmniejsza się przy $f/f_g = \infty$ od $Im(U_r)/U_0 = 0,25$ do 0, czyli ta sama zmiana średnicy przy $f/f_g = 0$ i przy $f/f_g = \infty$ przebiega na płaszczyźnie napięć zespolonych w kierunku przeciwnym. Oznacza to, że faza napięcia wyjściowego przetworników w przypadku zmian średnicy próbki zmienia się o 180° przy wzroście wartości f/f_g od 0 do ∞ . Stąd wynika jednak, że kierunek, w którym uwidaczniają się sygnały wywołane przez zmiany średnicy próbki raz pokrywa się z osią poziomą $Re(U_r)/U_0$



Rys. 2. Wpływ zmiany średnicy walcowego obiektu ferromagnetycznego o stałej przenikalności na zespolone napięcie wyjściowe przetwornika wiroprądowego [5]

Rys. 3. Wpływ zmiany przenikalności na zespolone napięcie przetwornika wiroprądowego zawierającego walcowy obiekt o stałej średnicy [5]

płaszczyzny napięć zespolonych, a drugim razem jest prostopadły do kierunku, w którym uwidaczniają się sygnały wywołane przez zmiany przenikalności, przewodności elektrycznej obiektu i częstotliwości pola zmiennego zasilającego przetwornik.

W przedstawione hodografy wrysowano liniami przerywanymi krzywe, obrazujące wpływ zmian średnicy obiektu na sygnał wyjściowy tych przetworników. Zakrzywienie wykresów wynika stąd, że wartości składowych unormowanych napięcia maleją liniowo ze zmniejszaniem się kwadratu średnicy obiektu, podobnie, jak wartości współczynnika wypełnienia. I tak np. zespolona wartość napięcia z $f/f_g = 8$ dla krzywej dla $\eta = 1$ przesuwa się dla krzywej dla $\eta = 0,75$ do $f/f_g = 6$, dla $\eta = 0,5$ przesuwa się do $f/f_g = 4$ i wreszcie dla $\eta = 0,25$ — do $f/f_g = 2$.

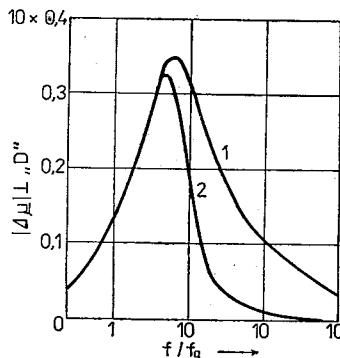
Jak widać, przy wartościach f/f_g bliskich zeru kierunek, w którym uwidaczniają się sygnały wywołane przez zmiany średnicy, jest prostopadły do kierunku, w którym uwidaczniają się sygnały wywołane przez zmiany przewodności elektrycznej i częstotliwości prądu wzbudzającego, natomiast sygnały wywołane przez zmiany średnicy w tym zakresie f/f_g przebiegają zgodnie z kierunkiem sygnałów pochodzących od zmian przenikalności.

Na rys. 3 przedstawiono unormowane napięcie przetwornika dla przypadku, gdy przy współczynniku $\eta = 0,7$ zredukowano przenikalność próbki z wartości $\mu_r = 2$ do $\mu_r = 1$ np. przez przyłożenie stałego magnetycznego pola nasycenia [5]. Rysunek 3 przedstawia zupełnie inny przebieg składowych napięcia przetwornika przy zmianach przenikalności magnetycznej w porównaniu do przebiegu tych składowych przy zmianach średnicy próbki (rys. 2) i to już poczynając od niskich wartości f/f_g . Już przy wartości $f/f_g = 6$ „kierunek średnic” na rys. 2 jest prawie prostopadły do „kierunku przenikalności” na rys. 3 podczas, gdy przy wysokich wartościach f/f_g kąt między „kierunkiem średnic” i „kierunkiem przenikalności” dąży do maksymalnej wartości 135°.

Z rysunku 3 możemy również wnioskować:

— zdolność (czułość) wykrywania zmiany μ_r zmniejsza się ze wzrostem wartości f/f_g ; np. zmiana μ_r od 2 do 1 na odcinku od $f/f_g = 4$ do $f/f_g = 2$ daje 15-krotnie większą zmianę wartości napięć zespolonych przetwornika pomiarowego niż na odcinku od $f/f_g = 400$ do $f/f_g = 200$,

— ze wzrostem wartości f/f_g „kierunek przenikalności” na płaszczyźnie napięć zespolonych zbliża się sukcesywnie do „kierunku przewodności elektrycznej i częstotliwości” — ale ze znakiem przeciwnym. Wzrost przenikalności uwidacznia się na płaszczyźnie napięć zespolonych przetwornika pomiarowego przy wysokich wartościach f/f_g w kierunku przeciwnym do kierunku, jaki daje wzrost przewodności elektrycznej i częstotliwości prądu w przetworniku.



Rys. 4. Przyrost przenikalności skutecznej próbki walcowej z materiału ferromagnetycznego, wywołany 1-procentową zmianą przewodności właściwej [5]

Na rys. 4 przedstawiono moduł przyrostu $|\Delta\mu_s|$ przenikalności skutecznej obiektu walcowego z materiału ferromagnetycznego, wywołany 1-procentową zmianą przewodności właściwej oraz składową przyrostu w kierunku prostopadłym do „kierunku zmian średnicy D_p ” [5].

Wartość składowej przyrostu $|\Delta\mu_s|$, prostopadłej do kierunku „zmian średnicy D_p ”, jest miarą rozróżnialności zmian przewodności na tle zmian średnicy obiektu D_p . Maksimum tej składowej występuje dla $f/f_g = 5$. Warunek ten w przypadku materiałów o dużej przenikalności magnetycznej jest w praktyce trudny do spełnienia, gdyż wymaga użycia prądu wzbudzającego o małej częstotliwości.

3. CZĘSTOTLIWOŚĆ PRACY PRZETWORNIKA

Wyżej opisano czynniki wpływające na napięcie wyjściowe przetworników wiropądowych przy wykrywaniu nieciągłości obiektów ferromagnetycznych. Ze względu na konieczność uzyskania dużej czułości urządzeń (a przede wszystkim przetworników) na pęknięcia badanych prętów, zakres ich średnic (0,5 mm — 3,0 mm) podzielono na dwa podzakresy i dla każdego z podzakresów dobrano odpowiednie częstotliwości pracy urządzeń. Występujące w obiektach, wykonanych z materiałów ferromagnetycznych, wahania rezystancji strat i przenikalności magnetycznej [17], spowodowane nierównomiernością struktury, powodują indukowanie się w uzwojeniu wyjściowym przetwornika sygnałów stanowiących zakłócenie przy wykrywaniu pęknięć. Również zmiany średnicy prętów powodują zmiany napięcia wyjściowego przetwornika.

Ze względu na to, że pęknięcia prętów i drutów powodują także zmiany rezystancji strat i przenikalności magnetycznej, konieczny jest dobór odpowiedniej częstotliwości pracy i takiego układu analizy sygnałów, aby otrzymać odpowiedni stosunek sygnałów pochodzących od pęknięć, do zakłóceń wywołanych nierównomiernością struktury.

Występujący w zależnościach (1) iloczyn $\eta\mu_r$ jest równy $\eta\mu_r = \left(\frac{D_0}{D_p}\right)^2 \mu_r$, przy czym:

D_0 — średnica obiektu, D_p — średnica cewki przetwornika. Wynika stąd, że zmiany napięcia wyjściowego przetwornika wywołane przez zmiany średnicy (η) i przenikalności magnetycznej μ_r pod względem fazowym zachodzą w tym samym kierunku i nie można odróżnić, która ze zmian (średnicy czy przenikalności) spowodowała zmianę napięcia przetwornika pomiarowego. Fakt ten korzystny jest przy wykrywaniu pęknięć w stali, gdyż istnienie wspólnego „kierunku $\mu_r D_p$ ” zmniejsza liczbę „kierunków” niepożądanych.

Powyższe rozumowanie jest słuszne tylko dla małych stosunków częstotliwości pracy do częstotliwości granicznej $1 \leq \frac{f}{f_g} \lesssim 5$ gdyż, jak to wynika z rys. 2 i 3, jedynie wówczas

sygnały wywołane przez zmiany średnicy obiektu na płaszczyźnie zmiennej zespolonej ujawniają się w kierunku zgodnym z kierunkiem, w którym uzewnętrzniają się zmiany przenikalności. Jednocześnie w tym zakresie stosunku f/f_g kąt fazowy pomiędzy sygnałami

pochodzącymi od pęknięć (kierunek $\frac{f}{f_g}$), a sygnałami pochodzącymi od zmian średnicy

i wahań przenikalności magnetycznej (kierunek $\mu_r D_p$) jest największy i bliski 90° . Przy uwzględnieniu powyższego, biorąc jednocześnie pod uwagę możliwość wykrywania wad podpowierzchniowych, przyjęto częstotliwość pracy defektoskopu, przeznaczonego do kontroli prętów i drutów stalowych o średnicach z zakresu 1,5 mm — 3,0 mm równą $f = 10$ kHz.

Częstotliwość graniczna, obliczona dla prętów ze stali chromowej średnicy $D_p = 2,3$ mm, przewodności właściwej $\gamma = 1,7 \cdot 10^6$ S/m i przenikalności magnetycznej względnej $\mu_r = 5$ wynosi $f_g = 11\,300$ Hz.

Przy kontroli drutów stalowych średnic w zakresie 0,5 mm—1,5 mm przyjęto natomiast częstotliwość pracy 25 kHz. Częstotliwości pracy defektoskopów zostały zweryfikowane na podstawie badań drutów i prętów z wadami sztucznymi w postaci rowków wzdłużnych.

Głębokość wnikania pola magnetycznego, obliczona według zależności (3) [9, 26]

$$\delta = \frac{503}{\sqrt{f\gamma\mu_r}} \text{ [mm]} \quad (3)$$

przy częstotliwości pracy $f = 10$ kHz wynosi $\delta = 1,72$ mm, a przy częstotliwości pracy $f = 25$ kHz $\delta = 1,09$ mm.

4. WŁAŚCIWOŚCI METROLOGICZNE I EKSPLOATACYJNE PRZETWORNIKA

Jako najistotniejsze zagadnienie, determinujące właściwości metrologiczne i eksploatacyjne przetwornika wiroprowadowego, można uznać czułość wykrywania wad i zmian średnicy. Zagadnienia te zostaną opisane przy szczególnym uwzględnieniu doboru częstotliwości pomiarowej, odpowiedniej do wykrywania pęknięć drutów i prętów. Częstotliwość pomiarowa powinna być dobrana tak, aby zapewnić możliwość eliminacji wpływu zmian średnicy i przenikalności obiektów badanych na wykrywalność pęknięć.

Właściwości metrologiczne przetwornika mogą być określone na podstawie:

— zmian wartości bezwzględnej napięcia wyjściowego przetwornika ΔX_R , wywołanej istnieniem wady,

— zmian wartości składowej napięcia wyjściowego przetwornika ΔX_{R1} prostopadłej do wartości bezwzględnej zmian napięcia wyjściowego przetwornika ΔX_s , wywołanej zmianą średnicy oraz

— przebiegu kąta fazowego Φ pomiędzy wielkością ΔX_R , spowodowaną istnieniem wady i wielkością ΔX_s , wywołaną zmianą średnicy, przy czym

$$\Delta X_{R1} = \Delta X_R \sin \Phi \quad (4)$$

Analiza powyższych wielkości pozwala określić warunki pracy przetwornika wiroprowadowego tak, aby uzyskać maksymalną czułość wykrywania wad, niezależnie od zmian średnicy i wahań przenikalności magnetycznej obiektu.

Analiza ta może być dokonana na podstawie pomiarów składowych napięcia wyjściowego przetwornika, U_R zgodnej w fazie z prądem zasilającym przetwornik i składowej U_I , przesuniętej o kąt $\pi/2$ względem prądu zasilającego. Składowe napięcia przetwornika nie zawierającego obiektu badanego oznaczono odpowiednio jako U_{R0} i U_{I0} . Zależności

$$ReX = \frac{U_R - U_{R0}}{U_{I0}} \quad ImX = \frac{U_I}{U_{I0}} \quad (5)$$

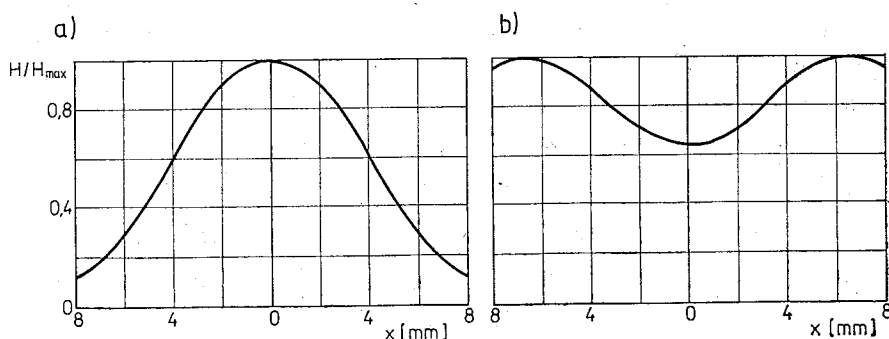
zostaną obliczone analogicznie, jak przy analizie właściwości przetworników wirowych o jednorodnym polu elektromagnetycznym.

Podane na rys. 6—28 charakterystyki przetwornika dotyczą przypadku, gdy znajdował się w nim pręt z wadą sztuczną w postaci rowka wzdłużnego o znacznej głębokości (ok. 36,5% średnicy) i zmianą średnicy około 21,7%. Konieczność wprowadzenia tak znacznych wad spowodowana jest małą czułością zastosowanej tu bezwzględnej metody pomiaru z użyciem miernika składowych napięć przetwornika. Metoda ta umożliwia wykonanie badań w szerokim zakresie częstotliwości (w tym przypadku od 1 kHz do 50 kHz). Metoda porównawcza może być zastosowana do badań prętów, zawierających wady o różnych głębokościach, przy wybranej wartości prądu wzbudzającego pole elektromagnetyczne.

Charakterystyki opracowanych przetworników przytoczone na rys. 6—28 zależą od następujących czynników: częstotliwości prądu I_w wzbudzające pole magnetyczne, wartości prądu I_w wzbudzającego pole elektromagnetyczne, stałego pola magnetycznego H_s podmagnesowującego pręt badany, obecności wady sztucznej w pręcie badanym i zmiany średnicy pręta badanego.

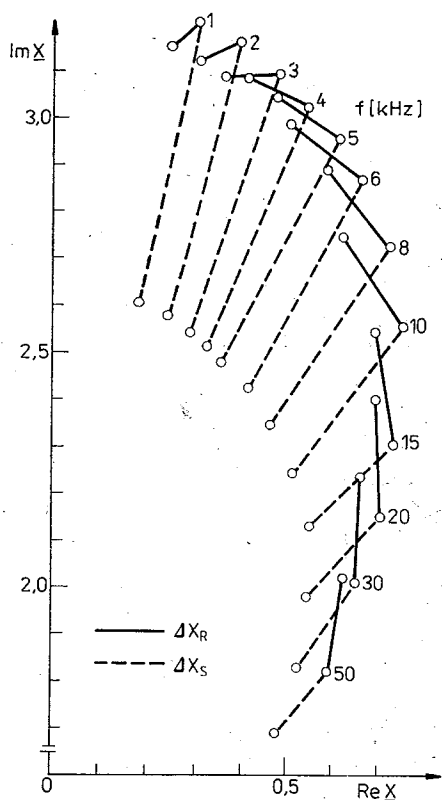
PRZETWORNIK PARAMETRYCZNY

Jako przetwornik parametryczny można wykorzystać uzwojenie wyjściowe przetwornika generacyjnego, pokazanego na rys. 1. Wartości współczynników wypełnienia przetwornika parametrycznego i generacyjnego przez obiekt badany są takie same, co umożliwia porównanie ich właściwości. Różnica pomiędzy analizowanymi przetwornikami polega na różnych rozkładach składowych osiowych pola magnetycznego w obszarze uzwojenia wyjściowego (rys. 5a i rys. 5b). Zależność opisującą składową osiową pola magnetycznego podano w pracy [11].

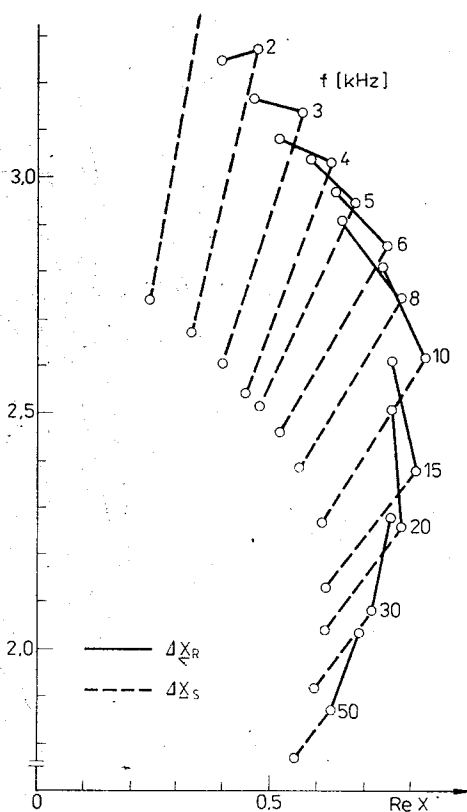


Rys. 5. Rozkład pola magnetycznego wzdłuż osi przetwornika a — parametrycznego (tylko uzwojenie wyjściowe przetwornika z rys. 1), b — generacyjnego, jak na rys. 1

Obliczone według zależności (5) składowe unormowane napięcia wyjściowego przetwornika parametrycznego dla dwóch wartości natężenia prądu wzbudzającego pole elektromagnetyczne ($I_w = 2$ mA, $I_w = 4$ mA) oraz dla $I_w = 4$ mA przy obecności stałego



Rys. 6. Zespolone napięcie przetwornika wirowprądowego parametrycznego, $H_s = 0$, $I_w = 2$ mA

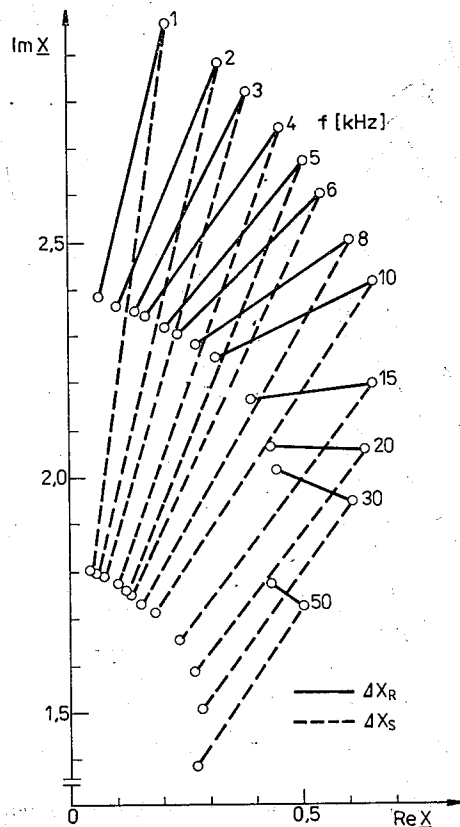


Rys. 7. Zespolone napięcie przetwornika wirowprądowego parametrycznego, $H_s = 0$, $I_w = 4$ mA

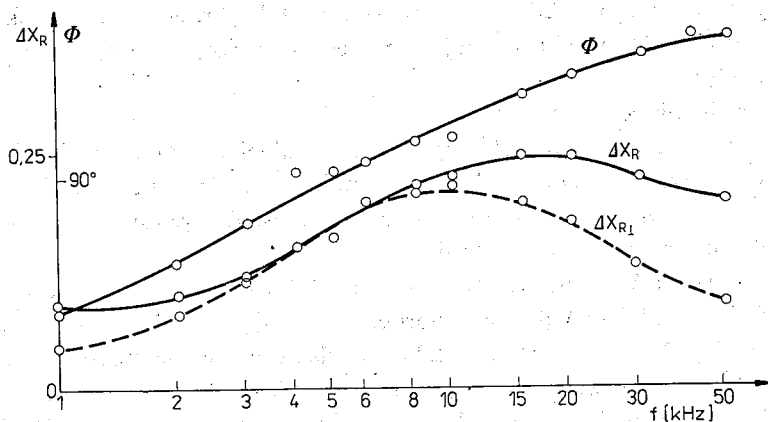
poła podmagnesowującego przedstawiono na rys. 6—8. Wykresy z rys. 6—8 skonstruowano w oparciu o charakterystyki obrazujące napięcie wyjściowe przetwornika dla czterech przypadków, gdy:

- 1) przetwornik nie zawiera pręta,
- 2) przetwornik zawiera odcinek pręta o nominalnej średnicy, pozbawiony wad,
- 3) przetwornik zawiera odcinek pręta o zmniejszonej średnicy (ok. 21,7%),
- 4) przetwornik zawiera odcinek pręta o nominalnej średnicy, zawierający wadę sztuczną (rowek głębokości około 36,5% średnicy).

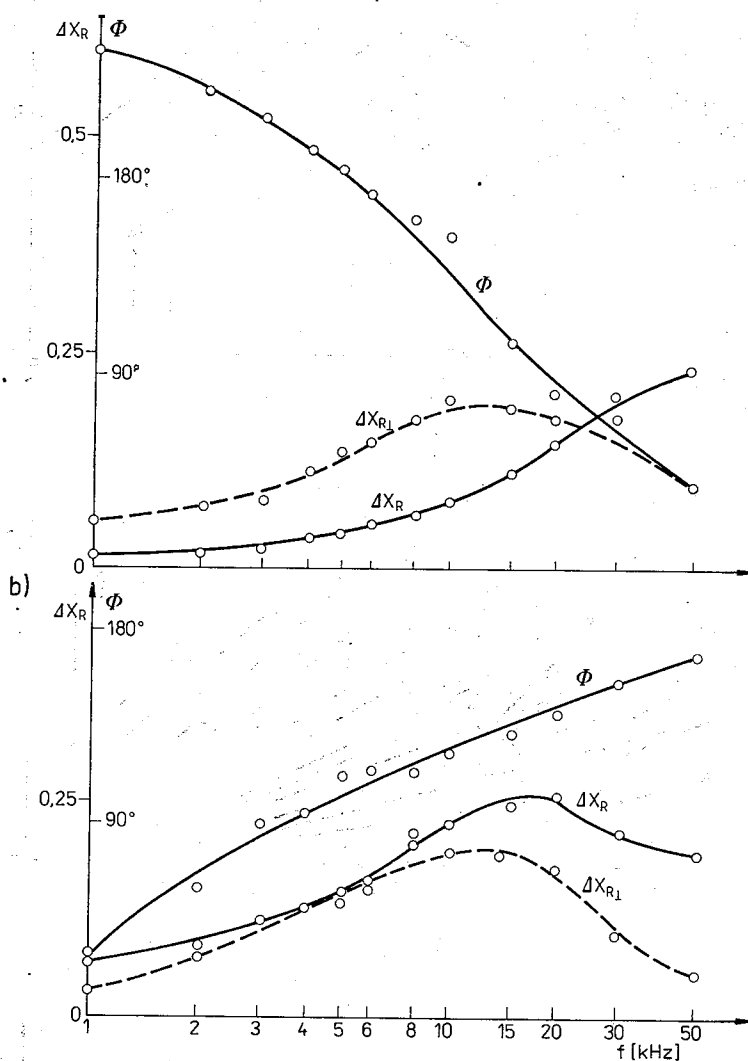
Współrzędne punktów opisanych wartością częstotliwości prądu zasilającego przedstawiają odpowiednio $Re X$ i $Im X$ dla przetwornika zawierającego odcinek pręta wzorcowego bez wady średnicy 2,3 mm. Wielkość ΔX_R przedstawia zmianę wartości X , wywołaną istnieniem wady w pręcie wzorcowym, natomiast wielkość ΔX_S odpowiada zmianie wartości X wynikającej ze zmniejszenia średnicy odcinka pręta wzorcowego ok. 21,7%. Przebieg składowych ΔX_R , $\Delta X_{R\perp}$ oraz Φ w funkcji częstotliwości prądu zasilającego przetwornik parametryczny przedstawiono na rys. 9 i 10.



Rys. 8. Zespolone napięcie przetwornika wiroprowadowego parametrycznego, $H_s = 2000$ A/m, $I_w = 4$ mA



Rys. 9. Zależność Φ , ΔX_R , ΔX_{R1} przetwornika wiroprowadowego parametrycznego od częstotliwości pola elektromagnetycznego, $H_s = 0$, $I_w = 2$ mA

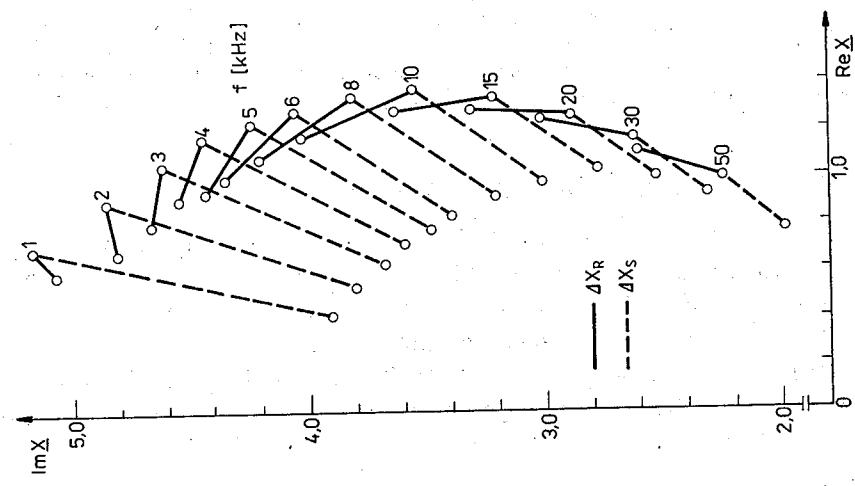


Rys. 10. Zależność Φ , ΔX_R , $\Delta X_{R\perp}$ przetwornika, wiroprowodowego parametrycznego od częstotliwości pola elektromagnetycznego: a) $H_s = 200$ A/m, $I_w = 4$ mA, b) $H_s = 0$, $I_w = 4$ mA

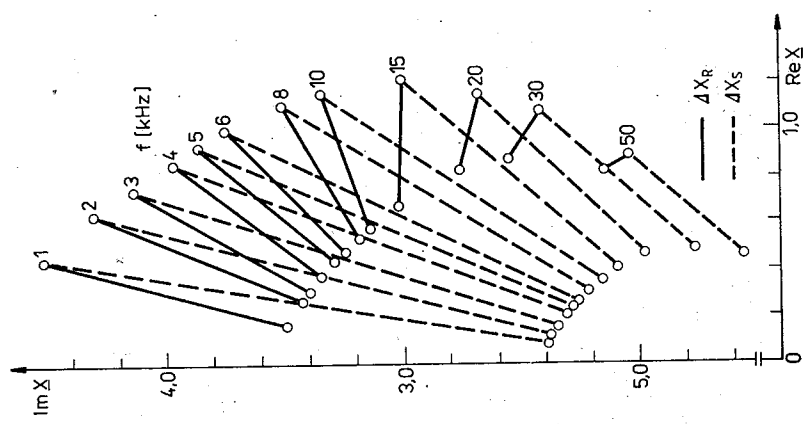
PRZETWORNIK GENERACYJNY

Na rys. 11—16 przedstawiono na płaszczyźnie zmiennej zespolonej składowe napięcia wyjściowego przetwornika dla trzech wartości prądu zasilającego przetwornik $I_w = 5$ mA, 10 mA i 15 mA, przy stałym polu magnetycznym $H_s = 0$ i $H_s = 2000$ A/m. W przetworniku znajdował się ten sam obiekt badany, co poprzednio.

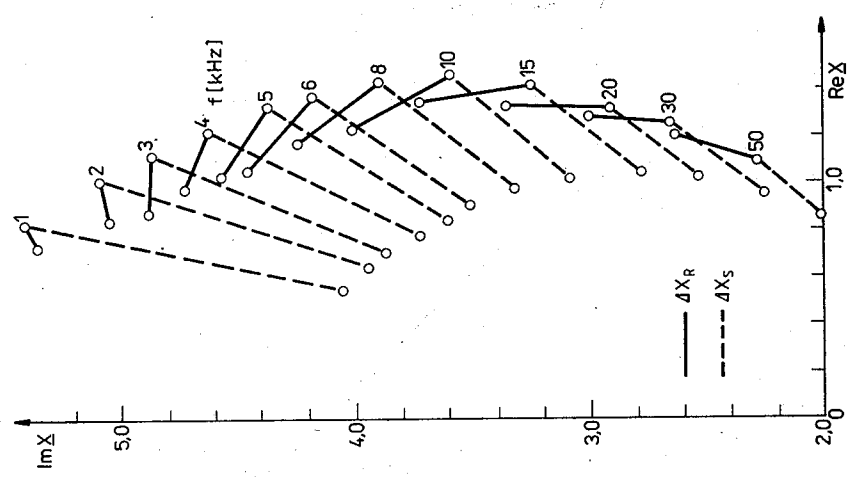
Przebieg składowych ΔX_R , $\Delta X_{R\perp}$ oraz Φ w funkcji częstotliwości prądu zasilającego przetwornik wiroprowodowy dla poszczególnych warunków pracy przetwornika przedstawiono na rys. 17—22.



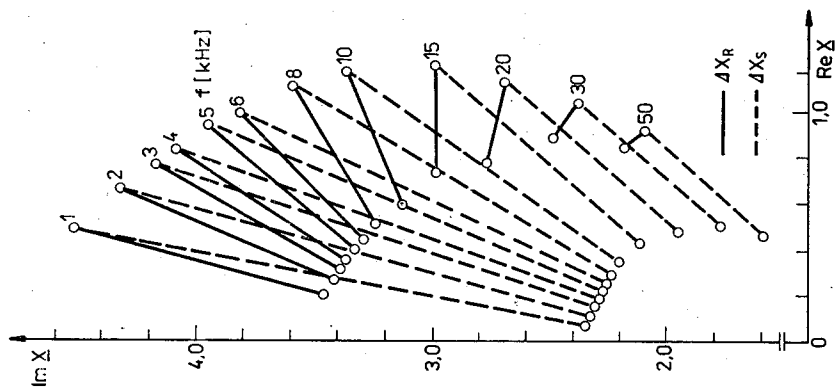
Rys. 11. Zespolone napięcie przetwornika wirowego generacyjnego, $H_s = 0$, $I_w = 5$ mA



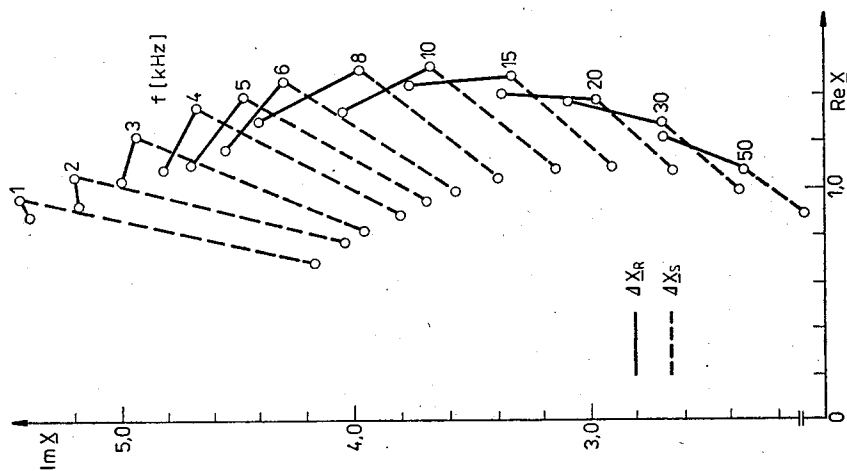
Rys. 12. Zespolone napięcie przetwornika wirowego generacyjnego, $H_s = 2000$ A/m, $I_w = 5$ mA



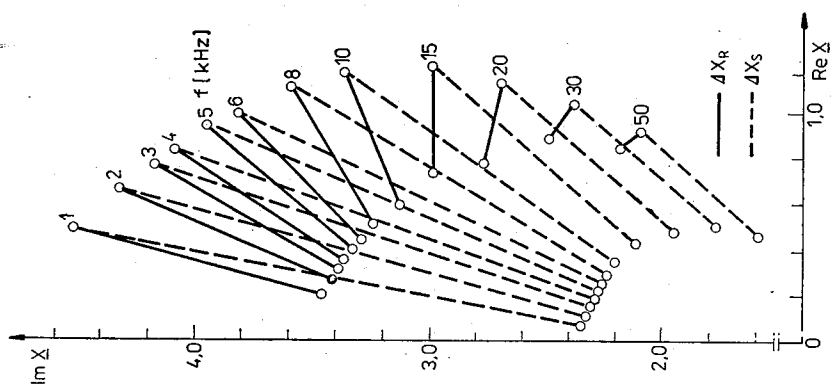
Rys. 13. Zespolone napięcie przetwornika wirowego generacyjnego, $H_s = 0$, $I_w = 10$ mA



Rys. 14. Zespolone napięcie przetwornika wirowego generacyjnego, $H_s = 2000$ A/m, $I_w = 10$ mA



Rys. 15. Zespolone napięcie przetwornika wirowego generacyjnego, $H_s = 0$, $I_w = 15$ mA



Rys. 16. Zespolone napięcie przetwornika wirowego generacyjnego, $H_s = 2000$ A/m, $I_w = 15$ mA

5. ANALIZA WŁAŚCIWOŚCI PRZETWORNIKÓW

WPLYW ROZKŁADU POLA MAGNETYCZNEGO PRZETWORNIKA WIROPRAĐOWEGO NA JEGO WŁAŚCIWOŚCI

Przetworniki o różnych rozkładach pola magnetycznego charakteryzują się znacznymi różnicami wpływu pręta stalowego na ich napięcie wyjściowe. Do porównania przyjęto dwa rodzaje przetworników: parametryczny i generacyjny. Rozkłady pola magnetycznego porównywanych przetworników przedstawiono na rys. 5. Porównanie właściwości może być przeprowadzone dla przypadku zbliżonej wartości średniej pola magnetycznego, obliczonej na długości uzwojenia wyjściowego. Przetwornik wiroprowdowy parametryczny, zasilany prądem o wartości $I_w = 2$ mA, przy braku stałego pola podmagnesowującego $H_s = 0$ charakteryzuje się średnią wartością osiowej składowej pola magnetycznego, obliczoną na długości uzwojenia pomiarowego, równą $H_s = 100$ A/m. Dla przetwornika wiroprowdowego generacyjnego przy $I_w = 15$ mA i $H_s = 0$ wartość ta wynosi $H_s = 105$ A/m.

Analizowane przetworniki wykazują zbliżone wartości składowych ich napięcia wyjściowego U_{I0} w przypadku braku obiektu badanego. W zakresie częstotliwości od 1 kHz do 50 kHz różnica wartości napięcia U_{I0} analizowanych przetworników wiroprowdowych nie przekracza 5%. Po uwzględnieniu błędu pomiarowego miernika składowych wynoszącego 4% oraz różnicy wartości średniej pola magnetycznego równej 5% można uznać, że składowe napięcia U_{I0} są sobie w przybliżeniu równe.

Składowe napięcia wyjściowego U_{R0} przetworników różnią się znacznie. Dla przetwornika generacyjnego wartość napięcia U_{R0} wynosi zero, natomiast dla przetwornika parametrycznego składowa napięcia wyjściowego U_{R0} przyjmuje wartości od 59 mV do 61 mV w rozważanym zakresie częstotliwości prądu zasilającego przetwornik od 1 kHz do 50 kHz. Różnice wartości napięcia U_{R0} wynikają ze względów układowych. Dla przetwornika parametrycznego rezystancja uzwojenia wyjściowego 30Ω wchodzi do obwodu pomiarowego, natomiast w przypadku przetwornika transformatorowego występuje ona w obwodzie wzbudzenia. Wartości U_{I0} dla rozpatrywanych przetworników pozwalają na stwierdzenie, że wartości średnie strumieni magnetycznych skojarzonych z uzwojeniem wyjściowym przetworników są sobie równe.

Wpływ obecności pręta w rozpatrywanych przetwornikach jest zdecydowanie różny. Wielkość $ReX = (U_R - U_{R0})/U_{I0}$ dla przetwornika parametrycznego przyjmuje początkowo wartości rosnące ze wzrostem częstotliwości od 0,3 dla $f = 1$ kHz, osiągając wartość maksymalną 0,75 dla $f = 10$ kHz. Dalszy wzrost częstotliwości powoduje zmniejszenie wartości ReX do 0,59 przy $f = 50$ kHz (rys. 6).

Natomiast dla przetwornika generacyjnego wartość ReX wynosi odpowiednio przy $f = 1$ kHz 0,96 osiągając maksimum 1,53 przy $f = 10$ kHz, a przy $f = 50$ kHz wynosi ono 1,09 (rys. 15). Występują również znaczne różnice wartości $ImX = U_I/U_{I0}$ dla analizowanych przetworników. Dla przetwornika parametrycznego wartości ImX w zakresie częstotliwości od 1 kHz do 50 kHz zmniejszają się od 3,21 przy $f = 1$ kHz do 1,81 przy $f = 50$ kHz (rys. 6). Odpowiednio dla przetwornika generacyjnego można zanotować zmiany wartości ImX od 5,41 przy $f = 1$ kHz do 2,33 przy $f = 50$ kHz (rys. 15).

Występują również znaczne różnice wartości ΔX_s spowodowanych zmianą średnicy pręta wzorcowego od $D_0 = 2,3$ mm do $D_0 = 1,8$ mm. Zmiany ΔX_s przetwornika generacyjnego są w przybliżeniu około 2-krotnie większe w porównaniu do przetwornika parametrycznego. Obecność wady w pręcie wywołuje zmiany ΔX_R również około 2-krotnie większe w przypadku przetwornika generacyjnego. Wartości maksymalne ΔX_R dla przetwornika parametrycznego występują przy $f = 15\text{--}20$ kHz i wynoszą $\Delta X_{R\max} = 0,25$ (rys. 9). Przetwornik generacyjny charakteryzuje się wartością maksymalną $\Delta X_{R\max} = 0,5$ przy $f = 12$ kHz (rys. 21). Podobna reakcja występuje w odniesieniu do składowej ΔX_R prostopadłej do ΔX_s i oznaczonej jako $\Delta X_{R\perp}$. Jedyną analizowaną wielkością, która przyjmuje zbliżone wartości dla obu rodzajów przetworników, jest kąt Φ zawarty między ΔX_R i ΔX_s .

Dokładne wyjaśnienie różnic właściwości przetworników wiroprowodowych o różnych rozkładach pola magnetycznego jest zagadnieniem bardzo trudnym. Można próbować wyjaśnić to przez wykorzystanie zależności (1), słusznych dla przetworników o jednorodnym polu magnetycznym. Wydaje się, że możliwe jest potraktowanie przenikalności magnetycznej względnej μ_r występującej w zależnościach (1) jako wielkości μ_{sr} , charakterystycznej dla określonego obiektu badanego, umieszczonego w konkretnym przetworniku wiroprowodowym o niejednorodnym polu magnetycznym. Wartość μ_{sr} należy uśrednić w całym obszarze cewki pomiarowej przetwornika. Oczywiście, pozostaje wyznaczenie wartości μ_{sr} , na którą ma wpływ zarówno przenikalność magnetyczna względna materiału μ_r , z którego obiekt badany jest wykonany, jak również rozkład pola magnetycznego przetwornika wiroprowodowego i związane z tym rozkładem pole odmagnesowujące. Dokładne ujęcie zjawisk związanych z wyznaczeniem właściwości magnetycznych obiektów badanych wykracza poza ramy niniejszej pracy.

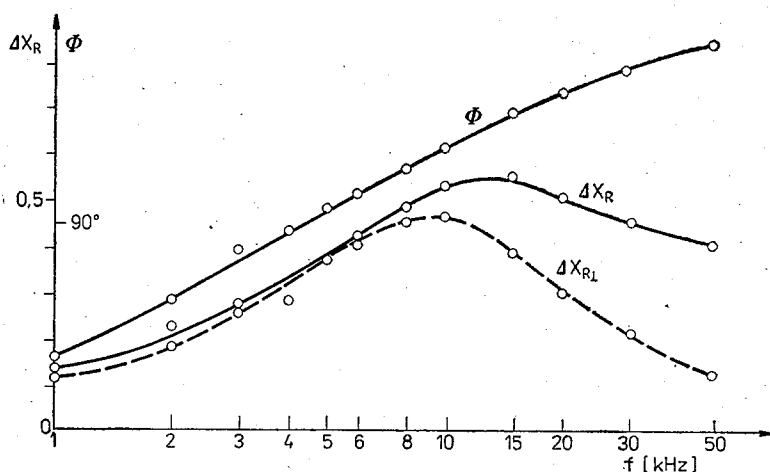
WPLYW CZĘSTOTLIWOŚCI PRĄDU WZBUDZAJĄCEGO POLA ELEKTROMAGNETYCZNE NA WŁAŚCIWOŚCI PRZETWORNIKA

Wartość częstotliwości prądu wzbudzającego pole elektromagnetyczne w przetwornikach wiroprowodowych ma zasadniczy wpływ na zmiany napięcia wyjściowego przetworników w zależności od parametrów pręta. Zmniejszenie się średnicy pręta z wartości nominalnej 2,3 mm do wartości 1,8 mm powoduje zmianę wartości X o wartość ΔX_s . Maksymalne zmiany wartości ΔX_s występują dla obu przetworników dla najmniejszych wartości częstotliwości prądu. W miarę wzrostu częstotliwości wartości ΔX_s maleją (rys. 6—8, 11—16).

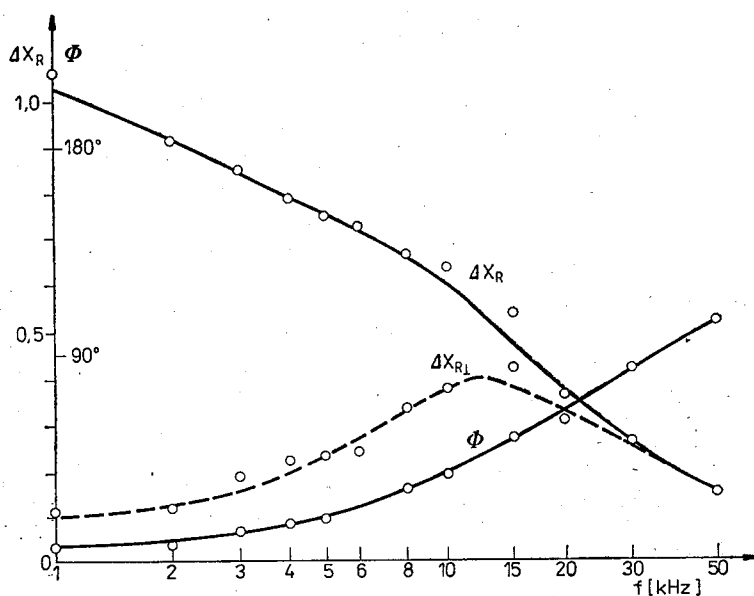
Występowanie w pręcie wady o głębokości 0,8 mm powoduje zmianę X o wartość ΔX_R . Dla przetwornika wiroprowodowego generacyjnego, przy wartości stałego pola podmagnesowującego H_s równej zero, wartości ΔX_R zmieniają się od około 0,1 przy $f = 1$ kHz osiągając maksimum $\Delta X_R = 0,5$ przy f zbliżonej do 10 kHz, a następnie maleją osiągając przy $f = 50$ kHz wartość około 0,37 (rys. 17, 19, 21).

Wartości kąta Φ zawartego między ΔX_R i ΔX_s zmieniają się od około 20° przy $f = 1$ kHz, osiągając 90° przy częstotliwości zbliżonej do 5 kHz, do około 160° przy $f = 50$ kHz.

Składowe ΔX_R prostopadłe do ΔX_s oznaczone jako $\Delta X_{R\perp}$ przyjmują wartości maksymalne około 0,45 przy częstotliwości równej 10 kHz (rys. 17, 19, 21).

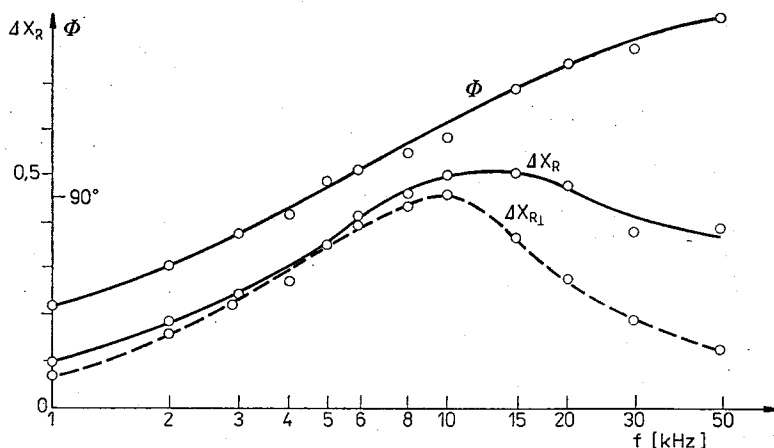


Rys. 17. Zależność Φ , ΔX_R , $\Delta X_{R\perp}$ przetwornika wiroprowodowego generacyjnego od częstotliwości pola elektromagnetycznego, $H_s = 0$, $I_w = 5$ mA

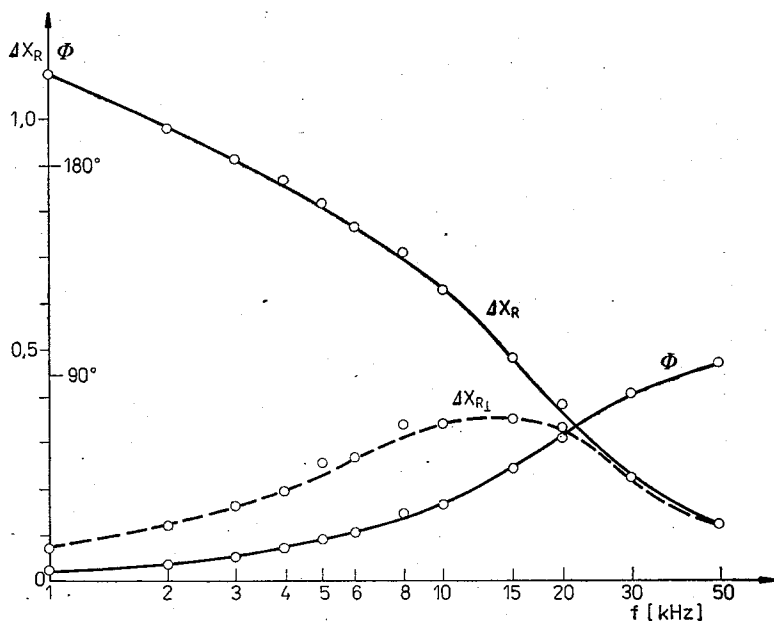


Rys. 18. Zależność Φ , ΔX_R , $\Delta X_{R\perp}$ przetwornika wiroprowodowego generacyjnego od częstotliwości pola elektromagnetycznego $H_s = 2000$ A/m, $I_w = 5$ mA

Dla przetwornika wiroprowodowego parametrycznego charakter zmian wielkości ΔX_s , ΔX_R , $\Delta X_{R\perp}$ i Φ jest podobny z tym, że maksymalne wartości ΔX_R i $\Delta X_{R\perp}$ są dwa razy mniejsze. Zmiany wartości X w zależności od częstotliwości prądu wzbudzającego pola elektromagnetyczne, przetworników zawierających pręt o średnicy 2,3 mm i 1,8 mm przebiegają podobnie, jak to wynika z zależności (1), słusznych dla przetworników o jednorodnym polu magnetycznym. Możliwe jest, przy spełnieniu pewnych warunków, wykorzystanie zależności (1) do analizy właściwości przetworników wiroprowodowych o niejedno-

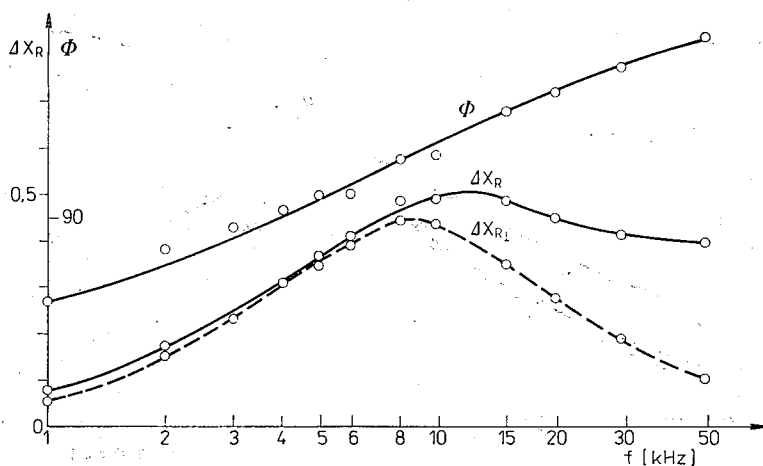


Rys. 19. Zależność Φ , ΔX_R , ΔX_{R_L} przetwornika wiroprowadowego generacyjnego od częstotliwości pola elektromagnetycznego, $H_s = 0$, $I_w = 10$ mA

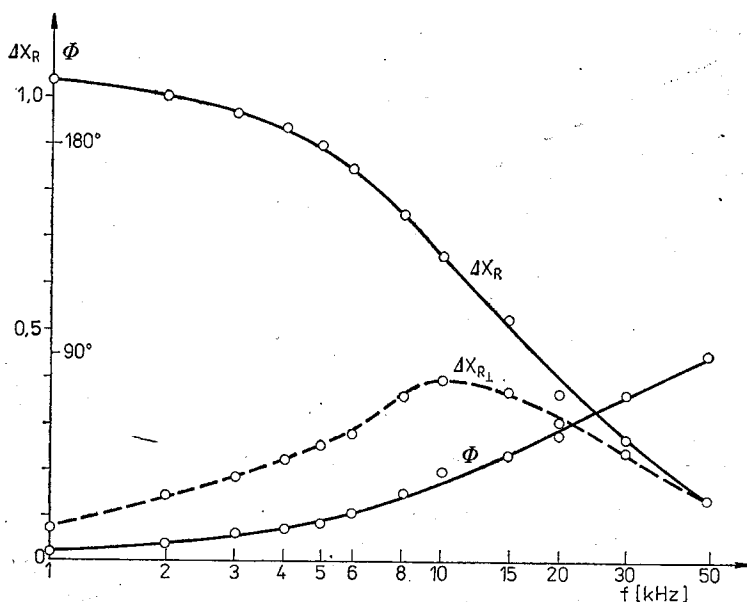


Rys. 20. Zależność Φ , ΔX_R , ΔX_{R_L} przetwornika wiroprowadowego generacyjnego od częstotliwości pola elektromagnetycznego, $H_s = 2000$ A/m, $I_w = 10$ mA

rodnym polu elektromagnetycznym. W tym celu należy przenikalność magnetyczną względną μ_r występującą w zależnościach (1) zastąpić przenikalnością magnetyczną średnią μ_{sr} , charakterystyczną dla obiektu o kształcie walca, umieszczonego w polu magnetycznym o określonym rozkładzie. Przenikalność skuteczna μ_s występująca w zależnościach (1), charakterystyczna dla obiektu o kształcie walca, umieszczonego w jednorodnym polu magnetycznym, wymaga zastąpienia przenikalnością skuteczną μ_s obiektu umieszczonego w niejednorodnym polu magnetycznym o określonym rozkładzie. Pozostaje wyznaczenie



Rys. 21. Zależność Φ , ΔX_R , ΔX_{R_L} przetwornika wirowodowego generacyjnego od częstotliwości pola elektromagnetycznego, $H_s = 0$, $I_w = 15$ mA



Rys. 22. Zależność Φ , ΔX_R , ΔX_{R_L} przetwornika wirowodowego generacyjnego od częstotliwości pola elektromagnetycznego, $H_s = 2000$ A/m, $I_w = 15$ mA

wartości przenikalności magnetycznej średniej μ_{sr} oraz przenikalności skutecznej μ_s obiektu, umieszczonego w niejednorodnym polu magnetycznym.

Przedstawiono tu jedną z metod postępowania, która może doprowadzić do uzyskania zadowalających wyników analizy teoretycznej przetworników o niejednorodnym polu elektromagnetycznym. Zasygnalizowane propozycje wymagają wnikliwych badań i analizy opartej na teorii pola elektromagnetycznego, co wykracza poza ramy niniejszego opracowania.

WPLYW WARTOŚCI PRĄDU WZBUDZAJĄCEGO POLE ELEKTROMAGNETYCZNE NA WŁAŚCIWOŚCI PRZETWORNIKÓW

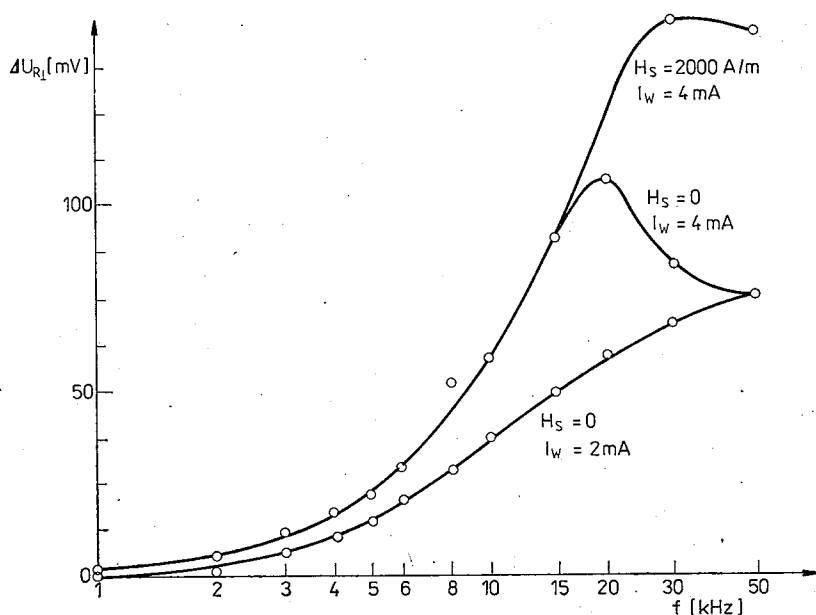
Wzrost wartości prądu wzbudzającego pole elektromagnetyczne powoduje nieznaczny wzrost wartości unormowanych składowych napięcia wyjściowego przetwornika. Zwiększenie wartości prądu powoduje wzrost wartości pola magnetycznego przetwornika. Ze względu na nieliniową charakterystykę magnesowania materiału prętów stalowych wzrost wartości pola magnetycznego powoduje zwiększenie wartości przenikalności magnetycznej amplitudowej pręta. Większa wartość przenikalności magnetycznej powoduje większy przyrost strumienia magnetycznego, skojarzonego z uzwojeniem pomiarowym i wzrost wartości X .

Wartości składowych ΔX_R i $\Delta X_{R\perp}$ nie zmieniają się przy wzroście wartości prądu. Jeżeli czułość wykrywania wad przetwornika wiroprowadowego określić za pośrednictwem składowych ΔX_R czy $\Delta X_{R\perp}$, to można dojść do wniosku, że czułość wykrywania wad analizowanych przetworników wiroprowadowych nie zależy od wartości prądu wzbudzającego.

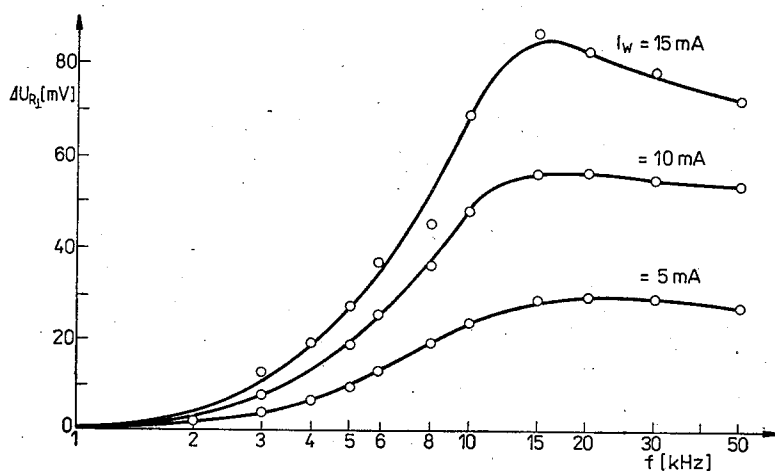
W praktycznych układach defektoskopów, wykorzystujących napięcia wyjściowe przetwornika wiroprowadowego, czułość wykrywania wad powinna być jednak określana przez bezwzględną zmianę napięcia wyjściowego przetwornika, wywołaną istnieniem wady. Wartości napięć $\Delta U_{R\perp}$, wywołanych istnieniem wady, opisuje zależność:

$$\Delta U_{R\perp} = \Delta X_R \cdot U_{I0} \quad (6)$$

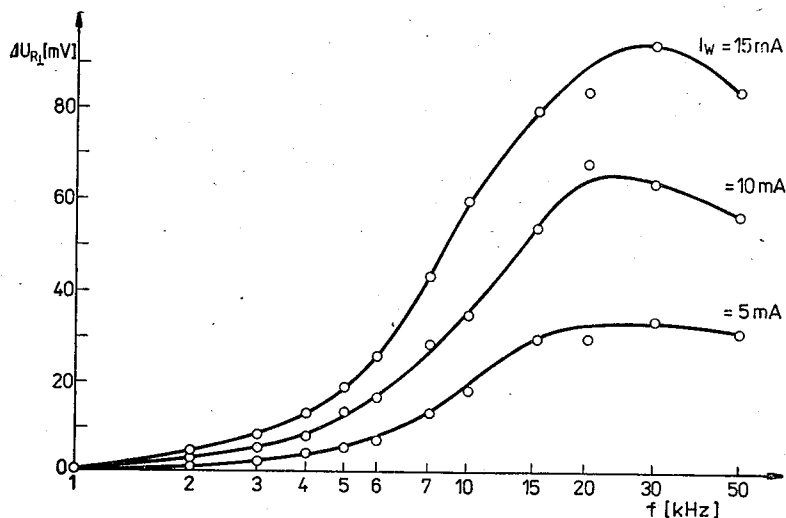
Zależność $\Delta U_{R\perp}$ od częstotliwości, dla różnych warunków pracy przetwornika, przedstawiono na rys. 23—25.



Rys. 23. Zależność $\Delta U_{R\perp}$ przetwornika parametrycznego od częstotliwości pola elektromagnetycznego



Rys. 24. Zależność $\Delta U_{R\perp}$ przetwornika generacyjnego od częstotliwości pola elektromagnetycznego, $H_s = 0$



Rys. 25. Zależność $\Delta U_{R\perp}$ przetwornika generacyjnego od częstotliwości pola elektromagnetycznego, $H_s = 2000 \text{ A/m}$

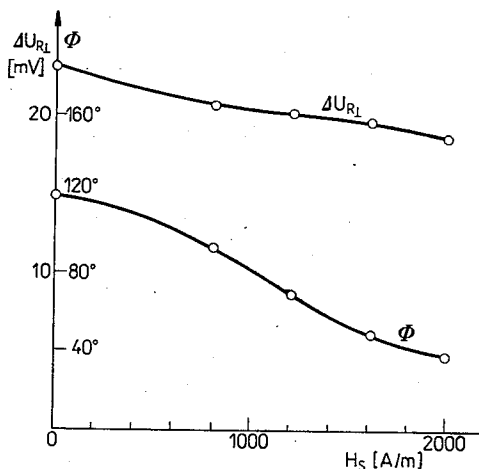
Z porównania wykresów $\Delta U_{R\perp} = f(f)$ i $\Delta X_{R\perp} = f(f)$ wynika, że maksymalne wartości $\Delta U_{R\perp}$ i $\Delta X_{R\perp}$ występują przy różnych wartościach częstotliwości, natomiast wzrost wartości prądu I_w powoduje wzrost wartości napięcia $\Delta U_{R\perp}$, co świadczy o wzroście czułości wykrywania wad.

WPLYW STAŁEGO POLA MAGNETYCZNEGO NA WŁAŚCIWOŚCI PRZETWORNIKÓW

Przy małych zmianach wartości zmiennego pola magnetycznego i obecności stałego pola magnetycznego, właściwości magnetyczne materiałów ferromagnetycznych określa się za pomocą przenikalności magnetycznej rewersyjnej. Przy wzroście wartości stałego

pola magnetycznego wartość przenikalności magnetycznej rewersyjnej maleje. Obecność stałego pola magnetycznego U_s powoduje zmniejszenie wartości ReX i ImX (rys. 8, 12, 14, 16).

Inny jest przebieg zależności ΔX_R od częstotliwości. Maksymalne wartości ΔX_R , wywołane istnieniem wady, występują przy najmniejszych wartościach częstotliwości (rys. 18, 20, 22). Kąt zawarty między ΔX_R i ΔX_s przyjmuje wartości od kilku stopni przy częstotliwości $f = 1$ kHz do około 90° przy $f = 50$ kHz. Wartości ΔX_{RL} , osiągane w przypadku istnienia stałego pola magnetycznego H_s , są nieco mniejsze od wartości ΔX_{RL} , występujących w warunkach bez pola H_s . Natomiast maksymalne wartości napięć ΔU_{RL} są większe przy obecności stałego pola magnetycznego H_s , ponieważ występują przy większych wartościach częstotliwości.



Rys. 26. Zależność napięcia wyjściowego ΔU_{RL} przetwornika wirowego generacyjnego oraz kąta Φ od stałego pola elektromagnetycznego H_s , $I_w = 5$ mA, $f = 10$ kHz

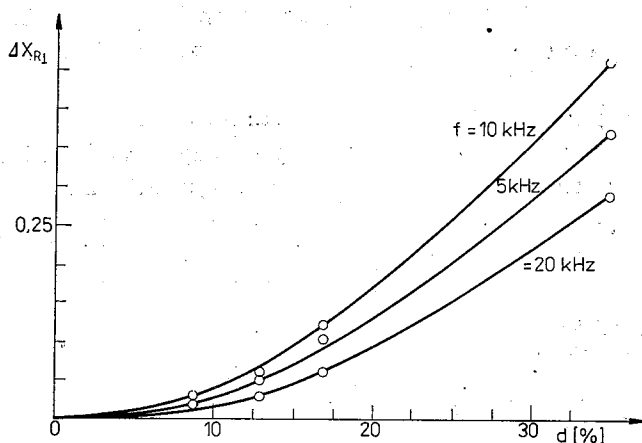
Wyniki obliczeń wartości napięć ΔU_{RL} i kąta Φ przetwornika generacyjnego zawierającego pręt, przy różnych wartościach stałego pola magnetycznego H_s przy częstotliwości $f = 10$ kHz oraz prądzie wzbudającym $I_w = 5$ mA przedstawiono na rys. 26. Z przedstawionych wykresów wynika istotna zależność wartości kąta Φ od wartości stałego pola magnetycznego H_s . W niewielkim stopniu natomiast zmienia się wartość napięcia ΔU_{RL} .

WPLYW WAD SZTUCZNYCH O RÓŻNYCH GŁĘBOKOŚCIACH

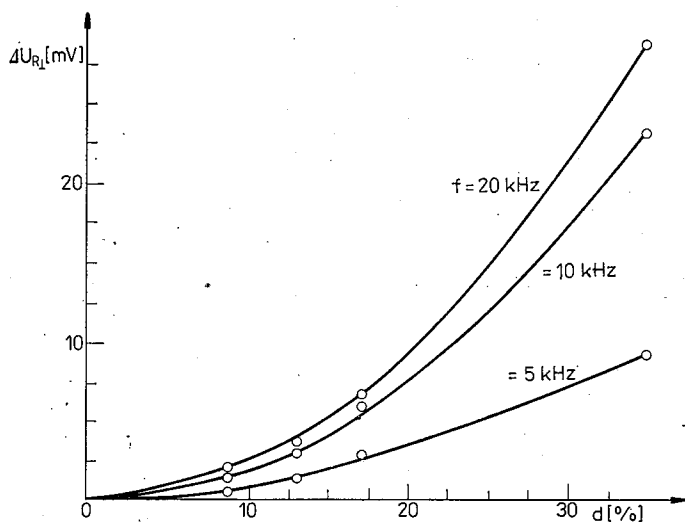
Obecnie zostanie przedstawiony wpływ wad sztucznych o różnych głębokościach na właściwości przetwornika generacyjnego, pracującego w układzie porównawczym. Układ taki zastosowano w zrealizowanych modelach defektoskopów. Charakterystyki przetwornika przedstawiono na rys. 27 i 28.

Z wykresów przedstawionych na rys. 28 wynika, że maksymalne wartości napięć ΔU_{RL} występują dla $f = 20$ kHz.

Wyniki zamieszczone na rys. 27 i rys. 28 potwierdzają wyniki przedstawione na rys. 17 i rys. 24.



Rys. 27. Zależność $\Delta X_{R\perp}$ przetwornika generacyjnego od głębokości wad sztucznych, $H_s = 0$, $I_w = 5$ mA

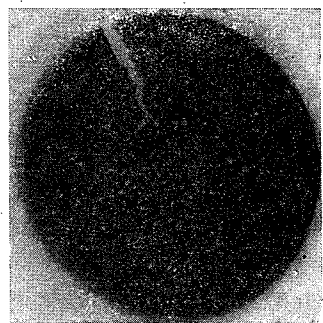
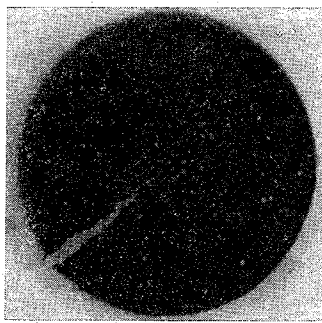
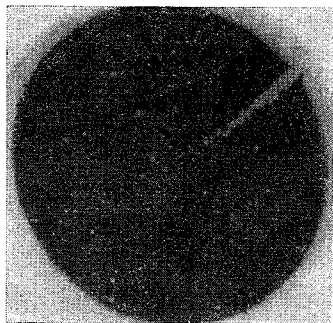
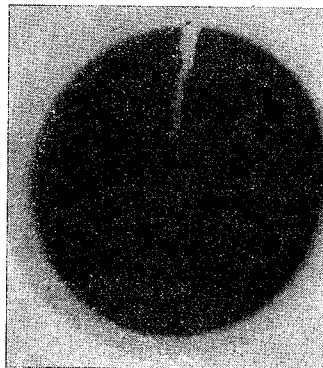
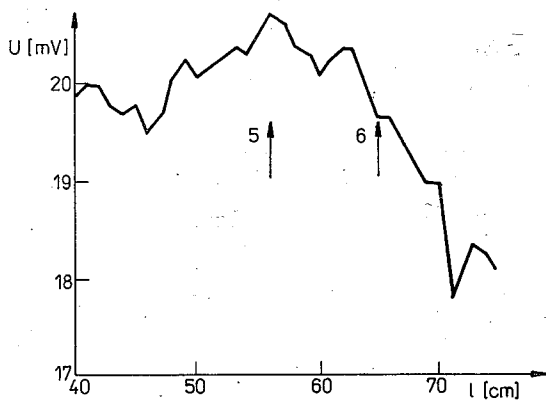
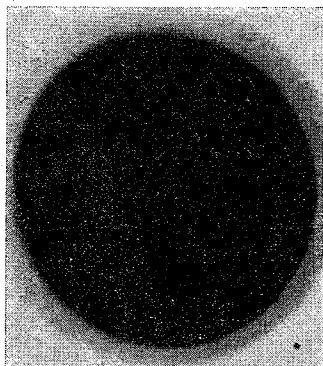
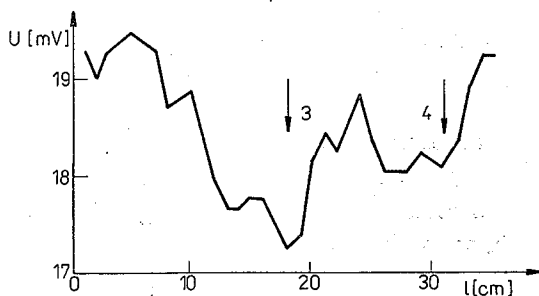
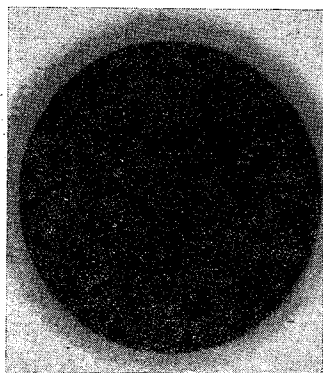
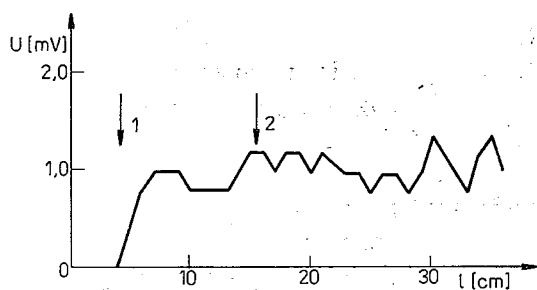


Rys. 28. Zależność napięcia $\Delta U_{R\perp}$ przetwornika generacyjnego od głębokości wad sztucznych, $H_s = 0$, $I_w = 5$ mA

6. WYBRANE CHARAKTERYSTYKI PRZETWORNIKA DLA OBIEKTÓW Z WADAMI NATURALNYMI

Na rys. 29 przedstawiono wybrane charakterystyki napięcia przetwornika zawierającego pręty średnicy 2,3 mm, wykonane ze stali chromowej oraz fotogramy zgładów metalograficznych poprzecznych, wykonanych dla miejsc oznaczonych na wykresach strzałkami.

Na podstawie przeprowadzonych badań metalograficznych można było dokonać ogólnej analizy rodzajów wad, występujących w przykładowym obiekcie badań, tj. w prętach ciągnionych szlifowanych, wykonanych ze stali chromowej. Wyniki tej analizy prowadzą do następujących wniosków:



Rys. 29. Napięcie przetwornika generacyjnego zawierającego pręty ze stali chromowej o średnicy 2,3 mm oraz fotogramy zglądów metalograficznych

1. Wady występujące w prętach mają charakter pojedynczych pęknięć o znacznej szerokości, przebiegających w płaszczyźnie w przybliżeniu równoległej do promienia.
2. Głębokość pęknięć wynosi najczęściej 30%—40% średnicy pręta.
3. Szerokość pęknięć wynosi około 11% do około 15% ich głębokości.
4. Wszystkie zaobserwowane pęknięcia mają charakter pęknięć powierzchniowych.
5. Obserwuje się różne natężenie występowania pęcherzy powietrznych, zarówno w miejscach pozbawionych pęknięć, jak i w miejscach zawierających pęknięcia.
6. W poszczególnych partiach prętów, wynoszących 100 sztuk, obserwuje się po kilka prętów z pęknięciem na całej długości (2 m). Należy nadmienić, że z 2-metrowego odcinka pręta stali chromowej wykonywane jest np. 100 przepustów próżnioszczelnych (przez szkło ołowiowe), przeznaczonych do zastosowania w produkcji lamp elektronowych i izolatorów do agregatów chłodniczych.

Przedstawione tu przykładowe wyniki badań dokumentują przydatność opracowanego przetwornika [19] i układu analizy [20, 23] sygnałów do wykrywania nieciągłości prętów i drutów stalowych. W szczególności można zauważyć, że stosunek sygnałów, pochodzących od pęknięć do sygnałów zakłócających, powstających w wyniku niejednorodności struktury wynosi około 26 dB, co świadczy o dobrej wykrywalności pęknięć.

7. WIELKOŚCI WPŁYWAJĄCE NA WYKRYWALNOŚĆ PĘKNIĘĆ. PODSUMOWANIE

Jak już wskazano, rozkład prądów wirowych jest uzależniony od rozmaitych wielkości. Najważniejszymi z nich są: przewodność elektryczna γ , przenikalność magnetyczna μ , średnica D_p elementu metalowego oraz częstotliwość wzbudzenia f . Według teorii procedury kontrolnej metodą prądów wirowych ze wspomnianych wyżej wielkości wynika dalsza, fundamentalna dla badania prądami wirowymi wielkość, tak zwana częstotliwość graniczna f_g . Oblicza się ją dla materiału pełnego (druty i pręty) wg zależności (2) [9, 26].

Jak wykazano, wpływ przewodności, przenikalności i wymiarów obiektu na rozkład prądów wirowych pozostaje w bardzo silnej zależności od stosunku częstotliwości f do częstotliwości granicznej f_g . Oznacza to, że przy właściwym doborze stosunku f/f_g wpływ niepożądanych wielkości zakłócających może być zmniejszony. W ten sposób, po obliczeniu częstotliwości granicznej f_g dla każdego badanego obiektu, można wybrać częstotliwość wzbudzenia odpowiednią dla wykrywalności pęknięć.

Przetworniki wiroprowodowe przelotowe stosowane są w zasadzie do wykrywania nieciągłości wzdłużnych, gdyż tego rodzaju wady powodują zakłócenie przepływu prądów wirowych. Jednak niektóre większe wady poprzeczne też są wskazywane, ze względu na to, że stanowią one miejsca o zwiększonej reluktancji, w okolicy których następuje rozproszenie strumienia magnetycznego.

Podstawową wadą defektoskopii indukcyjnej, opartej na wykorzystaniu zjawiska prądów wirowych jest tzw. sumowanie wad znajdujących się w obrębie przetwornika pomiarowego — przy czym jako wady uważa się tu zarówno nieciągłości występujące w danym przekroju, jak i wahania właściwości fizykochemicznych i rozmiarów. Dla wykrywalności

pęknięć ma jednak znaczenie nie tylko wielkość bezwzględna sygnału kontrolnego, lecz raczej odstęp tego sygnału kontrolnego od innych, nie dających się uniknąć wielkości zakłócających. Poniżej zostanie opisany wpływ rozmaitych wielkości oddziałujących na wynik badania.

WPLYW ZMIAN ROZMIARÓW

Napięcia, które są indukowane w uzwojeniu wtórnym przez pęknięcie, jak również zmianę rozmiarów, są wielkościami zespolonymi. Oba napięcia różnią się między sobą pod względem fazowym: kąt fazowy pomiędzy obydwooma efektami zależny jest od stosunku f/f_g . Dzięki selektywnemu układowi fazowemu wpływ zmiany średnicy może zostać w znacznym stopniu wyeliminowany. Jako sygnał kontrolny wykorzystana zostaje jedynie składowa napięcia pochodząca od pęknięcia, o kierunku prostopadłym do wektora napięcia wywołanego przez zmiany średnicy.

Możliwość wyodrębniania sygnałów pochodzących od pęknięć od sygnałów wywołanych zmianami średnicy obiektu, a tym samym i dobór optymalnego stosunku f/f_g wynika z kompromisu pomiędzy możliwie największą czułością wskazań pochodzących od pęknięć, a różnicą faz Φ możliwie najbardziej zbliżoną do 90° .

WPLYW PRZENIKALNOŚCI MAGNETYCZNEJ

Wpływ wahań przenikalności magnetycznej istnieje tylko w materiałach ferrytycznych i austenitycznych. W metalach tych przenikalność względna μ_r jest większa od 1. Przenikalność magnetyczna względna μ_r różnych obiektów ferromagnetycznych, które w praktyce podlegają badaniom, może przybierać wartości do kilku tysięcy. Na skutek np. naprężeń wewnętrznych występujących w materiale przy przeciąganiu, prostowaniu, walcowaniu i innych procesach lub też przy obróbce cieplnej, mogą pojawiać się silne, ograniczone co do miejsca wahania przenikalności. Wahania przenikalności magnetycznej mogą być spowodowane niejednorodnością składu chemicznego i struktury (nawet w obrębie tego samego wycięcia).

Zmiany takie w każdym przypadku powodują zakłócenia sygnału kontrolnego. Ażeby wykluczyć tego rodzaju szkodliwe wpływy podczas badań, próbkę ferromagnetyczną magnesuje się aż do nasycenia w odpowiednim urządzeniu. Przenikalność magnetyczna względna próbki dąży przy tym do wartości $\mu_r = 1$ w praktyce — do $\mu_r = 2 \div 3$. Próbka ferromagnetyczna zachowuje się wtedy pod względem jej właściwości magnetycznych, jak nieferromagnetyczna. Wpływ przenikalności zostaje w ten sposób wyeliminowany. Spadek przenikalności względnej do wartości $\mu_r = 1$ występuje również w metalach ferromagnetycznych przy podgrzewaniu do temperatur powyżej punktu Curie. W temperaturach tych, jak wiadomo, metal ferromagnetyczny zachowuje się tak, jak nieferromagnetyczny. Oddziaływujące na sygnał kontrolny wahania przenikalności nie występują więc. Magnesowanie nie jest zatem nieodzowne. Należy nadmienić, że kontrola wiroprądowa może być prowadzona dla obiektów o wysokich temperaturach.

WPŁYW ZMIAN PRZEWODNOŚCI ELEKTRYCZNEJ

Pojawiające się w praktyce zmiany przewodności wewnątrz badanej próbki nie są z reguły duże, co więcej — zmiany te przebiegają stopniowo, co oznacza, że brak tu jest stromego, ograniczonego co do miejsca, narastania, jak to zwykle ma miejsce przy pęknięciu.

Sygnały wywołane przez zmiany przewodności, jak wykazały doświadczenia — są podobne do tych, które wytwarzane są przez pęknięcia. Ich położenie fazowe jest niemal takie samo. Rozdzielenie tych wpływów nie może zatem następować poprzez analizę fazową sygnałów przetwornika. Tego rodzaju przebiegające stopniowo wpływy zakłócające zostają jednak przy zastosowaniu przetworników w układzie porównawczym lub samoporównawczym znacznie stłumione.

WPŁYW NIEDOKŁADNOŚCI PROWADZENIA

Ponieważ pole magnetyczne we wnętrzu przetwornika (składowa wzdłuż promienia) nie jest dokładnie homogeniczne, przy niedokładnym prowadzeniu badanych próbek występują sygnały zakłóceńowe. Często mają one to samo położenie fazowe, co i sygnały zakłóceńowe wytwarzane przez zmiany przekroju i przy tłumieniu tychże są one tym samym również w znacznym stopniu eliminowane.

Okazało się jednak w praktyce, że sygnały zakłóceńowe, wywołane wskutek niedokładności prowadzenia mogą również charakteryzować się różnymi kierunkami fazowymi. Stąd nie we wszystkich przypadkach jest możliwe tłumienie tego wpływu. Możliwości uniknięcia tego rodzaju wpływów zakłócających stanowią z jednej strony możliwie największy współczynnik wypełnienia cewki, tzn. że przetwornik pomiarowy powinien być możliwie najlepiej wypełniony przez obiekt badany, z drugiej zaś strony należy starać się o niezawodne, pozbawione uderzeń i szarpnięć, prowadzenia badanego obiektu w urządzeniu transportującym.

WPŁYW CHROPOWATOŚCI POWIERZCHNI

Dla wykrywania wad materiałowych, np. pęknięć, decydujące znaczenie ma to, w jaki sposób wady te pod względem swojej głębokości i swego kształtu odróżniają się od chropowatości powierzchni, nie przedstawiających zagrożenia dla późniejszego użytkowania wyrobu. Ogólnie wykrywalność wad jest tym lepsza, im mniejsza jest chropowatość powierzchni badanej próbki. Dla uzyskania dostatecznej wykrywalności pęknięć obiektu tak zwany stosunek sygnału użytecznego do zakłócenia wynosić powinien co najmniej 2:1. Oznacza to, że nieciągłości, które mają być wykryte muszą wywołać sygnały co najmniej dwa razy większe od sygnałów dla największych ujawniających się chropowatości powierzchni.

Stłumienie, względnie zmniejszenie sygnałów zakłóceńowych, pochodzących od chropowatości powierzchni następuje z jednej strony przez odjęcie tych sygnałów w przetwornikach w układzie porównawczym lub samoporównawczym. Z drugiej strony można osiągnąć stłumienie, względnie zmniejszenie tego wpływu, za pomocą układu selektywnego pod względem częstotliwości.

ODDZIAŁYWANIE POCZĄTKU I KOŃCA BADANEJ PRÓBKII

Przy wchodzeniu i wychodzeniu próbki do i z układu przetwornika pomiarowego indukowane zostaje każdorazowo przez początek obiektu i jego koniec napięcie w uzwojeniach wtórnych. Zjawisko to występuje zarówno dla przetworników w układzie bezwzględny, jak i w układzie porównawczym i samoporównawczym. Napięcie to jest w każdym przypadku znacznie większe niż wywołane przez wady napięcie pomiarowe. Zjawisko to może być zarówno pożądanym, jak i niepożądanym.

Pożądanym jest „efekt końca” obiektu przy rejestracji sygnałów pochodzących od wad za pośrednictwem urządzenia piszącego. Na taśmie może być dzięki temu zarejestrowany początek próbki i jej koniec na podstawie wytworzonego przez „efekt końca” impulsu napięcia o względnie dużej wartości.

„Efekt końca” jest niepożądanym w badaniach, w których impulsy napięcia wykorzystywane są do automatycznego sterowania jednostek, opracowujących zamierzone wartości napięć. „Efekt końca” może tu nawet wywoływać złudzenie sygnału nieciągłości i powodować błędne sterowanie jednostką opracowującą zmierzone wartości napięć. Aby tego uniknąć, dzięki zastosowaniu elektronicznych jednostek blokujących, proces kontroli rozpoczyna się dopiero wtedy, kiedy badana próbka zostaje całkowicie wprowadzona do przetwornika względnie, że przed wprowadzeniem końca badanej próbki wyłączone zostają układy elektronicznego opracowania sygnałów. Na skutek powyższego istnieją na początku badanej próbki i na jej końcu końcówki niesprawdzone. Długość niesprawdzonych końcówek jest z jednej strony zależna od wymiarów stosowanego układu przetwornika pomiarowego, a z drugiej zaś od materiału podlegającego badaniu. W wyniku odpowiedniego ekranowania przetwornika możliwe jest zmniejszenie długości końcówek niesprawdzonych.

8. MOŻLIWOŚĆ TŁUMIENIA SYGNAŁÓW ZAKŁÓCAJĄCYCH

Poszczególne wpływy zakłócające wynik badania oraz sposoby ich tłumienia zostały już omówione w poprzednich rozdziałach. Poniżej zostaną opisane trzy spośród różnych możliwości tłumienia wielkości zakłócających.

SELEKCJA FAZOWA

Napięcie zakłócające, wywołane np. przez zmiany przekroju ma pod względem fazowym inny kierunek, aniżeli napięcie pochodzące od pęknięcia. W celu oddzielenia napięcia od wady od napięcia zakłócającego, to ostatnie może zostać „obrócone” do kierunku poziomego na płaszczyźnie zmiennej zespolonej poprzez odpowiednie przesunięcie fazy. Wektor napięcia, który leży w linii poziomej, nie zostaje doprowadzony dalej do wykorzystania przez układy analizy fazoczułej. Jedynie składowa napięcia pomiarowego o kierunku pionowym jest wykorzystywana.

METODA PORÓWNAWCZA

Kolejną możliwość oddzielenia napięcia pochodzącego od pęknięcia od napięć zakłócających reprezentuje ich zróżnicowane zachowanie się w czasie. Wady takie, jak pęknięcia, przedstawiają najczęściej pewną, silną, ściśle ograniczoną co do miejsca zmianę w materiale. W związku z tym wady te powodują szybko zmieniające się napięcia w przetworniku pomiarowym.

Natomiast zupełnie odmienny jest przebieg napięcia w jednym uzwojeniu wtórnym, które powstaje w przypadku stopniowej zmiany przewodności właściwej. Na podstawie faktu odmiennego w czasie zachowania się poszczególnych napięć można poprzez zastosowanie metody porównawczej uzyskać to, że napięcia wytwarzane przez pęknięcia są podawane do układu pomiarowego, zaś napięcia niepożądane wywoływane np. przez przebiegające stopniowo zmiany przewodności w poważnym stopniu zostają stłumione.

SELEKCJA CZĘSTOTLIWOŚCIOWA

Dodatkowe polepszenie stosunku sygnału od pęknięć do zakłóceń może być uzyskane przy zastosowaniu układu selektywnego częstotliwościowego. Stosownie do każdorazowo istniejącej stałej prędkości kontrolnej — napięcie od wad i napięcia zakłócające mają różne częstotliwości. Przy odpowiednim dostrojeniu układu selektywnego częstotliwościowo przepuszcza on jedynie tylko napięcia pochodzące od wad tj. powyżej pewnej określonej częstotliwości. Napięcia zakłócające, które charakteryzują się innym zakresem częstotliwości zostają w poważnym stopniu stłumione.

GŁĘBOKOŚĆ WNIKANIA

Na wykrywalność nieciągłości materiału duży wpływ ma głębokość wnikania prądów wirowych. Głębokość wnikania wyznacza się wg wzoru (3) [9, 26]; zależy ona od częstotliwości pracy, jak również i przenikalności magnetycznej μ_r i przewodności γ , a więc także i temperatury badanego obiektu.

Przez dobór częstotliwości prądu wzbudzającego można dowolnie wybrać głębokość wnikania pola elektromagnetycznego, tj. głębokość objętego pomiarem obszaru powierzchniowego elementu badanego. Z tego faktu wynika, że metoda prądów wirowych nadaje się zwłaszcza do wykrywania wad na powierzchni materiału, lub tuż pod nią. Do wykrywania usterek w rdzeniu obiektu metoda ta nie nadaje się.

SZYBKOŚĆ KONTROLI

Zasadniczo można stwierdzić, że metoda prądów wirowych umożliwia uzyskanie największych szybkości kontroli w stosunku do innych nieniszczących metod kontroli. Szybkość kontroli dająca się uzyskiwać w każdym przypadku określona jest po pierwsze przez szerokość pasma przenoszenia układu elektronicznego opracowywania sygnałów. Aby uniknąć zmniejszenia czułości wskazań, określone szybkości kontroli nie mogą być przekroczone przy rozmaitych, uwarunkowanych przez obiekt badany częstotliwościach pracy. W razie przekroczenia maksymalnej szybkości kontroli występuje spadek czułości.

Szybkość kontroli może być dalej ograniczona przez jakość mechanicznych urządzeń prowadzących. Tak np. drobne niedokładności w przenośniku rolkowym (niewyważenia itp.), które przy nieznacznych szybkościach kontroli nie mają ewidentnego wpływu na wynik badania, zakłócają w sposób istotny badanie przy wyższych szybkościach kontroli. Oznacza to, że możliwa dzięki układom elektronicznym maksymalna szybkość kontroli może być osiągana tylko wtedy, kiedy mechaniczne prowadzenie badanych obiektów jest dostatecznie dokładne.

Ograniczenie szybkości kontroli od dołu teoretycznie nie istnieje. Badania można również prowadzić i w pracy statycznej (szybkość kontroli równa zero). W praktyce ten rodzaj pracy stosuje się tylko dla wzorcowania układów elektronicznych opracowywanych sygnałów. Praktycznie jednak szybkość kontroli zostaje ograniczona od dołu przez konieczność dopasowywania szybkości. To dopasowanie szybkości służy do tego, aby jeszcze bardziej poprawić stosunek sygnału do zakłóceń. Dopasowania można dokonywać przy szybkościach kontroli maksymalnie 50 m/s — aż do 0,008 m/s.

BIBLIOGRAFIA

1. M. Cousin: *An On-Line Eddy Currents System for the Quality Evaluation of Hot Rods*. Materials Evaluation, 1985, Nr 12, str. 1649—1654
2. K. Dydyński, M. Dźwiarek, A. Lewińska-Romicka, S. Piskorski: *Defektoskop do badania elementów niemagnetycznych*. Patent PRL P-145699, 1984-06-18
3. Ju. K. Fiedosienko: *Mietod pribliżennowo razziota mnogoslojnych wichirotokowych preobrazowatelej*. Diefektoskopija, 1980, Nr 5, str. 5—16
4. F. Förster: *Analyse der zerstörenden Werkstoffprüfung mit zerstörungsfreien elektromagnetischen Verfahren — Verfolgung des Risswachstums beim Dauerschwingversuch*. Materialprüfung, 1970, Nr 5, str. 149—156
5. F. Förster: *Analyse der zerstörenden Werkstoffprüfung mit zerstörungsfreien elektromagnetischen Verfahren*. Materialprüfung 1977, Nr 7, str. 237—245, Nr 8, str. 290—295, Nr 11, str. 470—473
6. W. G. Gerasimow: *Czuwstwielnost' prachodnych daczikow pri kontrole paramietrow mnogoslojnych cylindriczeskich izdelij*. Defektoskopija, 1969, Nr 3, str. 82—92
7. W. G. Gerasimow, J. J. Ostanin, A. D. Pokrovskij, V. V. Suchorukow, L. A. Cernow: *Nierazruszajuscij kontrol kaczestwa izdelij elektromagnitnymi mietodami*. Leningrad: Energija, 1978
8. W. G. Gerasimow, E. D. Sołomiencow, W. W. Suchorukow: *Pribliżenij mietod razziota wieliczin signalow wichirotokowych diefektoskopow z prachodnymi daczikami*. Diefektoskopija, 1967, Nr 4, str. 42—46
9. H. Heptner, H. Stroppe: *Magnetyczne i indukcyjne badania metali*. Katowice: Wyd. Śląsk, 1972
10. N. G. Leszczenko, B. B. Binokurov, W. G. Zołotuchin: *Dwuchczastotnyj diefektoskop dla kontrola stalnoj pruzinnoj prawoloki*. Diefektoskopija 1977, Nr 2, str. 50—54
11. A. Lewińska-Romicka: *Problemy nieniszczących badań rur niemagnetycznych metodą prądów wirowych*. Rozprawy Elektrotechniczne 1986, Nr 3, str. 801—812
12. A. Lewińska-Romicka: *Przetworniki wiropądowe do nieniszczących badań rur nieferromagnetycznych*. Rozprawy Elektrotechniczne 1988, nr 3, str. 839—873
13. F. D. Mundt: *Eddy Current Measurements with a Coil Encircling a Two — Conductor Rod*. Materials Evaluation, 1972, Nr 9, str. 189—194
14. J. Prestwood, K. Sumerling: *NDT for Irradiated Reactor Fuel Pins by Eddy Currents and Gamma Scanning*. Non-destructive Testing, 1975, Nr 4, str. 90—93

15. N. P. Rtiszczewa, W. W. Suchorukow: *Czislennyj analiz režimow raboty wicher-tokowo priobrazowatiela prachodnowo tipa pri kontrole fierromagnitnykh prutkow*. Defiektoskopija, 1979, Nr 12, str. 57—65
16. W. A. Sandowski: *Formirowanije elektromagnitnowo pola w kontrolirujemom obrazce*. Defiektoskopija, 1984, Nr 2, str. 46—49
17. W. Solnica, A. Lewińska-Romicka: *Niektóre problemy nieniszczących badań jakości prętów ze stali chromowej H26 w oparciu o metodę prądów wirowych*. Rozprawy Elektrotechniczne 1980, nr 4, str. 875—889
18. W. Solnica, A. Lewińska-Romicka: *System nieniszczących badań i sposoby rejestracji wyników kontroli drutów niemagnetycznych*. Rozprawy Elektrotechniczne, 1977, nr 3, str. 535—552
19. W. Solnica, A. Lewińska-Romicka, K. Dydyński, M. Owsiany: *Przetwornik pomiarowy do badania elementów ferromagnetycznych*. Patent PRL Nr 113703, 1978-10-13
20. W. Solnica, A. Lewińska-Romicka, K. Dydyński, S. Piskorski, M. Owsiany: *Defektoskop do badania elementów ferromagnetycznych*. Patent PRL Nr 113172, 1987-10-13
21. W. Solnica, A. Lewińska-Romicka, M. Owsiany: *Układ cyfrowej rejestracji wad do defektoskopów*. Patent PRL Nr 103326, 1975-11-13
22. W. Solnica, A. Lewińska-Romicka, M. Owsiany: *Układ do wykrywania wad drutów z materiałów przewodzących prąd elektryczny*. Patent PRL Nr 118366, 1977-07-30
23. W. Solnica, A. Lewińska-Romicka, S. Piskorski: *Sposób wzorcowania detektora fazy w defektoskopach wiroprądowych*. Patent PRL Nr 124916, 1980-05-28
24. W. Solnica, A. Lewińska-Romicka, S. Piskorski: *Urządzenie do nieniszczących badań jakości i znakowania wad materiałowych w drutach niemagnetycznych*. Rozprawy Elektrotechniczne, 1980, nr 3, str. 609—615
25. D. Waidelich: *The Role of Theory in Eddy Current Testing*. Materials Evaluation, 1970, Nr 12, str. 262—266
26. L. C. Wilcox, Jr: *Prerequisites for Quantitative Eddy Current Testing*. Materials Evaluation, 1967, Nr 10, str. 237—244

A. LEWIŃSKA-ROMICKA, S. PISKORSKI

PROPERTIES OF EDDY-CURRENT TRANSDUCERS FOR NON-DESTRUCTIVE TESTING OF FERROMAGNETIC RODS

Summary

The procedure and results of tests of eddy-current through-coil transformer transducers, developed for the detection of cracks in rods made of a ferromagnetic material are described. The method of determination of the service frequency of the transducers is presented.

A. LEWIŃSKA-ROMICKA, S. PISKORSKI

PROPRIÉTÉS DES CONVERTISSEURS CENTRIFUGES POUR LES ESSAIS NON-DESTRUCTIFS DES BARRES FERROMAGNÉTIQUES

Résumé

On a décrit la méthodologie et les résultats des essais de convertisseurs centrifuges, élaborés pour détecter des fissures de barres, faites avec des matériaux ferromagnétiques. On a montré le mode de détermination de la fréquence du travail des convertisseurs.

A. LEWIŃSKA-ROMICKA, S. PISKORSKI

EIGENSCHAFTEN DER WIRBELSTROMWANDLER FÜR ZERSTÖRUNGSFREIE
PRÜFUNGEN FERROMAGNETISCHER STÄBE

Zusammenfassung

Es wurden eine Prüfmethode für Wirbelstromwandler sowie diesbezügliche Versuchsergebnisse beschrieben, die für das Auffinden von Rissen bei Stäben aus ferromagnetischen Werkstoffen erarbeitet worden sind. Eine Bestimmungsmethode für die Arbeitsfrequenz der Wandler wurde dargestellt.

А. ЛЕВИНЬСКА-РОМИЦКА, С. ПИСКОРСКИ

СВОЙСТВА ВИХРЕТОКОВЫХ ПЕРЕХОДНЫХ ТРАНСФОРМАТОРНЫХ
ПРЕОБРАЗОВАТЕЛЕЙ ДЛЯ НЕРАЗРУШАЮЩИХ ИСПЫТАНИЙ ФЕРРОМАГНИТНЫХ
СТЕРЖНЕЙ

Резюме

Описана методика и результаты испытаний вихретоковых преобразователей разработанных с целью обнаруживания течин в стержнях изготавливаемых из ферромагнитных материалов. Представлен способ определения частоты работы преобразователей.

Generacja impulsów wysokonapięciowych w zestyku ślizgowym maszyn elektrycznych. Część III

ZBIGNIEW MUSZAŁSKI, LEON GOLUSIŃSKI, WIESŁAWA RAFAŁSKA,
WOJCIECH RAFAŁSKI

*Instytut Elektrotechniki Morskiej i Przemysłowej, Politechnika
Gdańska*

Otrzymano 1988.08.10

Autoryzowano do druku 1988.11.09

Podany uprzednio w części II rozwinięty schemat obwodowy zestyku ślizgowego, sprowadzono do równoważnego potencjalnie dwójnika w taki sposób, aby ujawnić oddziaływanie parametrów szczotki na podatność zestyku ślizgowego do pracy iskrowej.

Rozwiązanie układu równań wynikającego z omawianego dwójnika daje równanie ogólne zestyku ślizgowego. Analiza tego równania wskazuje, które elementy zestyku i w jaki sposób wpływają na jakość pracy zestyku.

1. UPROSZCZONY SCHEMAT ZASTĘPCZY ZESTYKU ŚLIZGOWEGO

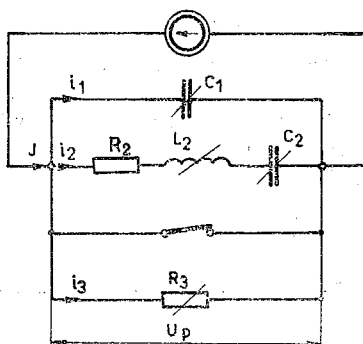
Doświadczenia eksploatacyjne sugerują, że w licznych przypadkach zmniejszenie intensywności iskrzenia zestyku ślizgowego maszyn elektrycznych można uzyskać przez — zmianę marki szczotki. Jeżeli przy tym katalogowe parametry zastosowanej „lepszej” szczotki nie odbiegają od analogicznych parametrów szczotki „gorszej” to znaczy, że poprawa jakości przewodzenia prądu może być uzyskana wskutek zmiany mechanicznych lub obwodowych parametrów szczotki [4]. Parametry mechaniczne jak sprężystość, częstotliwość drgań własnych oraz logarytmiczny dekrement tłumienia drgań szczotek nie będą tu rozpatrywane.

Na pracę iskrzeniową zestyku ślizgowego istotny wpływ wywiera przestrzeń międzystykowa, a ściślej mówiąc — jej parametry obwodowe. Analiza oddziaływania parametrów obwodowych szczotki w oparciu o rozwinięty schemat zastępczy zestyku ślizgowego ([4], rys. 6) jest utrudniona i stosunkowo mało przejrzysta. Skłania to do wprowadzenia uproszczeń i przedstawienia schematu zastępczego w takiej postaci, która uwidoczni wpływ parametrów szczotki na zjawiska zachodzące w przestrzeni międzystykowej.

Wprowadzeniu uproszczeń sprzyjają następujące przesłanki. Po utracie styczności szczotki z kolektorem, niektóre elementy rozwiniętego schematu zastępczego nie odgrywają żadnej roli. Tak więc „przelącznik” P_{ω_s} odłącza gałąź zawierającą powierzchniowe źródła

sem E_p oraz rezystancję przewężenia w punktach styczności R_0 . W czasie pracy iskrzeniowej zestyku występują składowe prądu przewodzonego o stosunkowo dużej częstotliwości, zatem „częstotliwościowo-elektronowy wyłącznik” W_f odłącza pojemność C_{s1} przy znacznie mniejszych częstotliwościach. Strefa IV jest strefą „politury” i choć prawdopodobnie ma pewien wpływ na rozwój początkowego wyładowania elementarnego, to jednak w rozpatrywanym zagadnieniu z całą pewnością nie odgrywa roli dominującej. Ponadto rozwój początkowego wyładowania elementarnego nie jest jeszcze zjawiskiem dostatecznie poznany [1]. Dominującej roli nie odgrywa również strefa I, to jest strefa linki szczotkowej oraz indukcyjność przestrzeni międzystykowej L_p w strefie III [4]. Zatem wymienione ogniwa i elementy można pominąć w dalszej analizie — bez obawy wprowadzenia poważniejszych błędów.

Tak uproszczony schemat zastępczy składający się z szeregowego połączenia dwóch dwójników tj. strefy II i III wg [4] może być sprowadzony do dwójnika potencjalnie równoważnego. Część dwójnika obrazującą wpływ niektórych parametrów obwodowych szczotki na charakter przebiegu spadku napięcia na impedancji przestrzeni międzystykowej podano



Rys. 1. Schemat obwodowy zestyku ślizgowego sprowadzony do potencjalnie równoważnego dwójnika

na rys. 1, przy czym parametry zastępcze RLC tego dwójnika wynikają z przekształcenia dwójników II i III.

2. MODEL MATEMATYCZNY ZESTYKU ŚLIZGOWEGO MASZYN ELEKTRYCZNEJ

Przebieg spadku napięcia na impedancji przestrzeni międzystykowej jest określony równaniem

$$u_p(\tau) = i_3(\tau)R_3 \quad (1)$$

Przebieg prądu $i_3(\tau)$ można określić z układu równań (2) wynikającego ze schematu obwodowego przedstawionego na rys. 1.

$$\left. \begin{aligned} I &= i_1 + i_2 + i_3 \\ i_1 &= C_1 \frac{du_{C1}}{dt} \\ i_2 &= C_2 \frac{du_{C2}}{dt} \\ i_3 R_3 &= i_2 R_2 + L_2 \frac{di_2}{dt} + u_{C2} \end{aligned} \right\} \quad (2)$$

Warunkami początkowymi są:

$$\begin{aligned} i_1|_{\tau=0} &= I & u_{C1}|_{\tau=0} &= 0 \\ i_2|_{\tau=0} &= 0 & u_{C2}|_{\tau=0} &= 0 \\ i_3|_{\tau=0} &= 0. \end{aligned}$$

Stosując przekształcenie Laplace'a w rozwiązaniu uzyskuje się równanie prądu

$$i_3(s) = \frac{I}{s} \frac{s^2 C_2 L_2 + s C_2 R_2 + 1}{s^3 C_1 C_2 L_2 R_3 + s^2 (C_2 L_2 + C_1 C_2 R_2 R_3) + s (C_2 R_2 + C_2 R_3 + s^3 C_1 R_3) + 1} \quad (3)$$

Oznaczając mianownik równania (3) przez m uzyskuje się

$$i_3(s) = I \left(\frac{s C_2 L_2}{m} + \frac{C_2 R_2}{m} + \frac{1}{sm} \right). \quad (4)$$

W trakcie stopniowej utraty styczności, charakter przebiegu prądu przepływającego przez przestrzeń międzystykową, może ulegać zmianie. W przypadku założenia rezonansowego przebiegu prądu mianownik równania (4) przyjmuje postać

$$m = C_1 C_2 L_2 R_3 (s - a_i) [(s - b_i)^2 + \omega_i^2] \quad (5)$$

przy czym

$$\left. \begin{aligned} a_i &= (U+V) - \frac{C_2 L_2 + C_1 C_2 R_2 R_3}{3 C_1 C_2 L_2 R_3} & b_i &= -\frac{1}{2} (U+V) - \frac{C_2 L_2 + C_1 C_2 R_2 R_3}{3 C_1 C_2 L_2 R_3} \\ \omega_i &= \frac{\sqrt{3}}{2} (U-V), \\ U &= \sqrt[3]{-q + \sqrt{q^2 + p^3}}, & V &= \sqrt[3]{-q - \sqrt{q^2 + p^3}} \end{aligned} \right\} \quad (6)$$

Przy powyższych, założeniach i spełnionym warunku $q^2 + p^3 > 0$

$$\left. \begin{aligned} p &= \frac{C_2 R_2 + C_2 R_3 + C_1 R_3}{3 C_1 C_2 L_2 R_3} - \frac{(C_2 L_2 + C_1 C_2 R_2 R_3)^2}{(3 C_1 C_2 L_2 R_3)^2}, \\ q &= \frac{(C_2 L_2 + C_1 C_2 R_2 R_3)^3}{(3 C_1 C_2 L_2 R_3)^3} - \frac{(C_2 L_2 + C_1 C_2 R_2 R_3)(C_2 R_2 + C_2 R_3 + C_1 R_3)}{6 (C_1 C_2 L_2 R_3)^2} + \\ &+ \frac{1}{2 C_1 C_2 L_2 R_3}. \end{aligned} \right\} \quad (7)$$

Ostatecznie transformata odwrotna dla równania (4) przyjmuje postać

$$i_3(\tau) = \frac{I}{C_1 C_2 L_2 R_3} [e^{a_i \tau} k_1 + e^{b_i \tau} k_2 \sin(\omega_i \tau + \varphi) - k_3]. \quad (8)$$

Jak wynika z oscylogramu [3 (rys. 4)] zanik prądu można interpolować wyrażeniem

$$I = \frac{I_m}{1 + \exp \frac{\tau - \tau_{11}}{\tau_0 \cdot d}} \quad (9)$$

gdzie τ [s] — czas liczony od początkowej chwili odrywania się szczotki

τ_{11} [s] — czas, po którym prąd osiągnie wartość $I = 0,5 I_m$

τ_0 [s] — czas, po którym długość przestrzeni międzystykowej osiągnie wartość $s = 4 \cdot 10^{-6}$ m

d — współczynnik zależny od środowiska, dla zestyku elektrografit-miedź w powietrzu $d = 7,66 \cdot 10^{-3}$; w tych warunkach $\tau_{11} = \tau_0 + 0,5 \Delta \tau$, gdzie

$$\Delta \tau = 5 \cdot 10^{-9} I_m [\text{s}];$$

I_m [A] — początkowa ustalona wartość prądu.

Parametry uproszczonego schematu obwodowego (rys. 1) przy uwzględnianiu parametrów obwodowych szczotki i przestrzeni międzystykowej przybierają następującą postać:

$$\begin{aligned} C_1 &= \frac{(L_{s1} + L_{s2})^2 C_s C_p}{(L_{s1} + L_{s2})^2 C_s + L_{s1}^2 C_p} \\ C_2 &= \frac{L_{s1}^2 C_p^2}{(L_{s1} + L_{s2})^2 C_s + L_{s1}^2 C_p} \\ L_2 &= \frac{[(L_{s1} + L_{s2})^2 C_s + L_{s1}^2 C_p]^2}{(L_{s1} + L_{s2}) L_{s1}^2 C_p^2} \end{aligned}$$

Wyznaczona w [4] indukcyjność $L_s(f) = 10^{\text{Dtg } \frac{\omega_l}{2\pi} + E}$ może być przedstawiona w postaci tożsamości

$$L_s = \frac{\zeta \xi L_s^2}{\zeta L_s + \xi L_s}$$

stąd $L_{s1} = \zeta L_s$; $L_{s2} = \xi L_s$ przy czym $\zeta = \frac{\xi}{\xi - 1}$. Przy założeniu, że $\zeta = \xi = 2$ indukcyjności szczotki $L_{s1} = L_{s2} = 2L_s$ wówczas parametry uproszczonego schematu obwodowego znacznie się uproszczą przyjmując postać

$$\left. \begin{aligned} C_1 &= \frac{4C_s C_p}{4C_s + C_p}; & C_2 &= \frac{C_p^2}{4C_s + C_p}; \\ L_2 &= \frac{(4C_s + C_p)^2}{C_p^2} L_s, \end{aligned} \right\} \quad (10)$$

stąd też iloczyn $C_1 C_2 L_2 R_3 = 4C_s C_p L_s R_p$. Ponadto $R_3 = R_p(\tau)$ przy czym parametry przestrzeni międzystykowej t.j. rezystancję $R_p(\tau)$ i pojemność elektryczną $C_p(\tau)$ oraz parametry szczotki t.j. rezystancję R_s , jej pojemność elektryczną $C_s(f)$ i indukcyjność $L_s(f)$,

omówiono w [4] $f = \frac{\omega_l}{2\pi}$, zaś pulsacja ω_l — z równania (6).

Z podstawienia wyrażeń (10) do równania (7) oraz pomijając składniki o kilka, i więcej rzędów mniejsze od największego, otrzymuje się odpowiednie współczynniki

$$\left. \begin{aligned} p &= \frac{C_p}{12(4C_s + C_p)C_sL_s} + \frac{1}{3(4C_s + C_p)L_s} - \frac{(4C_s + C_p)^2}{144C_s^2C_p^2R_p^2}, \\ q &= \frac{(4C_s + C_p)^3}{1728C_s^3C_p^3R_p^3} - \frac{1}{96C_s^2L_sR_p} + \frac{1}{12C_sC_pL_sR_p}. \end{aligned} \right\} \quad (11)$$

Po uwzględnieniu wyrażeń (10) występujące w równaniu (8) współczynniki k przyjmują odpowiednio postać:

$$\left. \begin{aligned} k_1 &= \frac{a_i^2(4C_s + C_p)^2L_s + a_iC_p^2R_s + (4C_s + C_p)}{a_i[(b_i - a_i)^2 + \omega_i^2](4C_s + C_p)}; \\ k_2 &= \sqrt{k_4^2 + k_5^2}; \\ k_3 &= \frac{1}{a_i(b_i^2 + \omega_i^2)}; \\ k_4 &= \frac{(b_i - a_i)[a_iC_p^2R_s + a_ib_i(4C_s + C_p)^2L_s + (4C_s + C_p)]}{a_i\omega_i[(b_i - a_i)^2 + \omega_i^2](4C_s + C_p)} + \\ &\quad + \frac{\omega_i(4C_s + C_p)L_s}{(b_i - a_i)^2 + \omega_i^2} - \frac{b_i}{a_i\omega_i(b_i^2 + \omega_i^2)}; \\ k_5 &= \frac{(b_i - a_i)(4C_s + C_p)L_s}{(b_i - a_i)^2 + \omega_i^2} - \frac{a_iC_p^2R_s + a_ib_i(4C_s + C_p)^2L_s + (4C_s + C_p)}{a_i[(b_i - a_i)^2 + \omega_i^2](4C_s + C_p)} + \\ &\quad + \frac{1}{a_i(b_i^2 + \omega_i^2)}; \\ \varphi &= \arctg \frac{k_5}{k_4}. \end{aligned} \right\} \quad (12)$$

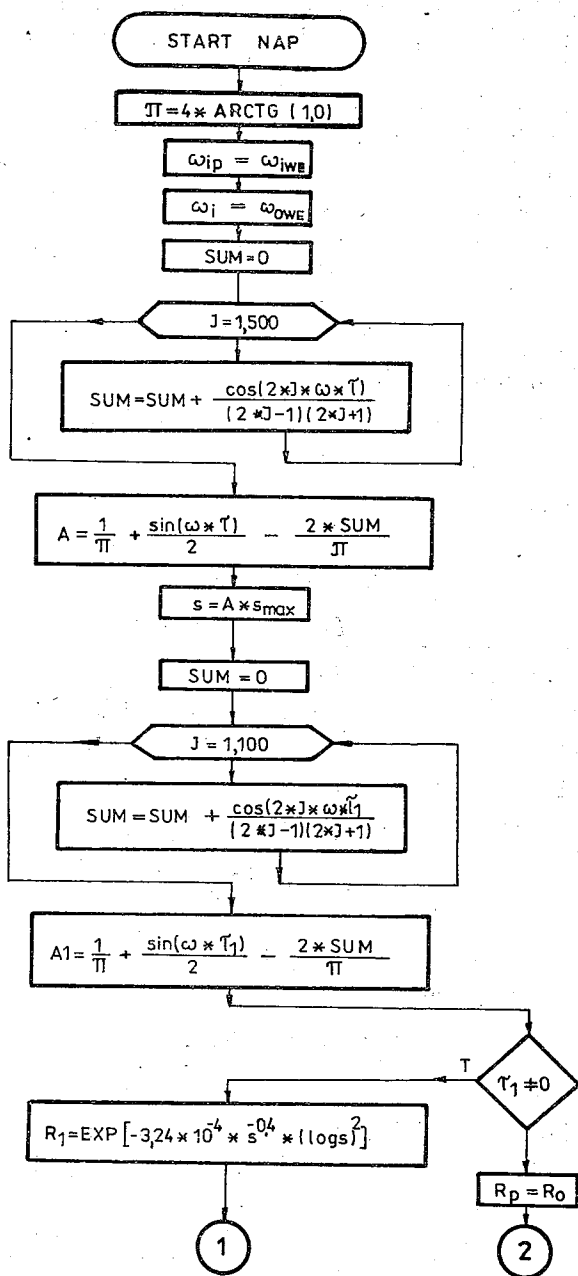
Tak więc ostatecznie przebieg spadku napięcia na przestrzeni międzystykowej zestyku ślizgowego jest wyrażony równaniem (13)

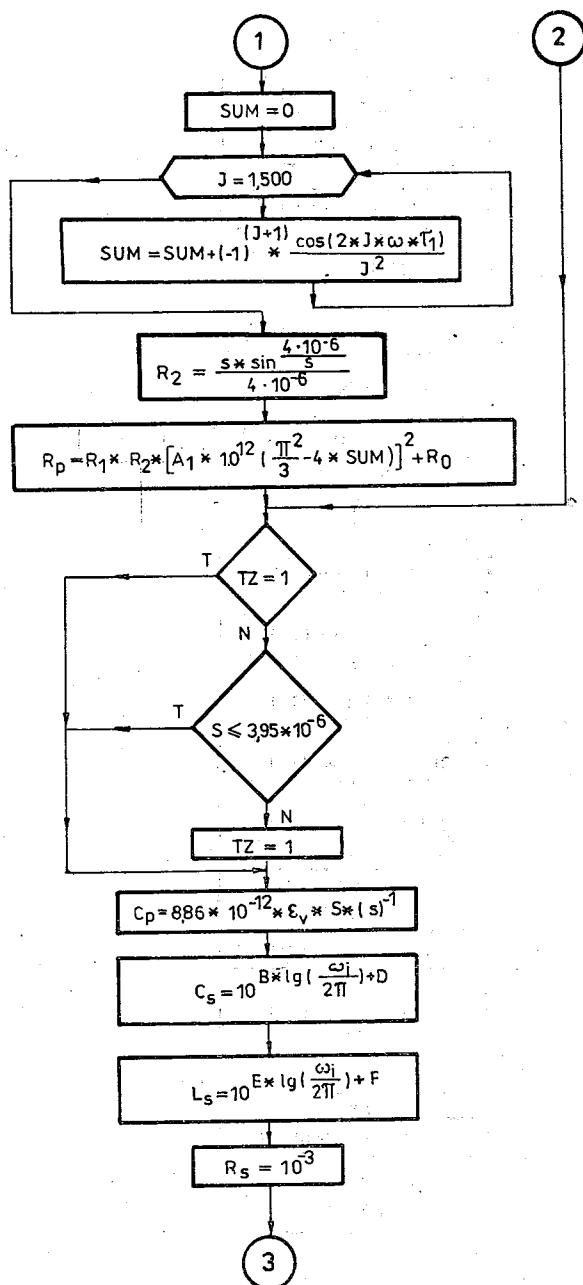
$$u_p(\tau) = i_3(\tau)R_p(\tau) \quad (13)$$

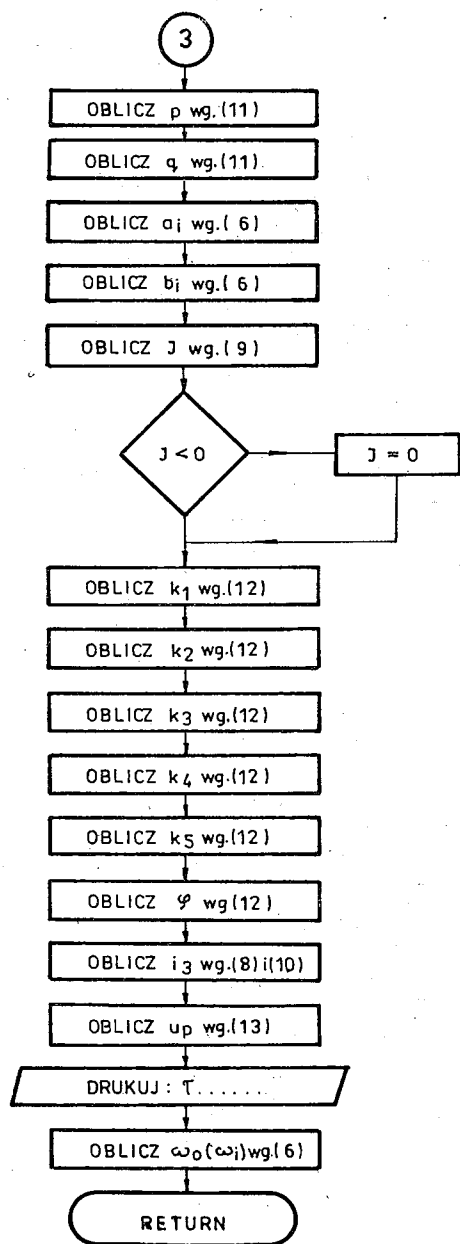
Omówione w [3] trzy stany pracy zestyku ślizgowego wymagają wprowadzenia pewnych dodatkowych warunków uwzględniających pewne zjawiska fizyczne.

W I stanie pracy wartość rezystancji przejścia może ulegać pewnym niewielkim zmianom wskutek zmiany liczby punktów styczności. Znacznie bardziej istotnym jest jednak II stan pracy, gdy rezystancja wzrasta wskutek utraty styczności i skokowo maleje dzięki zjawisku „*frittingu*”. W III stanie pracy — spadek napięcia wzrasta do chwili ponownego zapłonu łuku, lub do czasu rozładowania obwodu.

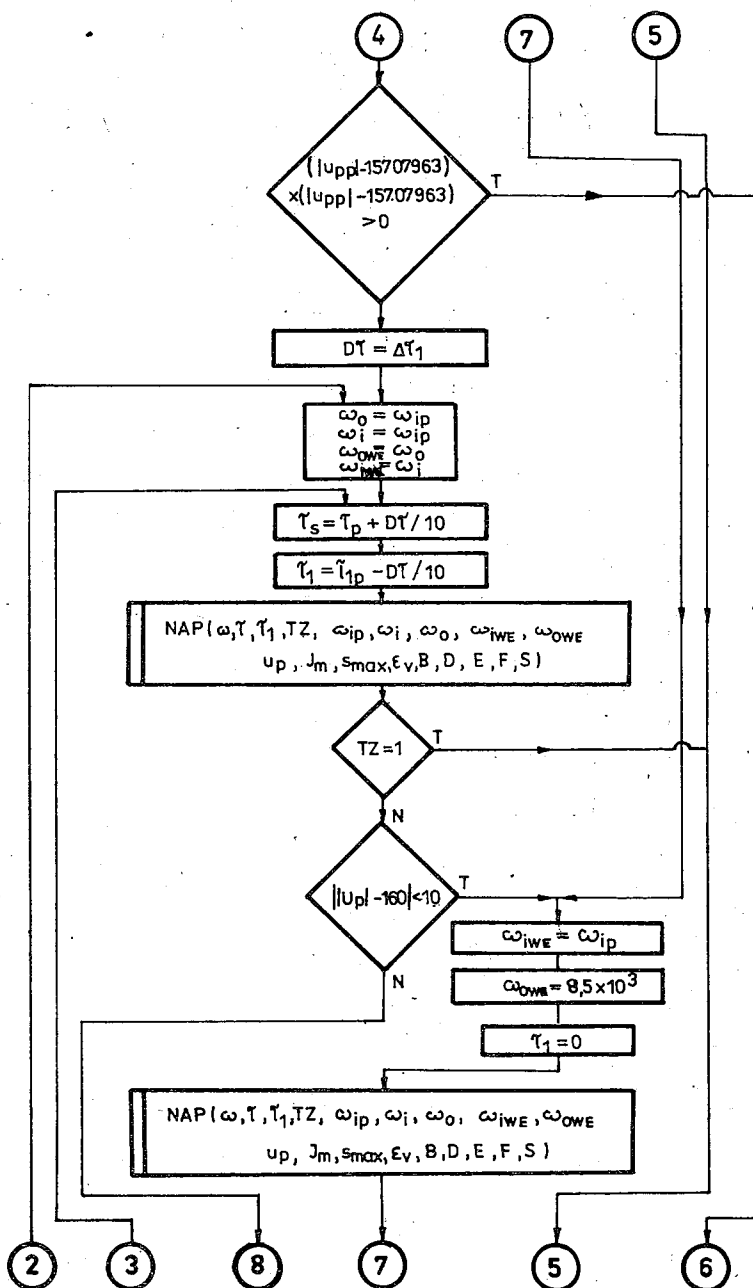
Zmiany wartości spadku napięcia wskutek zmiany zjawiska w przestrzeni międzystykowej, autorzy częściowo uwzględnili w algorytmie przedstawionym na rys. 2 i rys. 3 przez człony $R1$ i $R2$. Zwiększenie dokładności tych przebiegów, jak również uwzględnienie warunków prowadzących do ponownego zapłonu wymaga dalszych badań. Na rys. 4 przedstawiono przebieg spadku napięcia uzyskany z obliczeń na maszynie cyfrowej.



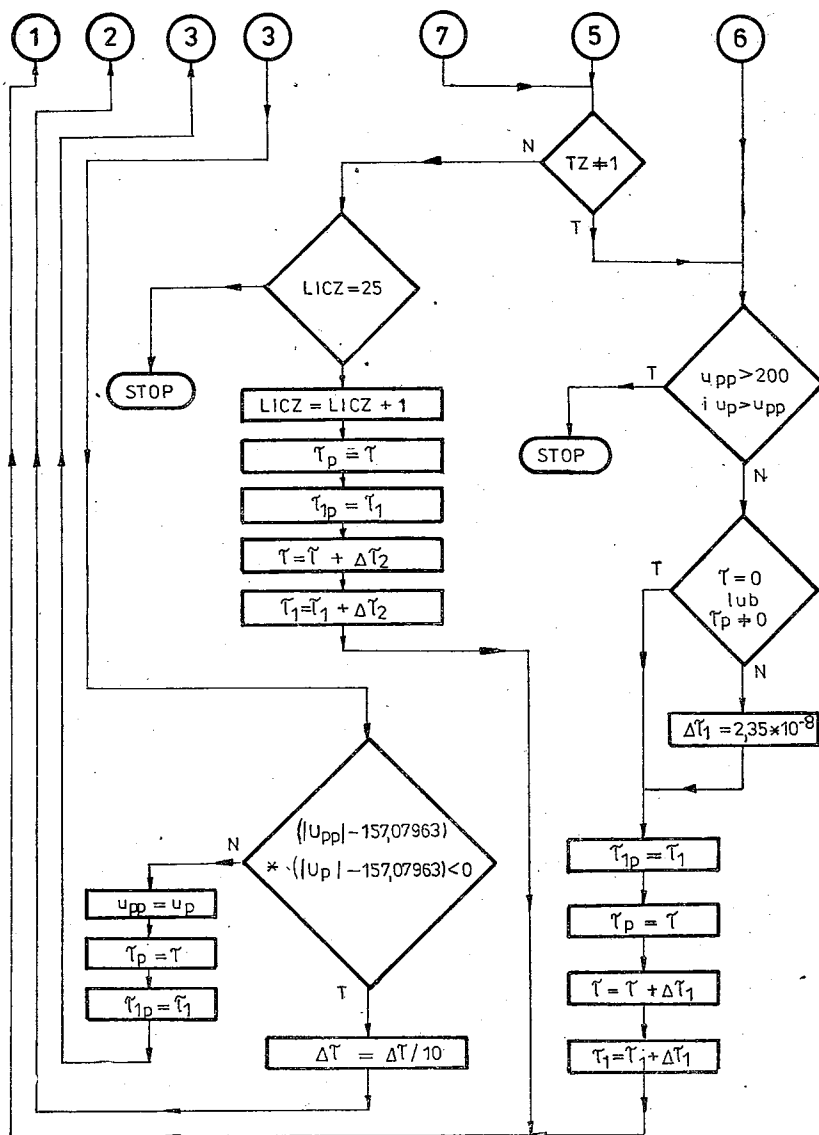




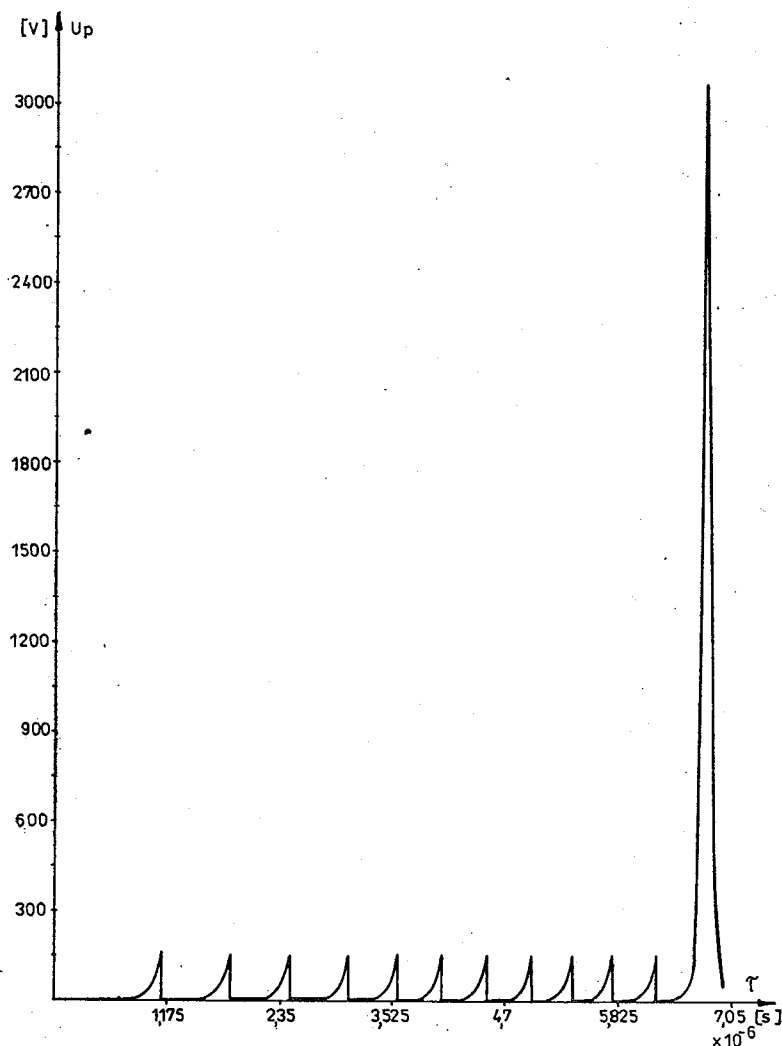




Rys. 2. Alogrytm wyznaczania przebiegu napięcia w przestrzeni międzystykowej zestawu ślizgowego — program NAP. Obliczenia wykonano dla następujących danych: $\omega = 8,5 \cdot 10^3$; $s_{max} = 10^{-4}$; $R_0 = 0,25$; $\varepsilon_v = 18$; $S = 2 \cdot 10^{-4}$; $B = -1,53$; $D = 4,22$; $E = -0,0823$; $F = -7,83$; $I_m = 30$ A



Rys. 3. Algorytm wyznaczania przebiegu napięcia w przestrzeni międzystykowej zestyku ślizgowego — program sterujący

Rys. 4. Spadek napięcia na przestrzeni międzystykowej $u_p(\tau)$

3. WNIOSKI

W [4] wykazano, że przy zmianie pulsacji składowej okresowej prądu przepływającego przez szczotkę — rząd wartości indukcyjności pozostaje stały, natomiast wartość pojemności szczotki może ulegać zmianom w granicach kilkunastu rzędów. Analizując więc wpływ parametrów szczotki na przebieg napięcia w przestrzeni międzystykowej zestyku ślizgowego można stwierdzić, że w początkowym stadium kiedy to pojemność szczotki jest rzędu $10^2 F$ czyli

$$C_s = 100F \gg C_p,$$

$$C_1 = C_p, \quad C_2 = \frac{C_p^2}{4C_s} = 2,5 \cdot 10^{-3} \cdot C_p^2,$$

$$L_2 = \frac{16C_s^2}{C_p^2} L_s = \frac{1,6 \cdot 10^5}{C_p^2} L_s.$$

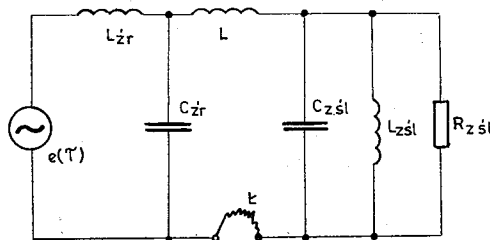
Natomiast w fazie końcowej, gdy rząd obu pojemności może być zbliżony, to znaczy gdy $C_s \cong C_p$ parametry te mogą przyjąć wartość

$$C_1 = 0,8C_p, \quad C_2 = 0,25C_p, \quad L_2 = 25L_s.$$

Porównanie powyższych zależności wykazuje, że parametry obwodowe szczotki w zestyku ślizgowym wywierają niebagatelny wpływ na charakter przebiegu napięcia — a tym samym i na podatność zestyku do pracy iskrzeniowej. Wpływ ten wynika głównie stąd, że wskutek zmiany pojemności szczotki, układ z dużą łatwością osiąga pulsacje rezonansowe.

Proponowany model zestyku ślizgowego wyjaśnia również, w jaki sposób na pracę zestyku ślizgowego oddziałują różne czynniki. Tak na przykład wartość C_p zmienia się nie tylko w zależności od długości przestrzeni międzystykowej, ale również w zależności od przenikalności elektrycznej składników środowiska ϵ_v . Zmiany te mogą przebiegać w przedziale dwóch rzędów. Zmiana charakteru atmosfery gazowej wpływa na zmianę przebiegu charakterystyki rezystancji przestrzeni międzystykowej $R_p(\tau)$, może zatem wpływać na zwiększenie lub zmniejszenie podatności zestyku do pracy iskrzeniowej.

W teorii procesów łączeniowych znane jest zjawisko ucinania prądów indukcyjnych. Podczas wyłączania prądu rzędu kilku do kilkunastu amperów, w układzie przedstawionym na rys. 5, może zaistnieć przerwanie prądu przed jego przejściem przez wartość zerową.



Rys. 5. Układ, który może doprowadzić do ucięcia prądu wyłączanego. L_{zr} , C_{zr} — indukcyjność i pojemność źródła; $R_{z.sl.}$, $L_{z.sl.}$, $C_{z.sl.}$ — rezystancja, indukcyjność i pojemność zestyku ślizgowego; L — indukcyjność połączeń dodatkowych; k — łuk elektryczny

Zjawisko to można w uproszczeniu wyjaśnić generowaniem się prądu w obwodzie zamkniętym; spadek napięcia na łuku L , C_{zr} , $C_{z.sl.}$ o częstotliwości

$$f_w = \left(2\pi \sqrt{\frac{LC_{zr}C_{z.sl.}}{C_{zr} + C_{z.sl.}}} \right)^{-1}. \quad (14)$$

Prąd wielkiej częstotliwości, pobudzony nagłymi zmianami spadku napięcia na łuku, może zrównoważyć prąd wyłączany — powodując jego wcześniejsze doprowadzenie do wartości zerowej, a tym samym i wyłączenie.

Układ przedstawiony na rys. 5 sugeruje daleko idącą analogię połączenia wspomnianego wyżej schematu łącznika z uproszczonym schematem zastępczym zestyku ślizgowego wg [2] w połączeniu z wirnikiem maszyny komutatorowej. W tym przypadku wielkości: $e(\tau)$, $L_{\Sigma r}$ i $C_{\Sigma r}$ są, związanymi z wirnikiem i komutatorem, parametrami maszyny elektrycznej.

Z uwagi na technikę łączeniową ucinanie prądu jest zjawiskiem negatywnym ponieważ prowadzi do przepięć. Natomiast w maszynach elektrycznych zjawisko to można rozpatrywać w dwóch aspektach. Po pierwsze w szybkoobrotowych komutatorowych maszynach mocy ułamkowej, gdy szczotka może uzyskiwać sporadyczny, chwilowy kontakt z komutatorem — zjawisko gaszenia i ponownego zapłonu łuku, może spełniać rolę „sterowanego zestyku plazmowego”. Jest to co prawda objaw naganny, ale może zwiększać równomierność biegu napędu. Po drugie w większych maszynach skrócenie czasu palenia się łuku jest objawem pozytywnym ponieważ — sprzyja mniejszemu uszkodzeniu powierzchni ślizgowej komutatora. Ponadto znana jest zdolność maszyn komutatorowych do pooparzelinowej samoregeneracji powierzchni ślizgowej komutatora. Nie ma natomiast takiej zdolności maszyna pierścieniowa, której proces niszczenia powierzchni ślizgowej, po przypadkowym powstaniu oparzeliny — stale wzrasta. Zmusza to w końcu do wyłączenia maszyny z eksploatacji i poddania zabiegowi konserwacji.

Objawy te sugerują, że zdolność samoregeneracji zestyku ślizgowego maszyna zawdzięcza — międzystykowej pojemności komutatora $C_{\Sigma r}$ i łatwo dostrajającej się pojemności szczotki $C_{z.st.}$.

Rozwinięty schemat obwodowy zestyku ślizgowego [4] wyjaśnia również obserwowane w maszynach elektrycznych zjawisko iskrzenia np. jednej szczotki z kilku pracujących równolegle. Można wyjaśnić to tym, że gdy jedna szczotka traci styczność — impedancja przejścia dla składowej wysokoczęstotliwościowej prądu na drodze do szczotki zachowującej styczność, może być znacznie większa od impedancji przejścia przez jej własną przestrzeń międzystykową.

BIBLIOGRAFIA

1. B. Bolanowski, A. Holj: *Analiza rozwoju wyladowania elementarnego*. VI Sympozjum — Podstawy teorii wyladowań elektrycznych w gazach — wrzesień 1982, P. W.
2. Z. Muszalski: *Rola pojemności szczotki w strukturze schematu zastępczego zestyku ślizgowego maszyn elektrycznych*. Przegląd Elektrotechniczny 1976 nr 8—9
3. Z. Muszalski, *Generacja impulsów wysokonapięciowych w zestyku ślizgowym maszyn elektrycznych*. Cz. I. Rozprawy Elektrotechniczne, 1986, z. 2
4. Z. Muszalski, *Generacja impulsów wysokonapięciowych w zestyku ślizgowym maszyn elektrycznych*. Cz. II. Rozprawy Elektrotechniczne, 1986, z. 2

Z. MUSZALSKI, L. GOLUSIŃSKI, W. RAFALSKA, W. RAFALSKI
HIGH VOLTAGE IMPULSE GENERATION ON THE SLIDE CONTACT
OF ELECTRICAL MACHINES
PART III

Summary

A fully developed circular diagram of a slide contact, presented in the part II, is reduced to the potentially equivalent two-terminal network so as to display the influence of brush parameters on the susceptibility of the slide contact to sparking. The solution of the system of equations resulting from the discussed two-terminal network gives the general equation of the slide contact. The analysis of this equation indicates which elements of the contact and in which way do influence the quality of the contact operation.

Z. MUSZALSKI, L. GOLUSIŃSKI, W. RAFALSKA, W. RAFALSKI
PROCESSUS DU DÉVELOPPEMENT DES IMPULSIONS DE HAUTE TENSION
DANS LE CONTACT GLISSANT DES MACHINES ÉLECTRIQUES
TROISIÈME PARTIE

Résumé

La deuxième partie a présenté un schéma-circuit développé du contact glissant. Dans celle-ci, on le remplace par un dipôle équivalent pour mieux étudier l'influence des paramètres du balai sur la susceptibilité de la production des étincelles dans le contact glissant. Comme résultat du système des équations, spécifique à ce dipôle, on obtient une équation générale pour le contact glissant. L'analyse de cette équation indique quels sont les éléments du contact qui influencent sur la qualité de son travail et quelle est leur nature.

Z. MUSZALSKI, L. GOLUSIŃSKI, W. RAFALSKA, W. RAFALSKI
ERZEUGUNG VON HOCHSPANNUNGSIMPULSEN IM GLEITKONTAKT
ELEKTRISCHER MASCHINEN
TEIL III

Zusammenfassung

Teil II dieser Bearbeitung umfaßte ein entwickeltes Kreisschema eines Gleitkontaktes. Hier wurde es durch einen gleich-wertigen Zweipol in der Art ersetzt, damit die Rückwirkung der Parameter der Bürste auf die Nachgiebigkeit des Gleitkontaktes für die Funkenbildung nachgebildet wird. Die Lösung des auf diesem Zweipol basierenden Gleichungs-systems führt zu einer allgemeinen Gleichung für den Gleitkontakt. Aus der Analyse dieser Gleichung ergibt sich, welche Elemente und auf welche Weise die Güte des Gleitkontaktes beeinflussen.

З. МУШАЛЬСКИ, Л. ГОЛУСИНСЬКИ, В. РАФАЛЬСКА, В. РАФАЛЬСКИ
ГЕНЕРИРОВАНИЕ ВЫСОКОЧАСТОТНЫХ ИМПУЛЬСОВ В СКОЛЬЗЯЩЕМ
КОНТАКТЕ ЭЛЕКТРИЧЕСКИХ МАШИН
ЧАСТЬ 3

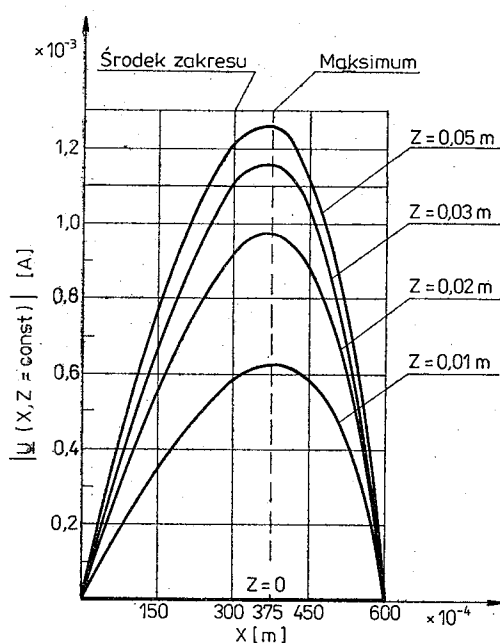
Резюме

Представленная во 2 части расширенная цепочная электрическая схема скользящего контакта приведена к потенциально эквивалентному двухполюснику таким образом, чтобы выявить влияние параметров щётки на склонность скользящего контакта к искрению. Решением системы уравнений вытекающей из рассматриваемого двухполюсника является общее уравнение скользящего контакта. Из анализа этого уравнения вытекают некоторые из элементов и каким образом влияют на качество работы этого контакта.

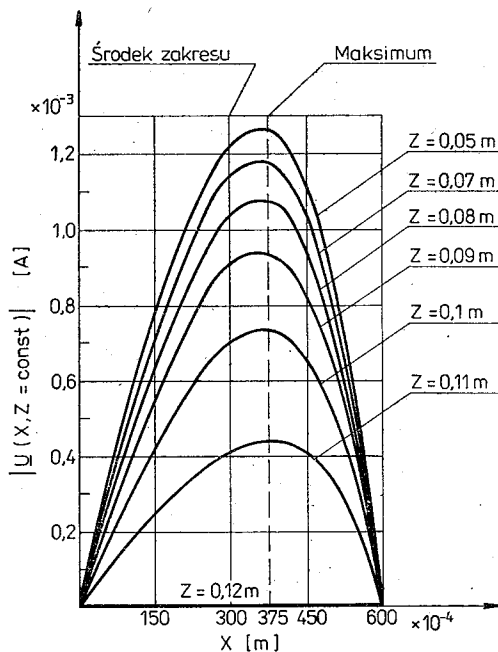
ERRATA DO ARTYKUŁU JERZEGO GOŁĘBIEWSKIEGO

Wykorzystanie funkcji Greena do analizy słabego zjawiska naskórkowości zachodzącego w cienkiej płycie prostokątnej pod wpływem poprzecznego pola o dowolnej niejednorodności

W zeszytzie 3-4/1987 na str. 662 wydrukowano błędne rysunki 3a i 3b. Poniżej zamieszczono właściwe wykresy:



Rys. 3a. Rozkład modułu funkcji prądu (funkcji strumienia) (31) dla $z \in \langle 0; 0,05 \rangle$ m



Rys. 3b. Rozkład modułu funkcji prądu (funkcji strumienia) (31) dla $z \in \langle 0,05; 0,12 \rangle$ m

Przepraszamy Autora i Czytelników

Redakcja

SPIS TREŚCI

Cz. R a j s k i: O pewnym zastosowaniu zasady minimum mocy	515
R. N a w r o w s k i: Rozkład pola elektrycznego i wytrzymałość elektryczna układu płyta-ostrze hiperboloidalne z uwzględnieniem ładunku przestrzennego	523
K. N o g a: Estymacja parametrów rozkładu Nakagami na podstawie danych pomiarowych sygnału	531
F. W y s o c k a - S c h i l l a k: Zastosowanie programowania nieliniowego do równomiernie falistej aproksymacji stałego opóźnienia grupowego skutecznym opóźnieniem grupowym dolnoprzepustowego filtra LC	541
M. G ł o w a c k i, M. P i e r z c h a ł a, J. S t a n c l i k: Warunki eliminacji dynamicznych zniekształceń intermodulacyjnych we wzmacniaczach z ujemnym sprzężeniem zwrotnym oraz kompensacją biegunem i zerem	555
M. S a d o w s k i: Stabilizacja częstotliwości generatora przestrajanego w dyskretnej pętli sprzężenia fazowego	565
R. A r t i u c h: Kształtowanie transmitancji pętli synchronizacji fazy	581
W. N a k w a s k i: Efekt falowodowy w laserach paskowych z rezonatorami Fabry'ego-Perota	595
St. N o w a k: Wpływ czynników konstrukcyjnych i technologicznych na temperaturowy współczynnik rezystancji scalonych sieci rezystywnych grubowarstwowych	615
St. P a s z c z y Ń s k i: Model opisujący efekty piezorezystywne w rezystorach grubowarstwowych	645
K. Z a m ł y Ń s k i: Modelowanie termicznego sprzężenia zwrotnego w monolitycznych układach scalonych przy pomocy programu SPICE-2G	673
A. L e w i Ń s k a - R o m i c k a, S. P i s k o r s k i: Właściwości przetworników wiroprowodowych do nieniszczących badań prętów ferromagnetycznych	693
Z. M u s z a l s k i, L. G o ł u s i Ń s k i, W. R a f a l s k a, W. R a f a l s k i: Generacja impulsów wysokonapięciowych w zestyku ślizgowym maszyn elektrycznych. Część III	723

CONTENTS — TABLE DES MATIERES — INHALT

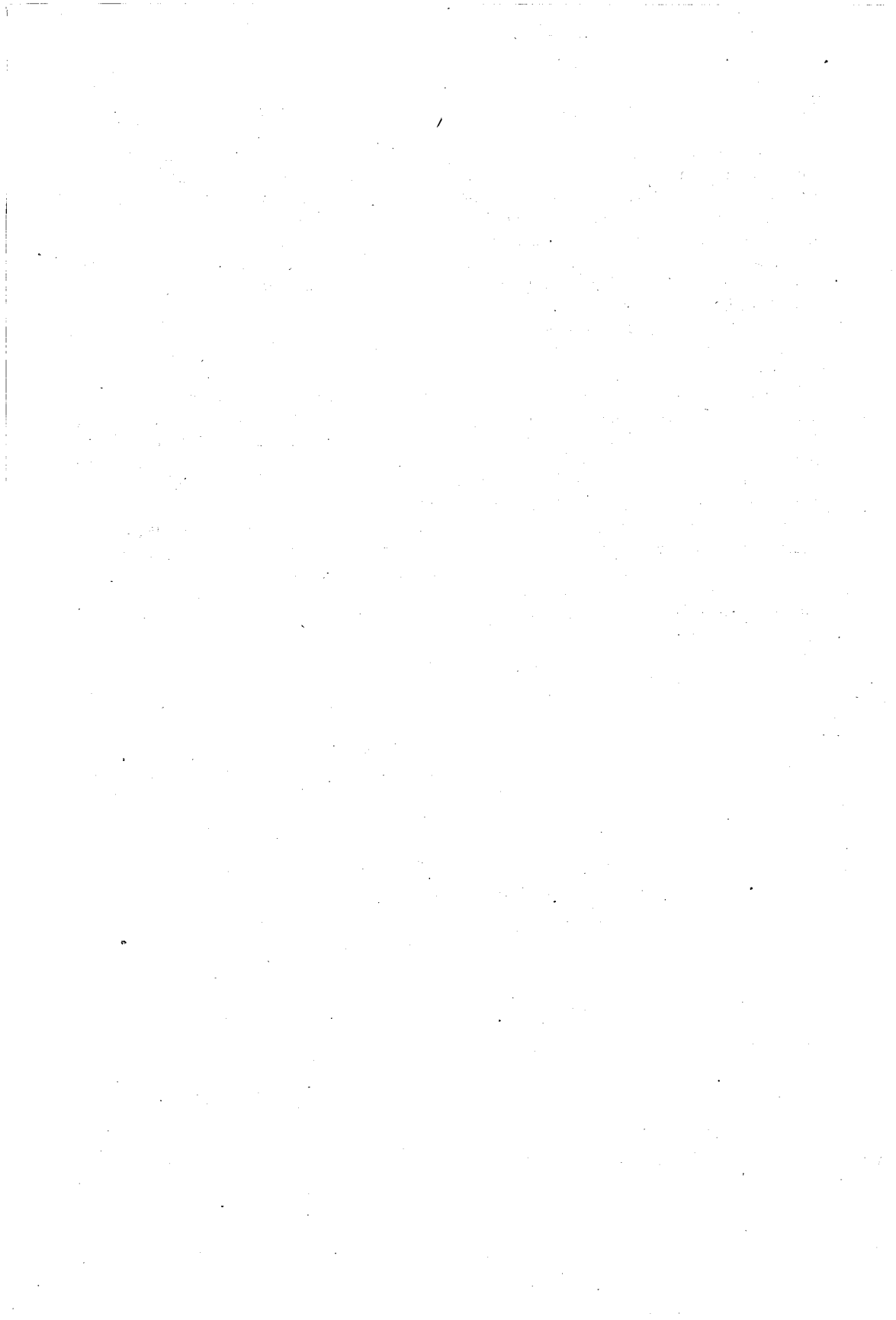
C. R a j s k i: On application of minimum power principle	515
Sur une certaine application du principe du pouvoir minimal	
Über gewisse Anwendung des Prinzips minimaler Leistung	
R. N a w r o w s k i: Electric field distribution and electric withstand strenght of a plate-hyperboloid point system with account taken of space charge	523
Répartition du champ électrique et de la rigidité diélectrique du système d'électrode plaque — pointe d'hyperboloïde prenant en considération la charge d'espace	
Verteilung des elektrischen Feldes und elektrische Festigkeit des Systems Platte-Hyperboloidspitze unter Berücksichtigung einer Raumladung	
K. N o g a: Estimation of the Nakagami distribution parameters basing on measured signal data	531
Estimation des paramètres de répartition de Nakagami en tenant compte des données de mesure du signal	
Parameterestimation der Nakagami-Verteilung auf Grund von Messdaten eines Signals	
F. W y s o c k a - S c h i l l a k: Application of nonlinear programming to the equiripple approximation of constant group delay by an effective group delay of a low-pass filter	541

	Application de la programmation non linéaire à l'approximation ondulée uniformément du délai de groupe constant par le délai de groupe effectif du filtre passe-bas LC	541
	Anwendung nichtlinearen Programmierung für die gleichmässig wellige Approximation der konstanten Gruppenlaufzeit mittels wirksam verzögerter Gruppenlaufzeit eines LC-Tiefpasses	
M. Głowacki, M. Pierzchała, J. Stanlik:	Conditions of elimination of the dynamic intermodulation distortion in feedback, pole-zero compensated amplifiers	551
	Conditions de l'élimination des distorsions dynamiques d'intermodulation dans un amplificateur avec contre réaction et compensation à l'aide des pôles et des zéros	
	Bedingungen für die Ausscheidung dynamischer Intermodulationsverzerrungen bei Verstärkern mit Gegenkopplung und Pol-Null-Kompensierung	
M. Sadowski:	Stabilization of a tuned oscillator frequency in a discrete phase-locked loop	565
	Stabilisation de fréquence d'un générateur réaccordé dans une boucle discrète de couplage de phase	
	Frequenzstabilisierung eines spannungsgesteuerten Generators mittels PLL-Technik	
R. Artiuch:	Shaping of phase locked loop transfer function	581
	Formation de la fonction de transfert d'une boucle de verrouillage de phase	
	Gestaltung der Übergangsfunktion einer Phasensynchronisierungsschleife	
W. Nakwaski:	Waveguiding effect in stripe-geometry Fabry-Perot laser diodes	595
	Effect de guide d'ondes dans les lasers longitudinaux avec résonateurs de Fabry-Perot	
	Wellenleitereffekt in Streifendiodenlasern mit dem Fabry-Perot-Resonator	
S. Nowak:	Effect of structure and manufacturing on the temperature factor of integrated thick-film resistance	615
	Influence des agents constructifs et technologiques sur le coefficient de température de la résistance des réseaux intégrés résistifs à couches épaisses	
	Einfluss konstruktiver und technologischer Faktoren auf den Temperaturkoeffizienten der Resistanz integrierter dickschichtiger Resistivnetze	
S. Paszczyński:	A model describing piezoresistive effects in thick film resistor	645
	Modèle dérivant les effets piézorésistifs dans les résistances à couches épaisses	
	Modell für die Beschreibung piezoresistiver Effekte in Dickschichtwiderständen	
K. Zamyński:	Modelling of thermal feedback in monolithic integrated systems by means of the program SPICE-2G	673
	Modelage de la rétroaction thermique dans les circuits intégrés monolithiques à l'aide du programme SPICE-2G	
	Thermische Modellierung der Rückkopplung mittels Programm SPICE-2G bei integrierten Monolithsystemen	
A. Lewińska-Romicka, S. Piskorski:	Properties of eddy-current transducers for non-destructive testing of ferromagnetic rods	693
	Propriétés des convertisseurs centrifuges pour les essais non-destructifs des barres ferromagnétiques	
	Eigenschaften der Wirbelstromwandler für zerstörungsfreie Prüfungen ferromagnetischer Stäbe	
Z. Muszalski, L. Golusiński, W. Rafalska, W. Rafalski:	High voltage impulse generation in the slide contact of electrical machines. Part III	723
	Processus du développement des impulsions de haute tension dans le contact glissant des machines électriques. III partie	
	Erzeugung von Hochspannungsimpulsen im Gleitkontakt elektrischer Maschinen. Teil III	

СОДЕРЖАНИЕ

Ч. Райски:	Об одном из применений принципа минимум мощности	515
Р. Навровски:	Образ электрического поля и электрическая прочность системы плоскость-гиперболоидальное острие с учетом пространственного заряда	523

К. Нога: Оценка параметров разложения Накагамы на основании данных измерения сигнала	531
Ф. Высоцка - Шиллак: Применение нелинейного программирования для решения задачи равноволновой аппроксимации постоянного группового времени запаздывания характеристикой рабочего времени запаздывания фильтра LC нижних частот	541
М. Гловацки, М. Пешчала, Ю. Станцлих: Условия элиминации динамических искажений в усилителе мощности с обратной связью и компенсацией полюсом и нулём	551
М. Садовски: Стабилизация частоты генератора перестраиваемого в дискретном контуре обратной связи	565
Р. Артюх: Формирование трансмитанса контура синхронизации фазы	581
В. Накваски: Волноводный эффект в полосковых лазерных диодах с резонатором Фабри-Перо	595
С. Новак: Влияние конструкционных и технологических параметров на коэффициент сопротивления интегральных резистивных сетей	615
С. Пащиньски: Модель описывающая пьезорезистивные эффекты в толстоплёночных резисторах	645
К. Замлынски: Моделирование температурной обратной связи в монолитных интегральных схемах при помощи программы SPICE-2G	673
А. Левиньска - Ромицка, С. Пискроски: Свойства вихретоковых переходных трансформаторных преобразователей для нерозрушающих испытаний ферромагнитных стержней	693
З. Мушальски, Л. Гольбусиньски, В. Рафальска, В. Рафальски: Генерирование высокочастотных импульсов в скользящем контакте электрических машин. Часть 3	723



Cena w wolnej sprzedaży **zł 2000,—**

Kwartalnik

Indeks 37483

POLSKA AKADEMIA NAUK
KOMITET ELEKTRONIKI I TELEKOMUNIKACJI

PLISSN 0035-9386

ROZPRAWY ELEKTROTECHNICZNE

TOM XXXV — ZESZYT 4

KWARTALNIK

PAŃSTWOWE WYDAWNICTWO NAUKOWE
WARSZAWA 1989

POLSKA AKADEMIA NAUK
KOMITET ELEKTRONIKI I TELEKOMUNIKACJI

ROZPRAWY ELEKTROTECHNICZNE

TOM XXXV — ZESZYT 4

KWARTALNIK

PAŃSTWOWE WYDAWNICTWO NAUKOWE
WARSZAWA 1989

RADA REDAKCYJNA

Przewodniczący

prof. dr inż. ADAM SMOLIŃSKI
członek rzeczywisty PAN

Członkowie

prof. dr hab. DANIEL JÓZEF BEM, prof. dr hab. inż. MICHAŁ BIAŁKO — czł. rzecz. PAN, prof. dr inż. KAZIMIERZ DZIĘCIOŁOWSKI, prof. dr hab. inż. STEFAN HAHN — czł. koresp. PAN, prof. dr inż. ANDRZEJ HAŁAS, prof. dr inż. ZDZISŁAW KACHLICKI, prof. dr hab. inż. BOHDAN MROZIEWICZ, prof. dr inż. JERZY OSIOWSKI, prof. dr inż. WITOLD ROSIŃSKI, prof. dr inż. STANISŁAW SŁAWIŃSKI, prof. dr hab. inż. STEFAN WĘGRZYN, prof. dr inż. ANDRZEJ ZIELIŃSKI, prof. dr inż. MARIAN ZIENTALSKI, prof. dr hab. inż. WIESŁAW WOLIŃSKI

REDAKCJA

Redaktor Naczelny

prof. dr hab. inż. WIESŁAW WOLIŃSKI

Zastępca Redaktora Naczelnego

doc. dr inż. KRYSZTYN PLEWKO

Sekretarz Odpowiedzialny

mgr KRYSZYNA LEŁAKOWSKA

ADRES REDAKCJI

00-665 Warszawa, ul. Nowowiejska 15/19
Politechnika, Instytut Telekomunikacji, Gmach Elektroniki, pok. 470
Dyżury Redakcji: środy i piątki, godz. 14—16
tel. 28 89 81 lub 210 07-737

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

PAŃSTWOWE WYDAWNICTWO NAUKOWE
Warszawa, Miodowa 10

Nakład 460 (385+75)	Oddano do składania w sierpniu 1989 r.
Ark. wyd. 28,0 Ark. druk. 20,5+1,0 wkł.	Podpisano do druku w sierpniu 1990 r.
Papier offset. kl. III, 80 g, 70×100	Druk ukończono w październiku 1990 r.
Zamówienie nr 644/12/89.	Cena zł 350,—

Drukarnia im. Rewolucji Październikowej, Warszawa

Wielomianowa metoda obliczania logarytmów dyskretnych w ciałach Galois charakterystyki dwa

MARIUSZ PIONTAS

Instytut Telekomunikacji, Politechnika Gdańska

Otrzymano 1988.06.29

Autoryzowano do druku 1989.02.14

Artykuł przedstawia problemy związane z obliczaniem logarytmów dyskretnych w ciałach skończonych charakterystyki dwa. Przedstawiono ulepszoną wersję algorytmu Wellsa wykorzystującą wagi Hamminga logarytmów.

Zakres zastosowań ciał skończonych nieustannie się powiększa, gdyż za ich pomocą można w zwarty i przejrzysty sposób opisywać wiele zagadnień. Ograniczając się tylko do tych dziedzin, które pozostają w bliskich związkach z telekomunikacją, możemy wymienić teorię kodów [9], wyszukiwanie informacji [11] i kryptografię. Problem logarytmowania w ciałach skończonych, którego dotyczy ta praca, zyskał na znaczeniu właśnie dlatego, że znaleziono dla niego praktyczne zastosowanie w kryptografii [15]. Diffie i Hellman w swej przełomowej pracy [14] podali metodę jawnej dystrybucji kluczy kryptograficznych poprzez publiczny kanał. Metoda ta polega na tym, że użytkownik i generuje losowo liczbę całkowitą X_i ze zbioru $\{1, \dots, p-1\}$, gdzie p jest dużą liczbą pierwszą. X_i pozostaje znana tylko swemu użytkownikowi, który umieszcza w ogólnie dostępnym zbiorze („książce adresowej”) liczbę

$$Y_i = \alpha^{X_i}(\text{mod } p), \quad (1)$$

gdzie α jest elementem pierwotnym ciała $\text{GF}(p)$. Takiej samej operacji dokonuje użytkownik j .

Para użytkowników może wymieniać informacje w sposób tajny za pomocą klucza K_{ij} , obliczanego według wzoru:

$$K_{ij} = \alpha^{X_i X_j}(\text{mod } p). \quad (2)$$

Klucz ten może być użyty np. w systemie DES¹⁾. Użytkownik j otrzymuje klucz przez podniesienie pobranej z publicznego zbioru wartości Y_i do potęgi X_j znanej tylko jemu

$$K_{ij} = Y_i^{X_j}(\text{mod } p) = \alpha^{X_i X_j}(\text{mod } p). \quad (3)$$

¹⁾ Skrót od Data Encryption Standard — standardowego algorytmu szyfracji i deszyfracji stosowanego powszechnie w USA

Taki sam klucz otrzymuje użytkownik i podnosząc pobraną z tego samego zbioru publicznego liczbę Y_j do znanej tylko sobie potęgi X_i . Nieautoryzowany użytkownik, aby uzyskać klucz, musiałby znaleźć logarytm dyskretny jednej z liczb Y_i lub Y_j .

Potęgowanie $\text{mod } p$ stanowi najlepiej znany przykład tzw. funkcji jednokierunkowej (ang. *one-way function*), tj. takiej bijekcji $f(x)$, obliczanie wartości której jest nieskomplikowane, natomiast obliczanie funkcji odwrotnej (tj. rozwiązanie równania $y = f(x)$ dla danego y) dla niemal wszystkich y jest zadaniem o znacznej złożoności obliczeniowej. Jeśli $b = \lceil \log_2 p \rceil$, to w przypadku potęgowania wymagane jest co najwyżej $2b$ mnożeń. Natomiast obliczanie logarytmu wymaga około $2^{b/2}$ mnożeń.

Niektóre systemy kryptograficzne opierają się właśnie na założeniu, że znajdowanie logarytmów jest złożone obliczeniowo. W specjalnych przypadkach można podać algorytmy o złożoności obliczeniowej mniejszej niż wykładnicza, np. Pohlig i Hellman [3] pokazali, jak znajdować logarytm elementu ciała w czasie proporcjonalnym do $\sqrt{p_1}$, gdzie p_1 jest największym pierwszym dzielnikiem $p-1 \subset p$ również np. [4,5]). Dla przypadku, gdy charakterystyka ciała K równa jest $\chi(K) = 2$, tj. gdy liczba elementów ciała jest potęgą liczby 2 (przypadek najczęściej rozpatrywany ze względu na łatwość implementacji i ułatwienie spowodowane tym, że dla dowolnego elementu x takiego ciała zachodzi równość $x+x=0$) algorytm logarytmowania o złożoności subwykładniczej został podany przez Coppersmitha [2]. Celem niniejszego opracowania nie jest znalezienie szybszego algorytmu, lecz pokazanie zależności między pewną metodą logarytmowania a dyskretną transformatą Fouriera i konsekwencji z tego wypływających.

Jak pokazał Wells [1], do obliczania logarytmu w ciałach Galois $\text{GF}(p)$, gdzie p jest liczbą pierwszą, można używać wielomianu $f(y)$ stopnia $p-2$ o współczynnikach z $\text{GF}(p)$ o następującej własności:

$$y \equiv \alpha^m (\text{mod } p) \Rightarrow m \equiv \sum_{j=1}^{p-2} c_j y^j (\text{mod } p), \quad c_j, x, y \in \{1, \dots, p-1\} \quad (4)$$

gdzie α jest elementem pierwotnym ciała.

Problem logarytmowania zostaje w ten sposób sprowadzony do interpolacji wielomianu w węzłach α^j , $j = 0, \dots, p-2$ [13]. Można dowieść, że dla danego p i danego α istnieje tylko jeden taki wielomian, a jego współczynniki dane są wzorem:

$$c_j = (1 - \alpha^j)^{-1}. \quad (5)$$

Ogólnie rzecz biorąc, współczynniki wielomianu, dla którego

$$f(\alpha^k) = g_k, \quad k = 0, \dots, p-2, \quad (6)$$

są określone zależnością:

$$\underline{c} = g \cdot M^{-1}(\alpha), \quad (7)$$

gdzie

$$\underline{c} = [c_0, \dots, c_{p-2}],$$

$$g = [g_0, \dots, g_{p-2}],$$

$$M_{i,j}(\alpha) = \alpha^{-ij} \quad 0 \leq i, j \leq p-2$$

$$M_{i,j}^{-1}(\alpha) = -M(\alpha^{-1}).$$

Oznacza to, że w celu otrzymania współczynników wielomianu przyjmującego pożądane przez nas wartości należy obliczyć odwrotną dyskretną transformatę Fouriera (RDFT) wektora, który na i -tej pozycji ma wartość, jaką funkcja wielomianowa powinna przyjąć w punkcie α^i .

Metoda znajdowania logarytmów przedstawiona powyżej jest nieprzydatna do celów praktycznych, gdyż otrzymany wielomian ma $p-2$ niezerowych współczynników i obliczanie logarytmu elementu ciała wymaga więcej operacji (mnożeń), niż znajdowanie logarytmu przez bezpośrednie wyszukiwanie.

Przedstawiony przez Wellsa sposób wyszukiwania logarytmu daje się bezpośrednio przenieść na ciała skończone będące rozszerzeniami ciał podstawowych, a w szczególności na ciała o charakterystyce 2. Zilustrujmy to przykładem. Na początku musimy wybrać reprezentację elementów ciała oraz zdefiniować dodawanie i mnożenie. W przypadku ciał $\text{GF}(p)$ posługiwaliśmy się reprezentacją najbardziej naturalną utożsamiając elementy ciała z liczbami z pierścienia $\mathbb{Z}/p\mathbb{Z}$ i wykonując działania modulo p . Spośród wielu możliwych reprezentacji [6,10] dla uproszczenia zapisu wybieramy następujące odwzorowanie

$$\text{dec}(\cdot): \text{GF}(2^n) \rightarrow \mathbb{Z}/2^n\mathbb{Z} \quad (8)$$

$$A(x) = \sum_{i=0}^{n-1} a_i \odot x^i \Rightarrow \text{dec}(A(x)) = \sum_{i=0}^{n-1} a_i 2^i \quad \text{dla } A(x) \in \text{GF}(2^n),$$

tj. elementowi ciała będzie odpowiadała w naszym zapisie liczba dziesiętna, otrzymana przez podstawienie wartości charakterystyki ciała do potraktowanego jako wielomian wektorowego zapisu elementu.

Działania dodawania i mnożenia, oznaczane $\oplus$ i $\odot$, definiowane są jako działania na wielomianach modulo wielomian generujący $P(x)$.

Dalsze rozważania dotyczyć będą sytuacji, gdy rząd grupy multiplikatywnej K^* ciała K jest liczbą pierwszą, tzn. gdy $|K^*| = 2^p - 1 = M_p$ jest liczbą pierwszą Mersenne'a [12]. Wiadomo, że p jest wtedy również liczbą pierwszą. Ciała takie są szczególnie interesujące, gdyż:

- ciało K nie ma żadnych podciał właściwych,
- każdy element grupy K^* różny od jedności jest jej generatorem
- do logarytmowania w ciele niemożliwe jest użycie algorytmu Pohliga-Hellmana [2,3].

Dla dalszych rozważań niezbędne jest wprowadzenie pojęcia *śladu*. Przy przyjętych założeniach definicja śladu [8] redukuje się do następującej postaci: *śladem elementu* β w ciele $\text{GF}(2^p)$ nazywamy sumę

$$\text{tr}(\beta) = \sum_{i=0}^{p-1} \beta^{2^i}, \quad (9)$$

gdzie symbol $\sum$ oznacza sumowanie w ciele $\text{GF}(2^p)$. Dowodzi się [16], że ślad przyjmuje wartości z ciała podstawowego $\text{GF}(2)$.

Przejdźmy do przykładu, który pokaże, jak działa algorytm Wellsa w rozszerzeniu ciała $\text{GF}(2)$.

Przykład

Wybermy wielomian generujący $P(x) = x^3 + x^2 + 1$ i skonstruujmy ciało $GF(2^3)$. W tabelce przedstawiono elementy grupy multiplikatywnej $GF^*(2^3)$ w postaci wielomianowej, reprezentację dziesiętną elementów ciała i ich odwrotności oraz ślady elementów.

i	$\alpha^i = A(x)$	$\text{dec}(\alpha^i)$	$\text{dec}(\alpha^{-i})$	$\text{tr}(\alpha^i)$
1	x	2	6	1
2	x^2	4	3	1
3	$x^2 + 1$	5	7	0
4	$x^2 + x + 1$	7	5	1
5	$x + 1$	3	4	0
6	$x^2 + x$	6	2	0
7	1	1	1	1

$\oplus$	0	1	2	3	4	5	6	7
0	0	1	2	3	4	5	6	7
1	1	0	3	2	5	4	7	6
2	2	3	0	1	6	7	4	5
3	3	2	1	0	7	6	5	4
4	4	5	6	7	0	1	2	3
5	5	4	7	6	1	0	3	2
6	6	7	4	5	2	3	0	1
7	7	6	5	4	3	2	1	0

$\odot$	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7
2	0	2	4	6	5	7	1	3
3	0	3	6	5	1	2	7	4
4	0	4	5	1	7	3	2	6
5	0	5	7	2	3	6	4	1
6	0	6	1	7	2	4	3	5
7	0	7	3	4	6	1	5	2

Dla tak zdefiniowanego ciała wielomianowa funkcja logarytmująca otrzymana za pomocą wzoru [5] jest równa

$$f(y) = 7 \oplus 5 \odot y \oplus 3 \odot y^2 \oplus 1 \odot y^3 \oplus 6 \odot y^4 \oplus 4 \odot y^5 \oplus 2 \odot y^6. \quad (10)$$

Niech $\mathcal{K}_i$, $i = 0, \dots, p$, będzie rodziną zbiorów, zdefiniowanych następująco:

$$\mathcal{K}_i = \{j: 0 \leq j \leq 2^p - 2 \text{ i } w(\text{bin}(j)) = i\} \quad (11)$$

gdzie: $w(\cdot)$ jest wagą Hamminga słowa binarnego; $\text{bin}(\cdot)$ jest funkcją zamieniającą liczbę dziesiętną na liczbę binarną.

Zbiór $\mathcal{K}_i$ jest zatem zbiorem takich liczb $j = 0, \dots, 2^p - 2$, których rozwinięcie binarne zawiera dokładnie i jedynek.

Zdefiniujmy teraz zbiór $p+1$ wektorów $k_i[j]$, $i = 0, \dots, p$, $j = 0, \dots, 2^p - 2$ w następujący sposób:

$$k_i[j] = \begin{cases} 1 & \text{gdy } j \in \mathcal{K}_i \\ 0 & \text{w przeciwnym przypadku} \end{cases} \quad (12)$$

Rozbijmy teraz zbiór liczb modulo $2^p - 1$ na klasy elementów sprzężonych $\mathcal{A}_j$ [8]:

$$\mathcal{A}_j = \{j, 2 \cdot j, \dots, 2^{m_j-1} \cdot j\}, \quad (13)$$

gdzie m_j jest najmniejszą liczbą dodatnią spełniającą równość $2^{m_j} \cdot j \equiv j \pmod{2^p - 1}$. Jak łatwo stwierdzić, $m_j = p$. Klasę elementów sprzężonych, będziemy dla zwięzłości nazywać dalej *hordą*.

Przykład:

Niech $P(x) = x^5 + x^2 + 1$. Mamy wtedy

	$\text{tr}(\alpha^i), i \in \mathcal{A}_i$
$\mathcal{A}_0 = \{0\}$	0
$\mathcal{A}_1 = \{1, 2, 4, 8, 16\}$	0
$\mathcal{A}_3 = \{3, 6, 12, 17, 24\}$	1
$\mathcal{A}_5 = \{5, 9, 10, 18, 20\}$	1
$\mathcal{A}_7 = \{7, 14, 19, 25, 28\}$	0
$\mathcal{A}_{11} = \{11, 13, 21, 22, 26\}$	1
$\mathcal{A}_{15} = \{15, 23, 27, 29, 30\}$	0
$\mathcal{H}_0 = \{0\}$	$= \mathcal{A}_0$
$\mathcal{H}_1 = \{1, 2, 4, 8, 16\}$	$= \mathcal{A}_1$
$\mathcal{H}_2 = \{3, 5, 6, 9, 10, 12, 17, 18, 20, 24\}$	$= \mathcal{A}_3 \cup \mathcal{A}_5$
$\mathcal{H}_3 = \{7, 11, 13, 14, 19, 21, 22, 25, 26, 28\}$	$= \mathcal{A}_7 \cup \mathcal{A}_{11}$
$\mathcal{H}_4 = \{15, 23, 27, 29, 30\}$	$= \mathcal{A}_{15}$
$\mathcal{H}_5 = \{31\}$	

Te hordy, które składają się na zbiór $\mathcal{H}_j$ nazwijmy *hordami stowarzyszonymi*. Stosując RDFS do nietrywialnych wektorów, które na pozycjach wyznaczonych zbiorami $\mathcal{A}_i$ i $\mathcal{H}_j$ mają wartość 1, a na pozostałych 0, otrzymujemy

$$\begin{aligned}
 F^{-1}\{a_1\} &= a_0 + a_5 + a_7 + a_{11} \\
 F^{-1}\{a_3\} &= a_0 + a_3 + a_7 + a_{15} \\
 F^{-1}\{a_5\} &= a_0 + a_1 + a_5 + a_{15} \\
 F^{-1}\{a_7\} &= a_0 + a_1 + a_3 + a_7 \\
 F^{-1}\{a_{11}\} &= a_0 + a_1 + a_{11} + a_{15} \\
 F^{-1}\{a_{15}\} &= a_0 + a_3 + a_5 + a_{11} \\
 F^{-1}\{k_1\} &= F^{-1}\{a_1\} = a_0 + a_5 + a_7 + a_{11} \\
 F^{-1}\{k_2\} &= F^{-1}\{a_3 + a_5\} = a_1 + a_3 + a_5 + a_7 \\
 F^{-1}\{k_3\} &= F^{-1}\{a_7 + a_{11}\} = a_3 + a_7 + a_{11} + a_{15} \\
 F^{-1}\{k_4\} &= F^{-1}\{a_{15}\} = a_0 + a_3 + a_5 + a_{11} \\
 \mathcal{E}_1 &= \mathcal{A}_0 \cup \mathcal{A}_5 \cup \mathcal{A}_7 \cup \mathcal{A}_{11} \\
 \mathcal{E}_2 &= \mathcal{A}_1 \cup \mathcal{A}_3 \cup \mathcal{A}_5 \cup \mathcal{A}_7 \\
 \mathcal{E}_3 &= \mathcal{A}_3 \cup \mathcal{A}_7 \cup \mathcal{A}_{11} \cup \mathcal{A}_{15} \\
 \mathcal{E}_4 &= \mathcal{A}_0 \cup \mathcal{A}_3 \cup \mathcal{A}_5 \cup \mathcal{A}_{11}
 \end{aligned}$$

Fakt, iż hordy transformują się na hordy nie jest przypadkiem. Zachodzi bowiem stwierdzenie:

Twierdzenie:

Odwrotna dyskretna transformata Fouriera wektora $\underline{a}_k$ dana jest wzorem:

$$F^{-1}\{\underline{a}_k\} = \sum_{j=0}^{(2^p-2)/p} d_{jk} \cdot \underline{a}_j, \text{ gdzie } d_{jk} = 0 \text{ lub } 1 \quad (14)$$

oraz

$$\sum_{j=1}^{(2^p-2)/p} d_{jk} = (2^p-2)/2 \cdot p = (2^p-1)/p. \quad (15)$$

Dowód:

Niech $\underline{b}_k = F^{-1}\{\underline{a}_k\}$. Mamy:

$$\underline{b}_k[j] = \sum_{i=0}^{2^p-2} \alpha^{-ij} \underline{a}_k[i] = \sum_{l=0}^{p-1} \alpha^{-jk2^l} \underline{a}_k[k \cdot 2^l] = \text{tr}(\alpha^{-jk})$$

Funkcja $f_k: \mathbb{Z}/M_p\mathbb{Z} \rightarrow \mathbb{Z}/M_p\mathbb{Z}$ dana wzorem $f_k(j) = -kj \pmod{M_p}$ dla $k \pmod{M_p} \neq 0$ bijekcją. Zatem elementy wektora $\underline{b}_k$ stanowią permutację wektora $\underline{tr}$, dla którego $\underline{tr}[i] = \text{tr}(\alpha^i)$. Funkcja f_k przekształca hordy w hordy, co oznaczają, że np. w zamieszczonym na poprzedniej stronie przykładzie każdy ze zbiorów $\mathcal{A}_1, \mathcal{A}_3, \mathcal{A}_5, \mathcal{A}_7, \mathcal{A}_{11}, \mathcal{A}_{15}$ jest przekształcany przez tę funkcję na jeden z tych zbiorów.

O wyjściowym wektorze $\underline{tr}$ wiemy, że:

— $\underline{tr}[0] = \text{tr}(\alpha^0) = \text{tr}(1) = p \pmod{2}$;

— z pozostałych 2^p-2 pozycji wektora dokładnie połowa to jedynki i połowa to zera;

— jedynki występują na pozycjach odpowiadających hordom, tzn. jeżeli pozycja odpowiadająca jednemu z elementów hordy zawiera jedynkę, to zawierają ją również pozycje odpowiadające pozostałym elementom hordy.

Funkcja f_k powoduje, że jedynki występujące na miejscach odpowiadających jednej hordzie przechodzą na miejsca odpowiadające innej hordzie. Tak więc otrzymany po permutacji wektora $\underline{tr}$ wektor $\underline{b}_k$ jest sumą wektorów $\underline{a}_j$, przy czym oprócz $\underline{a}_0$ w skład tej sumy wchodzi $(2^p-2)/(2 \cdot p)$ wektorów $\underline{a}_j$, co wynika z tego, że każda horda (oprócz $\underline{a}_0$) liczy p elementów. ■

Jako konsekwencje powyższego twierdzenia otrzymujemy, że wektory $\underline{c}_i$ również przyjmują wartości zerojedynkowe.

Nazwijmy wielomianem *i-logarytmującym* wielomian $f_i(x)$ o współczynnikach równych 1 na pozycjach wyznaczanych przez $\underline{c}_i$, i poza tym równych 0. Odpowiadająca mu funkcja wielomianowa zachowuje się następująco (p. (11)):

$$f_i(a^j) = \begin{cases} 1 & \text{gdy } w(\text{bin}(j)) = i \\ 0 & \text{w przeciwnym przypadku} \end{cases}$$

Z podstawowych własności binarnych przedstawień liczb dziesiętnych wynika, że dla $1 \leq j \leq 2^p-1$, $1 \leq i \leq p-1$ mamy:

$$\underline{c}_i[j] = \underline{c}_{p-i}[2^p-1-j], \quad (16)$$

$$\underline{k}_i[j] = \underline{k}_{p-i}[2^p-1-j], \quad (17)$$

a zatem korzystając z własności równań symetrycznych mamy:

$$f_{p-i}(x) = x^{2^p-1} \cdot f_i\left(\frac{1}{x}\right) = f_i\left(\frac{1}{x}\right). \quad (18)$$

Zatem wystarczy obliczyć wartości $[p/2]$ funkcji wielomianowych. Obliczenie wartości funkcji wielomianowej i -logarytmującej w punkcie y możemy zastąpić obliczaniem sumy śladów takich potęg elementu y , które równe są indeksom j zbiorów $\mathcal{A}_j$ wchodzących w skład zbioru $\mathcal{E}_i$, tj.

$$J_i = \{j: \mathcal{E}_i = \bigcup_j \mathcal{A}_j\}, f_i(y) = \sum_{j \in J_i} \oplus \text{tr}(y^j). \quad (19)$$

Jak widać ze wzoru (11), obliczanie śladu elementu ciała $\text{GF}(2^p)$ wymaga p potęgowań, nie jest zatem zadaniem skomplikowanym rachunkowo [7]. Wydaje się jednak, że w szczególnym przypadku, jaki rozważamy, prostszym i szybszym rozwiązaniem jest skorzystanie z addytywności śladu i obliczanie śladu wektorów bazowych w ciele traktowanym jako przestrzeń wektorowa nad $\text{GF}(2)$. Wyjaśnijmy to na przykładzie.

Przykład

Niech $P(x) = x^5 + x^2 + 1$. Wybierzmy najprostszą bazę o postaci $\{1, x, x^2, x^3, x^4\}$. Obliczmy ślady elementów bazowych:

$\text{tr}(1) = \sum_{i=0}^{p-1} 1 \equiv p \pmod{2}$ w $\text{GF}(2^p)$, a zatem w $\text{GF}(2^5)$ $\text{tr}(1) = 1$. Jak łatwo sprawdzić, zachodzi następująca równość

$$\text{tr}(y^2) = \text{tr}(y), y \in \text{GF}(2^p). \quad (20)$$

Mamy zatem

$$\text{tr}(x) = \text{tr}(x^2) = \text{tr}(x^4) = \alpha + \alpha^2 + \alpha^4 + \alpha^8 + \alpha^{16} = 2 \oplus 4 \oplus 16 \oplus 13 \oplus 27 = 0 \quad (21)$$

Zauważmy, że

$$\text{tr}(x) = 0 \Rightarrow \text{tr}(x^8) = 0 \Rightarrow \text{tr}(\alpha^8) = 0 = \text{tr}(x^3 + x^2 + 1) = 1 + 0 + \text{tr}(x^3) \Rightarrow \text{tr}(x^3) = 1 \quad (22)$$

Widzimy, że $\text{tr}(\beta) = \text{tr}\left(\sum_{i=1}^{p-1} a_i \cdot x^i\right) = a_0 \oplus a_3$.

Zatem dla ciała $\text{GF}(2^p)$ potrzeba co najwyżej $(p-1)^2$ potęgowań, aby obliczyć ślady $p-1$ elementów bazowych różnych od 1.

Posiadając sprawny algorytm znajdowania śladów, pokażemy, jak za ich pomocą znaleźć można wagę logarytmu elementu ciała.

Przykład

W ciele $\text{GF}(2^5)$ zdefiniowanym, jak w poprzednim przykładzie, znajdziemy wagę wykładnika elementu $y = 15$ (wektorowo $x^3 + x^2 + x^1 + x^0$).

$$y = 15$$

$$f_1(y) = \text{tr}(y^0) \oplus \text{tr}(y^5) \oplus \text{tr}(y^7) \oplus \text{tr}(y^{11}) = 1 \oplus 1 \oplus 1 \oplus 1 = 0$$

$$f_2(y) = \text{tr}(y^1) \oplus \text{tr}(y^3) \oplus \text{tr}(y^5) \oplus \text{tr}(y^7) = 0 \oplus 0 \oplus 1 \oplus 1 = 0$$

$$y^{-1} = 13,$$

$$f_1(y^{-1}) = \text{tr}(y^{-0}) \oplus \text{tr}(y^{-5}) \oplus \text{tr}(y^{-7}) \oplus \text{tr}(y^{-11}) = 1 \oplus 0 \oplus 1 \oplus 1 = 1$$

$$f_2(y^{-1}) = \text{tr}(y^{-1}) \oplus \text{tr}(y^{-3}) \oplus \text{tr}(y^{-5}) \oplus \text{tr}(y^{-7}) = 0 \oplus 1 \oplus 1 \oplus 0 = 0$$

$$f_1(y^{-1}) = 1 \Rightarrow w(\text{bin}(\log y)) = 4.$$

Jest tak istotnie, gdyż jak możemy sprawdzić, $15 = \alpha^{23}$, a $w(\text{bin}(23)) = 4$.

Pozostaje jeszcze problem odwzorowania wagi wykładnika na sam wykładnik. W tym celu należy dobrać ciąg elementów $x = x_0, x_1, \dots, x_z$ tak, aby na podstawie zbioru wartości $\{y_i: y_i = w(\text{bin}(\log(x_i)))\}$, $i = 1, \dots, z$ można było jednoznacznie określić $\log(x_0)$. Nie będziemy tu rozważać dokładnie tego problemu, zauważając jedynie, że zachodzi równość

$$w(\text{bin}(\log(\alpha^{2^{p-1}} \cdot x))) - w(\text{bin}(\log(x))) = \begin{cases} 1, & \text{gdy } \log(x) < 2^{p-1} \\ -1, & \text{gdy } \log(x) > 2^{p-1} \end{cases} \quad (23)$$

Nasuwa się przypuszczenie, że stosując wyszukiwanie połówkowe można znaleźć wykładnik elementu, którego waga wynosi ω , w $\omega + 1$ krokach ($0 < \omega < p - 1$). Opierając się na tym założeniu obliczymy średnią liczbę $\underline{EL}$ operacji znajdowania wykładnika niezbędnych do obliczania wartości logarytmu. W ciele $GF(2^p)$ istnieje $\binom{p}{l}$ elementów x o wadze wykładnika $w(\text{bin}(\log(x))) = l$. Prawdopodobieństwo, że losowo wybrany element ma wagę wykładnika równą l wynosi

$$P_l = P\{w(\text{bin}(\log(x))) = l\} = \binom{p}{l} / 2^p. \quad (24)$$

Niech L_l oznacza liczbę operacji znajdowania wagi wykładnika niezbędnych do znalezienia wykładnika, jeśli waga równa jest l . W oparciu o uczynione wyżej założenie i symetrię rozkładu (26) mamy:

$$L_l = \begin{cases} l+1 & \text{dla } 1 \leq l \leq [p/2] \\ p-l+1 & \text{dla } [p/2] < l \leq p-1 \end{cases}. \quad (25)$$

Szukana wartość równa jest z definicji

$$\underline{EL} = \sum_{l=1}^{p-1} L_l P_l. \quad (26)$$

Zarówno L_l jak i P_l są symetryczne, a stąd

$$L_l = L_{p-l} \wedge P_l = P_{p-l} \Rightarrow L_l P_l = L_{p-l} P_{p-l}. \quad (27)$$

Po przekształceniu otrzymujemy:

$$\underline{EL} = (p+2)/2 - \left(l(p+1)/2j \cdot \binom{p}{p/2} - 2 \right) / 2^p \quad (28)$$

$$\underline{EL} < p/2, \quad \text{dla } p > 6. \quad (29)$$

Przedstawiona metoda nie ma większego znaczenia praktycznego z uwagi na znaczną liczbę hord wchodzących w skład wielomianów i -logarytmujących. Wydaje się jednak, iż zastosowanie funkcji wielomianowych do znajdowania logarytmów dyskretnych może zaowocować szeregiem nowych, szybkich algorytmów. Szczególnie obiecujące wydaje się zastosowanie wielomianów i funkcji permutujących.

BIBLIOGRAFIA

1. A. L. Wells: *A Polynomial Form for Logarithms Modulo a Prime*, IEEE Trans. on Inf. Theory, vol. IT-30, s. 845—846, Novemb. 1984
2. D. Coppersmith: *Fast Evaluation of Logarithms in Fields of Characteristic Two*, IEEE Trans. on Inf. Theory, vol. IT-30, s. 587—601, July 1984
3. S. C. Pohlig, M. E. Hellman: *An Improved Algorithm for Computing Logarithms over $GF(p)$ and Its Cryptographic Significance*. IEEE Trans. on Inf. Theory, vol. IT-24, s. 106—109, Jan. 1978
4. J. E. Sachs, S. Berkovits: *Probabilistic Analysis and Performance Modelling of the 'Swedish' Algorithm and Modification*. D. Chaum. Ed. Advances in Cryptology. Proceedings of Crypto 83, Plenum Pub. Co. 1984
5. T. ElGamal: *A Subexponential-Time Algorithm for Computing Discrete Logarithms over $GF(p^2)$* , D. Chaum, Ed.: Advances in Cryptology. Proceedings of Crypto 83. Plenum Pub. Co. 1984
6. G. E. Collins, M. Mignotte, F. Winkler: *Arithmetic In Basic Algebraic Domains*. Computer Algebra. Symbolic and Algebraic Computation. (Kompiuternaia algebr. Simvolnye i algebraičeskie vyčislenia), Moskwa, Mir. 1986
7. K. Imamura: *On Simultaneous Computation of Addition Table and Trace Table in a Finite Field*, The Transactions of the IECE of Japan, 1981. vol. E64, s. 90—91, No. 2 February
8. R. E. Blahut: *Theory and Practice of Error Control Codes*, (Teoria i praktyka kodov, kontroliruiusčih ošibki). Moskwa: Mir, 1986
9. G. Birkhoff, T. C. Bartee: *Współczesna algebra stosowana*. Warszawa: PWN, 1983
10. Cz. Kościelny: *Programowa realizacja działań w ciałach skończonych do zastosowań w technice kodowania korekcyjnego i kryptografii*. Wrocław, 1983
11. W. M. Turski: *Niektóre zagadnienia struktur danych w: Problemy przetwarzania informacji*. t. 2. Warszawa: WNT, 1974
12. W. Narkiewicz: *Classical Problems in Number Theory*. Warszawa: PWN, 1986
13. A. V. Aho, J. E. Hopcroft, J. D. Ullman: *Projektowanie i analiza algorytmów komputerowych*, Warszawa: PWN 1983.
14. W. Diffie, M. E. Hellman: *New Directions in Cryptography*. IEEE Trans. on Information Theory, vol. IT-22, s. 644—654, No. 6, November 1976
15. T. Harlestad, R. Johannesson: *On computing logarithms over $GF(2^p)$* . BIT, Bind 21, s. 326—334, No. 3, 1981
16. K. Ireland, M. Rosen: *A Classical Introduction to Modern Number Theory*. New York—Heidelberg—Berlin, Springer Verlag, 1982

M. PIONTAS

A POLYNOMIAL METHOD OF OBTAINING A DISCRETE LOGARITHM IN GALOIS FIELD OF CHARACTERISTIC TWO

Summary

The paper gives a survey of problems connected with computing discrete logarithms in finite fields of characteristic two. The improved version of Wells' algorithm based on Hamming weight is presented.

M. PIONTAS

UTILISATION DES FONCTIONS POLYNÔMES PENDANT LE CALCUL DES LOGARITHMES DANS LES CORPS DE GALOIS

Résumé

Dans l'article sont présentés les problèmes liés à la recherche du logarithme discret dans les corps finis à la caractéristique deux. On a amélioré l'algorithme Wells grâce à la recherche de la mesure du logarithme de Hamming.

M. PIONTAS

EINE POLYNOMMETHODE ZUR LÖSUNG DISKRETER LOGARITHMEN IN
GALOISFELDERN DER CHARAKTERISTIK ZWEI

Zusammenfassung

In diesem Artikel werden Probleme behandelt, die verbunden sind mit der Suche nach diskreten Logarithmen in endlichen Körpern der Charakteristik Zwei. Der Algorithmus von Wells wird verbessert, um das Vorhandensein von Hamming-Gewichten im den diskreten Logarithmen zu finden.

М. ПИОНТАС

МНОГОЧЛЕННЫЙ МЕТОД НАХОЖДЕНИЯ ДИСКРЕТНОГО ЛОГАРИФМА
В ПОЛЯХ ГАЛУА ХАРАКТЕРИСТИКИ ДВА

Резюме

Приводится обзор методов нахождения дискретного логарифма в полях Галуа характеристики два. Представленный усовершенствованный алгоритм получен при использовании метода Хемминга веса логарифмов.

Redukcja matryc logicznych typu PLA metodą składania

KRZYSZTOF JASIŃSKI

Instytut Telekomunikacji, Politechnika Warszawska

Otrzymano 1988.09.14

Autoryzowano do druku 1989.03.22

W pracy przedstawiono podstawowe idee i algorytmy kompresji programowanych matryc logicznych PLA metodą składania (ang. folding). Sformułowano problem i podano jego formalny model. Naszkicowano strategię rozwiązania, opisano wybrane metody i implementacje algorytmów wraz ze schematami i specyfikacją podstawowych procedur. W szczególności przedstawiono i porównano dwie różne, ze względu na stosowaną strategię, metody: przybliżoną metodę heurystyczną typu branch and bound oraz metodę dokładną systematyczną typu best—first—search.

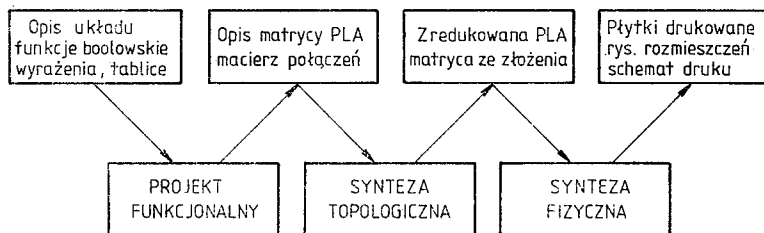
1. WSTĘP

Układy wielkiej skali integracji LSI stały się tak złożonymi strukturami, że ich stosowanie często wymaga posługiwania się systematycznymi metodami projektowymi, które pozwalają w rozsądnie krótkim czasie weryfikować poprawność zarówno projektu formalnego jak i jego elektrycznej realizacji. Takie metody stają się jeszcze bardziej potrzebne, albo wręcz nieodzowne, gdy wkraczamy w dziedzinę układów VLSI. Przy projektowaniu strukturalnym systemów wielkiej skali integracji coraz szerzej stosuje się zintegrowane matryce logiczne, a w szczególności Programowane Matryce Logiczne PLA (ang. PROGRAMMABLE LOGIC ARRAY), które okazały się bardzo efektywnym materiałem konstrukcyjnym do implementacji wielowyjściowych układów funkcji kombinacyjnych. Proces syntezy układu kombinacyjnego w strukturze PLA można podzielić na podzadania i wykonać w trzech krokach jako: projekt funkcjonalny, projekt topologiczny i synteza fizyczna układu. Schemat blokowy procesu projektowania układu w PLA przedstawiono na rys. 1.

Pierwszy etap tego procesu obejmuje przekształcenie opisu układu z postaci funkcji boolowskich w zbiór minimalnych, dwupoziomowych formuł kanonicznych (sum iloczynów lub iloczynów sum), które zapewniają optymalizację układu realizowanego w PLA pod względem zajmowanej powierzchni krzemu, szybkości działania, konfiguracji doprowadzeń, pobieranej mocy itp.

Do optymalizacji projektu na poziomie funkcjonalnym stosowane są przede wszystkim metody strukturalnego podziału układu oraz metody minimalizacji funkcji w postaci kanonicznej.

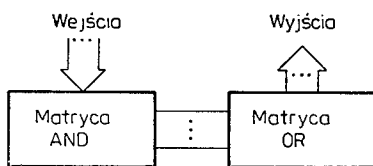
Po wyznaczeniu optymalnej postaci strukturalnej układu, odpowiedniej do realizacji w PLA, następnym zadaniem jest odwzorowanie jego logicznej struktury w topologiczną mapę połączeń PLA. Optymalizacja projektu na tym etapie jest właśnie przedmiotem niniejszego artykułu. Do wykonania tego zadania posłużymy się metodą kompresji PLA,



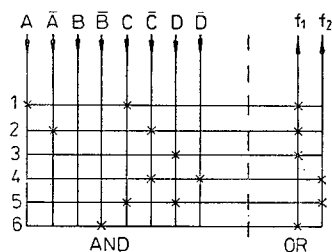
Rys. 1. Projektowanie układu w PLA

zwaną składaniem (ang. folding). Wykorzystując własności macierzy rzadkich oraz pomocnicze pojęcia teoriomnogościowe i grafowe sformułujemy problem optymalnego składania (foldingu) PLA, a następnie opiszemy dwie najbardziej efektywne strategie i algorytmy rozwiązania problemu składania (foldingu): *heurystyczny* i *branch and bound*.

Celem składania PLA jest wyznaczenie maksymalnego zbioru par kolumn (*i/lub* wierszy), które mogą być zrealizowane w tej samej kolumnie (*i/lub* wierszu) fizycznej macierzy logicznej. Jeżeli każda kolumna i każdy wiersz będą objęte składaniem, to otrzymana w rezultacie redukcji macierz PLA będzie wymagała do realizacji jedynie ok. 25% powierzchni zajmowanej przez pierwotną PLA. Ogólny model struktury PLA jest pokazany na rys. 2.



Rys. 2. Ogólna struktura PLA



Rys. 3. Przykład realizacji funkcji w PLA

Wejścia matrycy PLA są doprowadzone do linii pionowych matrycy AND. Wyjścia tej matrycy, reprezentujące kombinacje wejść PLA i ich negacji, są równocześnie wejściami matrycy OR. Wyjścia matrycy OR są sumami iloczynów opisującymi funkcje boolowskie realizowane w PLA. Przykład funkcji, opisanych formułami:

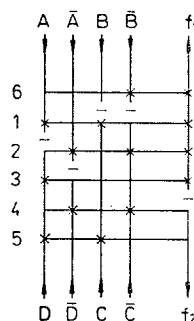
$$f_1 = AC + \bar{A}\bar{C} + D + \bar{B}, \quad f_2 = \bar{C}\bar{D} + CD,$$

zrealizowanych w PLA jest pokazany na rys. 3. Na rysunku tym x w matrycy AND reprezentuje zmienną wejściową w iloczynie, zaś x w matrycy OR odpowiada obecności danego iloczynu w sumie wyjściowej. Matryce PLA mogą być realizowane zarówno w tech-

nologii bipolarnej jak i w technice MOS, ale nie ma to istotnego wpływu na rozwiązanie problemu składania PLA.

Przy realizowaniu układów w PLA wykorzystywany jest symboliczny opis struktury logicznej w postaci tablicy prawdy. Tablica ta ma liczbę kolumn równą sumie wejść i wyjść matrycy PLA oraz liczbę wierszy równą liczbie iloczynów — składników sumy. Tablica prawdy matrycy logicznej z rys. 3 pokazana jest na rys. 4. W tej tablicy symbol „1” na

$\begin{matrix} k \\ w \end{matrix}$	1	2	3	4	5	6
1	1	0	1	0		
2	1	0	0	1		
3	0	1	0	1		
4	0	1	1	0		



Rys. 4. Tablica prawdy matrycy logicznej z rys. 3 Rys. 5. Matryca PLA z rys. 3 po założeniu kolumn

pozycji (i, j) oznacza w matrycy AND obecność j -tego wejścia w i -tym iloczynie, symbol „0” — obecność dopełnienia j -tego wejścia w i -tym iloczynie, zaś symbol „—” reprezentuje brak połączenia. Symbol „1” w kolumnie wyjściowej na pozycji (i, j) oznacza obecność i -tego iloczynu w sumie reprezentującej j -tą funkcję j .

Tablica prawdy wiernie odwzorowuje strukturę logiczną matrycy PLA i może być bezpośrednio wykorzystana do jej zaprogramowania na etapie syntezy fizycznej. Matryce PLA dla szerokiej klasy zastosowań mają niewielką gęstość punktów połączeniowych, a opisują je tzw. macierze rzadkie. Realizacje takie są wysoce nieoszczędne pod względem zajmowanej powierzchni a ponadto wolniejsze w działaniu. W związku z tym ważnym staje się zagadnienie redukcji wielkości matryc PLA, któremu poświęcony jest ten artykuł. W szczególności przedstawiono w nim algorytmy redukcji PLA, wykorzystujące metody optymalnego składania PLA, którego zasady są opisane poniżej (rozdz. 2).

W rozdziale 3 sformułowano problem optymalnego składania PLA. W rozdziale 4 jest przedstawiony model teoriografowy tego zadania. Rozdział 5 zawiera opis wybranych metod i algorytmów składania, opracowany na podstawie [2]—[6].

2. ZASADY SKŁADANIA PLA

Istnieje wiele typów składania PLA możliwych do realizacji w zależności od zastosowanej technologii. W znanych pracach na ten temat zdefiniowano kilka typów składania, w tym wielokrotne składanie kolumnowe i wierszowe, uwzględniające możliwość realizowania więcej niż dwóch symbolicznych wierszy lub kolumn w jednym wierszu lub kolumnie fizycznej matrycy logicznej. W artykule zajmujemy się jednym szczególnym rodzajem składania PLA, a mianowicie prostym składaniem kolumnowym. Polega ono na składaniu kolumn wejściowych albo wyjściowych. Przykład prostego składania kolumnowego po-

kazany jest na rys. 5. Przedstawia on tablicę z rys. 3. która po złożeniu ma o połowę mniej kolumn i może być fizycznie zrealizowana na dwukrotnie mniejszej powierzchni.

Techniczna realizacja składania kolumn struktury PLA odbywa się na etapie projektu układu scalonego, który może być wykonany przez użytkownika lub zlecony wytwórcy. W szczególności dokonuje się tego w trakcie tworzenia tzw. topografii albo inaczej maski układu, która stanowi pewien wzór, nakładany w procesie technologicznym na standardową strukturę PLA. Istota składania kolumn struktury PLA, zrealizowanej w technologii typu MOS, polega na przecięciu w odpowiednim miejscu pary linii polikrzemu i dołączeniu do przeciwnych ich końców drugiego bufora wejściowego, realizującego drugą zmienną wejściową.

Badając warunki składania kolumn łatwo zauważyć, że jest konieczne, aby dwie kolumny matrycy PLA były rozłączne, to jest by punkty połączeń w każdej z nich występowały w różnych wierszach. W zaprogramowanej matrycy PLA połączenia danej szyby z wejściem prostym i jego negacją nigdy nie mogą występować jednocześnie, zatem można ją reprezentować tzw. macierzą połączeń określoną w następujący sposób: każdemu wejściu i wyjściu przypisuje się symboliczną kolumnę, która zawiera tylko symbole połączenia — „1” lub braku połączenia — „0”. Warunek rozłączności dwóch kolumn często może być niewystarczający do ich złożenia, bowiem w ogólności składanie jednej pary kolumn implikuje warunki ograniczające równoczesne składanie innych par kolumn. Składane ze sobą kolumny symboliczne są realizowane w jednej fizycznej linii matrycy logicznej, podzielonej tak, iż do części odpowiadającej jednej z nich jest doprowadzenie od dołu, do drugiej zaś od góry. Można przyjąć, że wystarczającym warunkiem składania kolumny i z kolumną j jest istnienie permutacji wierszy macierzy, w wyniku której każdy punkt połączeniowy w kolumnie i znalazłby się powyżej każdego punktu połączeniowego w kolumnie j . Zatem składanie dwóch kolumn implikuje warunki w postaci relacji między wierszami, określającej położenie punktów połączeniowych w jednej kolumnie względem punktów połączeniowych w drugiej. Składanie jednej pary kolumn może wykluczyć składanie innej pary, mimo, że są one rozłączne ze względu na punkty połączeń. Dla przykładu rozważmy fragment macierzy połączeń z rys. 6. W macierzy istnieją dwie pary kolumn

w \ k	Wejścia				Wyjścia	
	1	2	3	4	5	6
1	1	-	1	-	1	-
2	0	-	0	-	1	-
3	-	-	-	1	1	-
4	-	-	0	0	-	1
5	-	-	1	1	-	1
6	-	0	-	-	1	-

Rys. 6. Przykład konfliktu warunków składania kolumn

rozłącznych: 1,2 oraz 3,4. Składanie dowolnej z pary kolumn implikuje relację między wierszami, wykluczającą składanie tej drugiej pary. Istotnie, z jednej strony bowiem składanie kolumn 1 z 2 zakłada, że wiersz 2 jest nad wierszem 4, zaś z drugiej strony składanie 3 z 4 wymusza warunek przeciwny. Niemożliwe jest więc składanie tych par jednocześnie.

W przypadku istnienia wielu par kolumn rozłącznych badanie warunków wystarcza-

jących do jednoczesnego ich składania staje się złożonym i trudnym problemem. Konstrukcja efektywnych metod jego rozwiązywania jest przedmiotem niniejszego artykułu. Niezbędne jest wprowadzenie pomocniczego aparatu pojęciowego, który umożliwi ściśle i dostatecznie ogólne sformułowanie problemu. To sformułowanie z kolei stanowi punkt wyjścia do konstrukcji metod i algorytmów.

3. SFORMUŁOWANIE PROBLEMU OPTIMALNEGO SKŁADANIA PLA

Wprowadźmy teraz niezbędne oznaczenia i definicje podstawowych pojęć. Niech W będzie zbiorem wierszy, zaś K zbiorem kolumn matrycy logicznej. Przyjmując wcześniejszą umowę możemy zdefiniować macierz połączeń m o wymiarach $|K| \times |W|$: $m(k, w) = 1$, jeśli kolumna k ma połączenie z wierszem w , zaś $m(k, w) = 0$, gdy brak takiego połączenia.

Definicja 1 Niech $R \subseteq K \times W$ będzie relacją: $R = \{(k, w): m(k, w) = 1\}$. Relacja R wyznacza zbiór wszystkich punktów połączeń w matrycy logicznej opisanej macierzą połączeń m .

Definicja 2 Parę kolumn $\{k_1, k_2\}$ nazywamy rozłącznymi, jeśli $R(k_1) \cap R(k_2) = \emptyset$, to jest gdy punkty połączeń w kolumnach k_1 i k_2 są rozłącznymi podzbiorami relacji R . Określmy w zbiorze wierszy W relację porządku „ $<$ ”. Zapis $w_2 < w_1$ oznacza, że wiersz w_1 ma pozycję wyższą niż w_2 . Przyjmijmy następujący zapis: $R(k_1) < R(k_2)$, który oznacza, że dla $w_1 \in R(k_1)$, $w_j \in R(k_2)$: $w_i < w_j$.

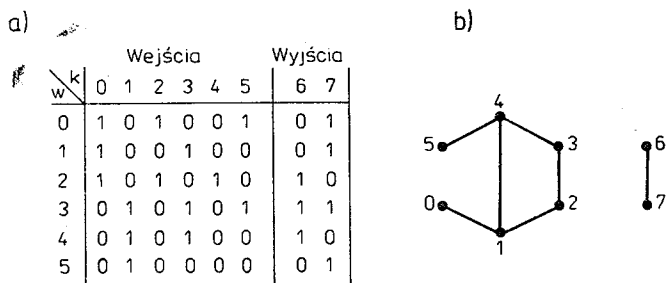
Definicja 3 Uporządkowaną parę $p = (k_1, k_2)$ nazywamy składalną, jeżeli $R(k_1) < R(k_2)$. Relacja $Q_p = R(k_1) \times R(k_2)$ określa warunki implikowane przez złożenie kolumn pary p .

Z definicji powyższej wynika, że kolumny pary składalnej są rozłączne.

Definicja 4 Zbiorem składalnym nazywamy zbiór 0 par składalnych, których elementy tworzą rozłączne 2-elementowe podzbiory kolumn.

Przykład

Dana jest macierz indywidualności przykładowej tablicy logicznej, jak na rys. 7. Badając macierz zauważymy następujące pary kolumn rozłącznych: (k_0, k_1) , (k_1, k_2) , (k_1, k_4) , (k_2, k_3) , (k_3, k_4) i (k_4, k_5) . Na rys. 7b pary te reprezentowane są krawędziami



Rys. 7. Przykład matrycy PLA: a) macierz połączeń; b) graf rozłączności kolumn

grafu. Z elementów tego zbioru można utworzyć zbiory składalne, w szczególności maksymalny zbiór składalny: $\{(k_0, k_1), (k_2, k_3), (k_4, k_5)\}$.

Niech O będzie składalnym zbiorem par, zaś $Q(O)$ relacją implikowaną przez O zdefiniowaną następująco:

$$Q(O) = \bigcup_{p \in O} Q_p \cup I_w,$$

gdzie I oznacza relację identyczności w zbiorze W .

Oznaczmy przez $Q'(O)$ tranzytywne domknięcie relacji $Q(O)$.

Definicja 5 Relacja „ $<$ ” określona w zbiorze W jest relacją częściowego porządku, jeśli:

- 1) $w < w$ dla każdego $w \in W$;
- 2) $w_1 < w_2$ oraz $w_2 < w_3 \rightarrow w_1 < w_3$ dla dowolnych $w_1, w_2, w_3 \in W$;
- 3) $w_1 < w_2$ oraz $w_2 < w_1 \rightarrow w_1 = w_2$ dla dowolnych $w_1, w_2 \in W$.

Definicja 6 Zbiór składalny O jest realizowalny, jeżeli relacja $Q'(O)$ jest relacją częściowego porządku.

Przykład

Dla macierzy indywidualności z rys. 7 zbiór składalny:

$$O = \{(k_1, k_2), (k_3, k_4)\}$$

implikuje następującą relację:

$$Q(O) = \{w_1 < w_3, w_1 < w_4, w_2 < w_3, w_2 < w_4, w_1 < w_2,$$

$$w_3 < w_2, w_4 < w_2, w_4 < w_3\} \cup \{w_i < w_i : i \in \{1, 2, 3, 4\}\}.$$

W tym przypadku relacja implikowana przez zbiór O jest przechodnia, czyli $Q(O) = Q'$, ale nie jest relacją częściowego porządku, zachodzi bowiem $w_2 < w_4$ oraz $w_4 < w_2$ lecz $w_2 \neq w_4$.

Posługując się zdefiniowanymi wyżej pojęciami możemy sformułować problem optymalnego składania matryc PLA — P_{OPT} — w następujący sposób:

Dana jest macierz połączeń m matrycy PLA; znaleźć realizowalny, składalny zbiór O^m o największej liczności.

4. TEORIO-GRAFOWY MODEL OPTYMALNEGO SKŁADANIA PLA

Problem optymalnego prostego składania kolumnowego jest obliczeniowo NP-zupełny. Oznacza to, że nie istnieje wyczerpujący algorytm, dla którego czas obliczeń jest ograniczony wielomianowo, a w konsekwencji już przy stosunkowo niedużym wymiarze zadania praktycznie nie można znaleźć najlepszego rozwiązania. Z tego powodu należy poszukiwać rozwiązań quasi-optymalnych, stosując metodę *heurystyczną* albo *quasi-wyczerpującą branch and bound*. W obu przypadkach wygodniej jest nie zmieniając istoty problemu przedstawić go w nieco innej postaci:

Dany jest zbiór par składalnych F . Należy zbadać czy istnieje realizacja składalnego zbioru O , który można otrzymać z F przez uporządkowanie jego par.

5. ALGORYTMY OPTIMALNEGO SKŁADANIA KOLUMN

5.1. METODA TEORIO-GRAFWA HEURYSTYCZNA

Wprowadźmy na początek pojęcie grafu mieszanego, który jest modelem dla konstrukcji algorytmu ([4]). Punktem wyjścia jest macierz połączeń matrycy PLA.

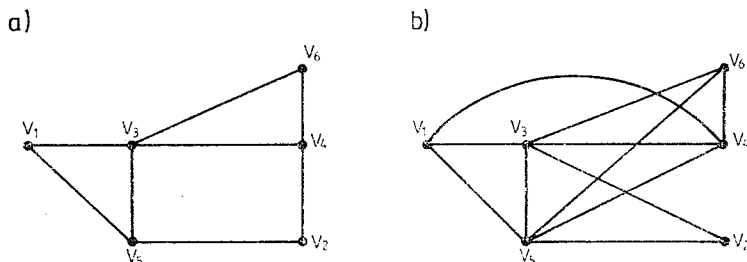
Niech $G = (V, E)$ będzie grafem, przy czym $V \cong K$ (tj. V i K są zbiorami izomorficznymi), V jest zbiorem wierzchołków, zaś $E = \{e = \{V_i, V_j\}: R(k_i) \cap R(k_j) \neq \emptyset\}$ jest zbiorem jego krawędzi nieskierowanych. Graf G nazywamy grafem przecięć kolumn.

Dla danego zbioru składalnego O i grafu przecięć kolumn $G(V, E)$ wprowadzimy pojęcie grafu mieszanego G_M , skojarzonego z O : $G_M = (V, E, A(O))$, w którym E jest zbiorem krawędzi nieskierowanych, zaś $A(O)$ zbiorem krawędzi skierowanych, odpowiadającym parom składalnym ze zbioru O :

$$A(O) = \{a = (v_i, v_j): (k_i, k_j) \in O\}.$$

Przykład

Na rys. 8 przedstawione są: graf przecięć kolumn G oraz graf mieszany G dla macierzy połączeń z rys. 7. Zbiór $A(O)$ zawiera parami rozłączne krawędzie, tj. takie, które nie posiadają wspólnych wierzchołków. Ponadto wierzchołki każdej skierowanej krawędzi z $A(O)$ są wspólne z krawędziami nieskierowanymi ze zbioru E . Zbiór A o tej własności nazywamy dopasowaniem grafu G . Ze względu na wzajemnie jednoznaczność odpowied-



Rys. 8. Graf macierzy połączeń: a) graf połączeń kolumn; b) graf mieszany

ność między wierzchołkami grafu a kolumnami matrycy PLA taka odpowiedniość zachodzi również między zbiorami O oraz $A(O)$. Dla danej matrycy PLA opisanej macierzą połączeń związek pomiędzy zbiorem składalnym O a zbiorem $A(O)$ grafu mieszanego G_M pozwala scharakteryzować w języku teorii grafów warunki realizowalności składalnego zbioru O . W tym celu posłużymy się poniższymi określeniami.

Definicja 7 Sekwencję wierzchołków $\alpha = (v_1, v_2, \dots, v_{2K-1}, v_{2K})$ mieszanego grafu $G(V, E, A)$ nazywamy ścieżką przemianą, jeśli $(v_{2K-1}, v_{2K}) \in A$ dla $k = 1, 2, \dots, K$ oraz $\{v_{2K}, v_{2K+1}\} \in E$ dla $K = 1, 2, \dots, K-1$

Definicja 8 Dany jest graf mieszany $G_M(V, E, A)$. Ścieżka przemianowa $\alpha = (v_1, v_2, \dots, v_{2K})$ tworzy cykl przemianowy, jeżeli $\{v_{2K}, v_1\} \in E$.

W przykładzie z rys. 8 sekwencja $\alpha = (5, 6, 3, 2)$ tworzy ścieżkę przemianową, która jest częścią cyklu, bowiem para $(2, 5) \in E$.

Istnienie cyklu przemiennego w grafie mieszanym oznacza, że nie są spełnione warunki realizowalności zbioru O . Uogólnienie tego wniosku zawiera się w poniższym twierdzeniu.

Twierdzenie 1 ([4]). Dany jest graf mieszany $G_M(G, A)$, gdzie $G = (V, E)$ jest grafem przecięć kolumn tablicy PLA, zaś A jest dopasowaniem w G . Składalny zbiór $O(A)$, odpowiadający dopasowaniu A , jest realizowalny wtedy i tylko wtedy, gdy w G_M nie istnieje żaden cykl przemienny (dowód w [4]).

Bezpośrednią konsekwencją twierdzenia 1 jest stwierdzenie, że problem optymalnego składania kolumn P_{OPT} jest równoważny następującemu problemowi teorio-grafowemu P_{GRAPH} : Dany jest graf $G(V, E)$ przecięć kolumn macierzy połączeń matrycy PLA. Znaleźć najliczniejszy zbiór A taki że:

- 1) A jest dopasowaniem w $G = (V, E)$;
- 2) $G_M = (V, E, A)$ nie zawiera cykli przemiennych.

Problem P_{GRAPH} , podobnie jak problem P_{OPT} jest problemem NP-zupełnym i obliczeniowo przekraczającym fizyczne możliwości dostępnych komputerów. Wyjściem kompromisowym jest algorytm heurystyczny, który pozwoliłby generować rozwiązania „dobre” jakościowo, nie wykluczając rozwiązań optymalnych. Proponowany algorytm pozwala znajdować maksymalne zbiory A tj. zbiory krawędzi zorientowanych, które mają własności 1) i 2) problemu P_{GRAPH} oraz nie istnieje zbiór $X \neq A$ i $|X| \geq |A|$ posiadający te własności. Działanie algorytmu polega na stopniowym budowaniu zbioru A przez powiększenie go każdorazowo o jedną krawędź. Taka strategia konstruowania rozwiązania nie wymaga powrotów i w rezultacie umożliwia istotne zmniejszenie złożoności obliczeniowej algorytmu. Ideę strategii przedstawimy w postaci szkicu algorytmu ZŁOŻENIE. Algorytm jest zapisany w pseudokodzie z zastosowaniem pomocniczych oznaczeń:

V — zbiór wierzchołków grafu $G(V, E)$;

A — zbiór krawędzi skierowanych, odpowiadający zbiorowi składalnemu $O(A)$;

VG, VD — zbiory wierzchołków grafu G , z których wybierane są odpowiednio początek i koniec krawędzi skierowanej, kandydującej do zbioru A ;

$SEL-G, SEL-D$ — procedury wyboru wierzchołków ze zbiorów VG, VD ;

$CYKL, TRANZ-DOM$ — procedury: wykrywania przemiennej ścieżki cyklicznej oraz wyznaczania tranzytywnego domknięcia zbioru krawędzi skierowanych.

Poniżej przedstawimy szkic głównego algorytmu.

Algorytm ZŁOŻENIE

BEGIN

$A := 0$; $VD := V$; $VG := V$; $TRANZ := 0$

PĘTLA: while ($VG \neq \emptyset$) do

begin 1

$v := SEL-G(VG)$; {wyznaczanie kolumny „wyższej”}

$VD1(V) := \{u \in VD : u \neq v, \{u, v\} \neq E\}$;

while $VD1(v) \neq \emptyset$ do

begin 2

$u := SEL-D(VD1(v))$; {wyznaczanie kolumny „niższej”}

if ($CYKL(v, u, TRANZ)$) {badanie istnienia naprzemiennej ścieżki

```

    then VD1(v): = VD1 - {u};      na podstawie domknięcia przechodniego}
    else
      begin 3
        TRANZ: = TRANZ = DOM(v, u, TRANZ); {uaktualnienie domk-
        A: = A ∪ {(u, v)};                nienia przechodniego}
        VG: = VG - {u, v}; VD: = VD - {u, v}; {usuwanie wierzchoł-
        go to PETLA                      ków wybranych kra-
      end 3                               wędzi)
    end 2
    VG: = VG - {v};
  end 1
  write A;
END.

```

Zastosowanie w powyższym algorytmie procedur SEL-G, SEL-D, CYKL oraz TRANZ-DOM wymagają osobnej dyskusji. Do zwięzłego i ścisłego ich omówienia niezbędne jest wprowadzenie dodatkowych oznaczeń i pojęć.

W procedurach SEL-G i SEL-D wyboru wierzchołków: początkowego i końcowego wierzchołka krawędzi kandydującej do zbioru A kryterium selekcyjnym jest parametr liczbowy, tzw. stopień wierzchołka w $G(V, E)$, definiowany jako liczba zawierających go krawędzi. W zależności od wybranej przesłanki heurystycznej stosuje się selekcję według minimum lub maksimum wartości tego parametru.

W grafie mieszanym $G_M(V, E, A)$ zbiór skierowanych krawędzi tworzy wzajemnie jednoznaczłą relację pomiędzy elementami zbiorów A_s i A_t , przy czym $A_s = pr_1 A = \{s: \exists t((s, t) \in A)\}$, $A_t = pr_2 A = \{t: \exists s((s, t) \in A)\}$. Zbiór A_s zawiera wierzchołki początkowe, zaś zbiór A_t końcowe wierzchołki krawędzi ze zbioru A.

Niech PRZ oznacza relację przyległości, która każdemu wierzchołkowi $v \in V$ w grafie $G(V, E)$ przyporządkowuje zbiór wierzchołków u , które z wierzchołkiem v tworzą krawędzi nieskierowane należące do zbioru E, tj.: $PRZ(v) = \{u, v\} \in E$. Niech $AP(v)$ oznacza zbiór ścieżek przemiennej o wierzchołku początkowym v . Funkcja TRANZ przyporządkowuje wierzchołkowi v zbiór

$$TRANZ(v) = \bigcup_{\alpha \in AP(v)} \{PRZ(u): u \in \alpha \cap A_t\}.$$

Elementami tego zbioru są wierzchołki przyległe do wierzchołków końcowych tych krawędzi zorientowanych, które należą do ścieżek przemiennej $\alpha \in AP(v)$. Zbiór $TRANZ(v)$ nazywamy domknięciem tranzytywnym wierzchołka v .

Stosując powyższe definicje i oznaczenia możemy zapisać procedury CYKL i TRANZ-DOM.

Procedura CYKL($v, u, TRANZ$)

```

begin
  CYKL := FALSE
  Su := {t ∈ A_m: t ∈ PRZ(u)};
  Fv := {t ∈ A_m: v ∈ TRANZ}

```

```

    if  $Su \cap Fv \neq \emptyset$  then RETURN
    else CYKL := TRUE;
  RETURN
end.

```

Procedura TRANZ-DOM(v, u, TRANZ)

```

begin
  TRANZ( $v$ ) := PRZ( $u$ );
  for  $t \in Su$  do
    begin
      TRANZ( $v$ ) := TRANZ( $v$ )  $v$  TRANZ( $t$ )
    end;
  for  $t \in Fv$  do
    begin
      TRANZ( $t$ ) := TRANZ( $t$ )  $v$  TRANZ( $v$ )
    end;
  RETURN;
end.

```

Pomijając jako zbyt złożoną analizę efektywności różnych strategii selekcji wierzchołków (SEL-D, SEL-G) można przyjąć dwa najlepiej uzasadnione warianty: 1) SEL-G według minimum stopnia, SEL-D według maksimum stopnia, oraz 2) obie selekcje według maksimum stopnia. Pierwszy wariant w większym stopniu uwzględnia przesłankę unikania tworzenia ścieżek cyklicznych. Drugi wariant akcentuje zasadę maksymalnego wykorzystania par składalnych do tworzenia zbioru składalnego.

5.2. METODA TEORIOGRAFOWA SYSTEMATYCZNA (*quasi-wyczerpująca*)

Dla wyjaśnienia istoty proponowanej systematycznej metody składania macryc PLA posłużymy się analogicznym modelem grafu jak w metodzie heurystycznej. Podstawą tej konstrukcji ([5]) jest pojęcie grafu relacji rozłączności kolumn $G_0 = (V, E_0)$, przy czym V — jest zbiorem wierzchołków, izomorficznym ze zbiorem kolumn macierzy połączeń, zaś E_0 reprezentuje zbiór par rozłącznych kolumn:

$$E_0 = \{\{v_i, v_j\}: R(k_i) \cap R(k_j) = \emptyset\}.$$

Do grafu G_0 możemy wprowadzić informacje o zbiorze składalnym O zastępując krawędź nieskierowaną odpowiednią krawędzią skierowaną, reprezentującą element zbioru składalnego, tj. parę składalną. W wyniku takiego przekształcenia otrzymujemy graf mieszany $G_{OM} = (V, E_{OM}, A)$, w którym występują dwa zbiory krawędzi: E_{OM} — zbiór krawędzi nieskierowanych oraz A — zbiór krawędzi skierowanych, izomorficzny ze zbiorem składalnym O . Posługując się pojęciami grafu mieszanego G_{OM} i ścieżki przemiennej w tym grafie możemy określić warunki realizowalności zbioru składalnego O i sformułować twierdzenie stanowiące podstawę proponowanego algorytmu.

Lemat 1: ([5]) warunkiem koniecznym realizowalności zbioru składalnego O jest istnienie w skojarzonym z nim grafie mieszanym $G_{OM} = (V, E_{OM}, A(O))$ ścieżki przemiennej

nej α , która zawiera krawędzi skierowane odpowiadające wszystkim elementom zbioru składalnego O .

Warunek podany w Lemacie 1 jest jedynie warunkiem koniecznym realizowalności zbioru składalnego. Jeśli skorzystamy z lematu i uwzględnimy def. 6, określającą warunek realizowalności zbioru składalnego, to możemy sformułować twierdzenie, z którego bezpośrednio wynika konstrukcja algorytmu.

Twierdzenie 2 Problem optymalnego składania kolumnowego macierzy m jest równoważny poszukiwaniu najdłuższej ścieżki przemiennej α w grafie mieszanym $G_{OM} = (V, E_{OM}, A)$ takiej, że zbiór par składalnych O odpowiadający zbiorowi $A \subseteq \alpha$ indukuje relację porządkującą wiersze macierzy m nie zawierającą cykli, tj. relację częściowego porządku.

Dowód jest oczywisty, gdyż znalezienie najdłuższej ścieżki przemiennej odpowiada znalezieniu zbioru składalnego o maksymalnej liczności, zaś dodatkowy warunek, aby indukowana przez ten zbiór relacja porządkująca wiersza nie zawierała cykli, zapewnia jego realizowalność.

5.2.1. OPIS ALGORYTMU

Wszystkie znane algorytmy składania PLA można podzielić na algorytmy heurystyczne i algorytmy wyczerpujące (m.in. typu *branch and bound*). Opisany w poprzednim punkcie heurystyczny algorytm składania umożliwia stopniowe rozbudowywanie realizowalnego zbioru par składalnych, których dołączenie nie niszczy częściowego porządku w zbiorze wierszy. Kolejne kroki polegają na selekcyjonowaniu nowych par i badaniu ich przydatności do realizacji. Element heurystyczny występuje na poziomie generowania i selekcji nowych par składalnych. Selekcja jest dokonywana ze względu na parametr, zwany stopniem wierzchołka. Uporządkowanie wierzchołków ze względu na wartość tego parametru stanowi heurystyczną przesłankę algorytmu. Ponieważ algorytm nie uwzględnia rozgałęzień i powrotów do rozwiązań wcześniejszych, jego efektywność jest zdeterminowana skutecznością procedury selekcji wierzchołków.

Algorytm oparty na strategii *branch and bound* w odróżnieniu od algorytmu heurystycznego rozszerza zbiór składalny wykorzystując wszystkie rozgałęzienia w drzewie poszukiwania rozwiązań. Jeżeli w procesie poszukiwania kolejne rozwiązanie nie jest lepsze, to cofamy się do najlepszego dotąd znalezionej. Teoretycznie optymalne rozwiązanie może być znalezione przez badanie wszystkich możliwości, jednak ze względu na ograniczenia czasowe nie jest to do przyjęcia. W praktycznych realizacjach tego algorytmu stosuje się pomocnicze funkcje ograniczenia górnego lub dolnego, których rolą jest „obcinanie” gałęzi na każdym poziomie drzewa poszukiwań. Wykonanie całości algorytmu jest zdeterminowane przez kryteria selekcji odpowiednio uporządkowanych par składalnych i oszacowanie górnego ograniczenia liczby par składalnych, którymi można uzupełnić zbiór z bieżącego odgałęzienia. Kryteria selekcyjne mogłyby uwzględniać również rozumowanie heurystyczne [Liu]. Algorytmy tego typu przy rozwijaniu drzewa poszukiwań uwzględniają powroty, z tego powodu ich złożoność czasowa jest trudna do oszacowania.

Odmienne niż we wspomnianych wyżej algorytmach składania podejście jest zastosowane w metodzie PPN (poszukiwanie pierwszego najlepszego rozwiązania), którą teraz

opiszemy. Metoda polega na poszukiwaniu zbioru składalnego o maksymalnej liczności, zawartego w najdłuższej ścieżce przemiennej łączącej dwa wierzchołki: źródłowy i docelowy, i badaniu jego realizowalności. Jeżeli wydzielony z danej ścieżki zbiór nie jest realizowalny, to badana jest następna najdłuższa ścieżka. Proces jest kontynuowany aż do znalezienia pierwszego realizowalnego zbioru składalnego. Szkic algorytmu jest opisany poniżej.

Algorytm główny

krok 1. Procedura inicjalizacji,

- 1.1. Wprowadzić dane (wczytać macierz m);
- 1.2. Wyznaczyć graf rozłączności kolumn G_0 ;
- 1.3. Znaleźć spójne komponenty grafu G_0 ;
- 1.4. Skonstruować najdłuższą ścieżkę przemiennej dla każdej pary wierzchołków grafu G_0 ;
- 1.5. Uporządkować skonstruowane ścieżki według ich długości,

krok 2. Wybrać najdłuższą ścieżkę i wydzielić z niej zbiór składalny.

krok 3. Dla otrzymanego zbioru składalnego wyznaczyć relację porządkującą.

krok 4. Zbadać, czy relacja porządkująca jest częściowym porządkiem. Jeżeli nie, to usunąć badaną ścieżkę z listy i przejść do kroku 2.

krok 5. Na podstawie relacji porządkującej uporządkować liniowo wiersze macierzy m i utworzyć macierz m złożonej matrycy PLA.

Algorytm składa się z dwóch części: procedury inicjalizacji i pętli głównej. W fazie procedury inicjalizacji wyznaczany jest graf rozłączności kolumn i jego spójne komponenty, a następnie dla każdej uporządkowanej pary wierzchołków konstruowana jest w tym grafie najdłuższa ścieżka. Utworzona lista ścieżek jest z kolei porządkowana według malejącej długości ścieżek. W fazie pętli głównej algorytmu z najdłuższej ścieżki wyznacza się zbiór składalny i bada jego realizowalność przeprowadzając test cykliczności. W przypadku, gdy zbiór składalny zawarty w ścieżce okaże się nierealizowalny, nie sprawdza się innych możliwych ścieżek pomiędzy tymi wierzchołkami, ale przechodzi się do badania najdłuższej ścieżki dla kolejnej pary wierzchołków. Taka metoda poszukiwań zawęża co prawda zbiór znajdowanych rozwiązań optymalnych, ale istotnie redukuje czas obliczeń.

Celem konstruowania najdłuższej ścieżki pomiędzy dwoma dowolnymi wierzchołkami jest znalezienie ścieżki zawierającej zbiór składalny o maksymalnej liczności. W opisanym algorytmie dla każdej z $n * (n - 1)$ par uporządkowanych wierzchołków tworzona jest jedna najdłuższa ścieżka. Ponieważ konstrukcja najdłuższej ścieżki jest procedurą stosunkowo czasochłonną, wprowadzamy modyfikację algorytmu, zwiększającą jego efektywność. Modyfikacja polega przede wszystkim na zastąpieniu poszukiwań najdłuższej ścieżki dla każdej z $n * (n - 1)$ par wierzchołków tworzeniem najdłuższej ścieżki o zadanym wierzchołku początkowym dla każdego z n wierzchołków. Ponadto na etapie testowania realizowalności zbioru składalnego zawartego w danej ścieżce w przypadku negatywnego wyniku testu, zamiast usuwania ścieżki z listy dokonuje się operacji skracania ścieżki o jedno ogniwo i ponownie umieszcza się ją na liście. Skrócona ścieżka będzie mogła być ponownie badana, gdy stanie się najdłuższą na liście ścieżek.

Zmodyfikowany algorytm główny

krok 1. Procedura inicjalizacji.

1.1. Wprowadzić dane (wczytać macierz m);

1.2. Wyznaczyć graf rozłączności kolumn G_0 ;

1.3. Skonstruować najdłuższe przemienne ścieżki wychodzące z każdego wierzchołka grafu i utworzyć ich listę;

1.4. Uporządkować listę ścieżek ze względu na ich długość,

krok 2. Wybrać najdłuższą ścieżkę z listy i wydzielić z niej zbiór składalny.

krok 3. Dla otrzymanego zbioru składalnego wyznaczyć relację porządkującą.

krok 4. Zbadać, czy relacja porządkująca jest częściowym porządkiem — jeśli nie to skrócić ścieżkę, umieścić ją z powrotem na liście i przejść do kroku 2.

krok 5. Na podstawie relacji porządkującej uporządkować liniowo wiersze macierzy m i utworzyć macierz m' złożonej matrycy PLA.

Efektywność implementacji opisanego algorytmu zależy w znacznym stopniu od metody wyznaczania najdłuższej ścieżki, ponieważ dowiedziono, że jest to problem NP-zupełny. Do rozwiązania tego zadania są niezbędne algorytmy heurystyczne lub probabilistyczne.

5.2.2. PROCEDURY KONSTRUKCJI NAJDŁUŻSZYCH ŚCIEŻEK

Przedstawimy teraz dwie probabilistyczne procedury wyznaczania najdłuższej ścieżki o zadanym wierzchołku początkowym. Procedury te są modyfikacją zaproponowanego przez Posa ([7]) algorytmu wyznaczania cyklu Hamiltona. Prowadzenie prostej (tj. bez cykli) ścieżki jest możliwe dzięki zastosowaniu tzw. dopuszczalnej transformacji ścieżki.

Dany jest nieskierowany graf $G = (V, E)$ i początkowy wierzchołek s ze zbioru V . Mechanizm tworzenia najdłuższej ścieżki w grafie, wychodzącej z s , polega na iteracyjnym badaniu jednego z wierzchołków przyległych do bieżącego wierzchołka końcowego, tj. ostatniego na dotychczas utworzonej ścieżce. Punktem wyjścia procedury jest trywialna ścieżka, która zawiera jeden wierzchołek. W kolejnych iteracjach są czynione próby jej wydłużenia aż do momentu wyczerpania wszystkich wierzchołków albo przekroczenia ustalonej liczby iteracji.

Opisane poniżej procedury konstrukcji najdłuższej ścieżki w G różnią się jedynie sposobem rekonstrukcji ścieżki, w której powstał cykl.

Procedura KS1

BEGIN

krok 1. Przyjąć ścieżkę P , składającą się z jednego wierzchołka s , za początkową oraz wierzchołek s za bieżący wierzchołek końcowy ścieżki, oznaczony symbolem k .

krok 2. Jeżeli liczba iteracji przekracza ustalone ograniczenie, to zakończyć procedurę; w przeciwnym razie wybrać losowo jeden z wierzchołków przyległych do bieżącego końca k i oznaczyć go przez i .

krok 3. W zależności od położenia wierzchołka i kontynuować tworzenie ścieżki w następujący sposób:

3.1. Jeżeli wierzchołek i nie należy do P , to przedłużyć P dodając krawędź

(k, i) i przyjąć i za bieżący wierzchołek końcowy. Jeżeli P zawiera wszystkie wierzchołki ze zbioru V , zakończyć procedurę; w przeciwnym razie wykonać krok 2.

3.2. Jeżeli wierzchołek i należy do P , to

i) jeżeli i jest wierzchołkiem poprzednim do bieżącego wierzchołka końcowego k , to sprawdzić czy jest on jedynym wierzchołkiem przyległym do k ; jeśli tak, to skrócić P przez usunięcie wierzchołka k i krawędzi (i, k) i przyjąć i za bieżący wierzchołek końcowy. Następnie przejść do kroku 2.

ii) Jeżeli wierzchołek i nie poprzedza bezpośrednio wierzchołka k , to zrekonstruować ścieżkę P stosując następującą transformację (rys. 9): Znaleźć wierzchołek, który bezpośrednio następuje po i , a następnie oznaczyć go przez m , po czym zastąpić krawędź (i, m) krawędzią (i, k) . Z kolei zmienić numerację kolejnych wierzchołków i kierunek krawędzi na odcinku ścieżki P od m do k . Podstawić m za bieżący wierzchołek końcowy. Przejść do kroku 2.

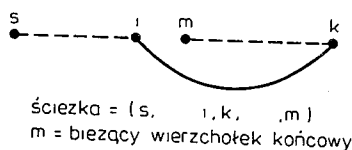
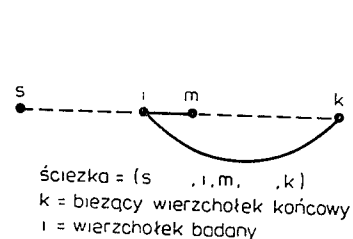
END

W przedstawionej metodzie konstrukcji ścieżki liczba iteracji jest ograniczona do $c \cdot n$, gdzie c jest dobrane empirycznie aby ograniczyć obszar wyszukiwania rozwiązań, zaś n jest liczbą wierzchołków grafu.

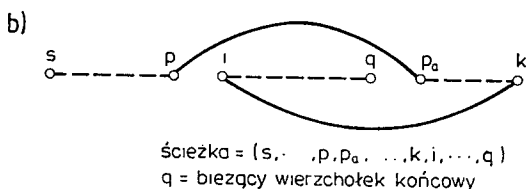
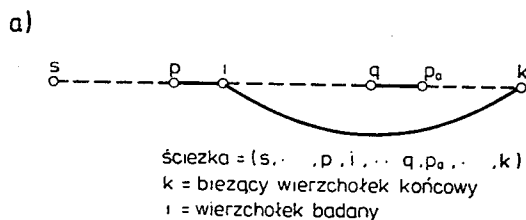
Procedura KS2

Procedura KS2 jest taka sama jak KS1 z wyjątkiem transformacji rekonstruującej ścieżkę (krok 3.2 ii). Opiszemy teraz mechanizm rekonstrukcji ścieżki zastosowany w procedurze KS2.

Niech s, k oraz i oznaczają odpowiednio: wierzchołek początkowy, bieżący wierzchołek końcowy i wierzchołek aktualnie badany, tak jak w procedurze KS1. Jeżeli i jest aktualnie



Rys. 9. Transformacja ścieżki z algorytmu KS1:
 a) ścieżka przed transformacją; b) ścieżka po transformacji

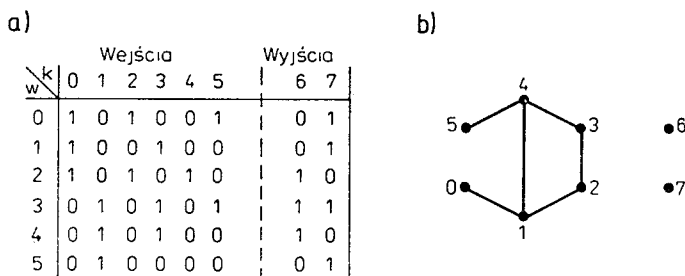


Rys. 10. Transformacja ścieżki z algorytmu KS2: a) ścieżka przed transformacją; b) ścieżka po transformacji

badanym wierzchołkiem, należącym do ścieżki P i niepoprzedzającym bezpośrednio k , to rekonstrukcję przeprowadza się w następujący sposób: dla wierzchołka p , bezpośrednio poprzedzającego badany wierzchołek i , znaleźć wierzchołki do niego przyległe i leżące na ścieżce P między wierzchołkiem i a wierzchołkiem k . Jeżeli nie istnieje żaden taki wierzchołek, to wykonać krok 2. Jeżeli istnieje więcej niż jeden taki wierzchołek, wybrać pierwszy lepszy i oznaczyć go przez p_a a wierzchołek poprzedzający go przez q . Zrekonstruować ścieżkę przez usunięcie z niej krawędzi (p, i) oraz (q, p_m) . Przyjęcie wierzchołka q za bieżący końcowy kończy rekonstrukcję ścieżki. Procedura KS2 ma tę zaletę, że podczas tworzenia ścieżki orientacja krawędzi pozostaje niezmienną. Proces rekonstrukcji jest pokazany na rys. 10.

5.3. PRZYKŁADY

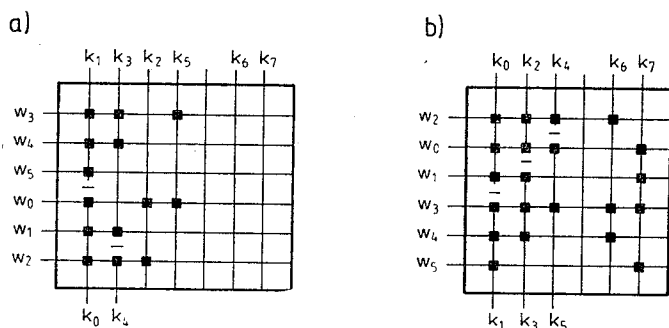
Dla poglądowego zilustrowania opisanych algorytmów i porównania ich efektywności posłużymy się dwoma prostymi przykładami matryc PLA. Rozważmy jako pierwszy przykład matrycę PLA1, której macierz połączeń i graf rozłącznych kolumn jest pokazany na rys. 11. Stosujemy algorytm heurystyczny z mieszaną strategią selekcji wierzchołków: pierwszy wierzchołek pary według maksimum stopnia oraz drugi, jeden z przyległych do niego, według minimum stopnia. Poszukując wierzchołka o maksymalnym stopniu mo-



Rys. 11. Przykładowa matryca PLA 1: a) macierz połączeń; b) graf rozłącznych kolumn

żemy wybrać k_1 lub k_4 . Przyjmijmy, że wybrany został wierzchołek k_1 . Dobierając do pary dla k_1 wybieramy przyległy do niego wierzchołek o minimum stopnia. Jedynym takim wierzchołkiem jest k_0 . W następnej iteracji poszukujemy kolejnej pary wierzchołków posługując się tym samym kryterium selekcji. Jako pierwszy wierzchołek nowej pary wybieramy k_4 i dobieramy do niego k_5 , który jest przyległym do k_4 wierzchołkiem o minimum stopnia. Niestety, para (k_4, k_5) nie jest zgodna z parą (k_1, k_0) ponieważ indukują one uporządkowanie cykliczne wierszy macierzy połączeń. Odrzucamy parę (k_4, k_5) i poszukujemy kolejnej pary postępując według tego samego schematu selekcji. W ten sposób wybieramy parę (k_4, k_3) . W rezultacie otrzymujemy zbiór składalny złożony z dwóch par: $\{(k_1, k_0), (k_4, k_3)\}$. Jest to najlepsze rozwiązanie generowane przez algorytm heurystyczny. Otrzymana w wyniku złożenia matryca PLA1⁻ jest przedstawiona na rys. 12a.

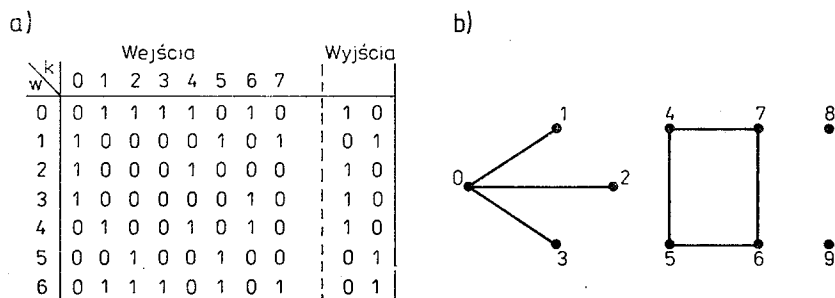
Z kolei stosując quasi-wyczerpujący algorytm PPN jako pierwszą propozycję rozwiązania otrzymujemy zbiór składalny złożony z trzech par: $\{(k_0, k_1), (k_2, k_3), (k_4, k_5)\}$, wydzielony z najdłuższej ścieżki: $(k_0, k_1, k_2, k_3, k_4, k_5)$. Zbiór ten jest realizowalny, a w wyniku złożenia otrzymujemy matrycę PLA1⁻ pokazaną na rys. 12b.



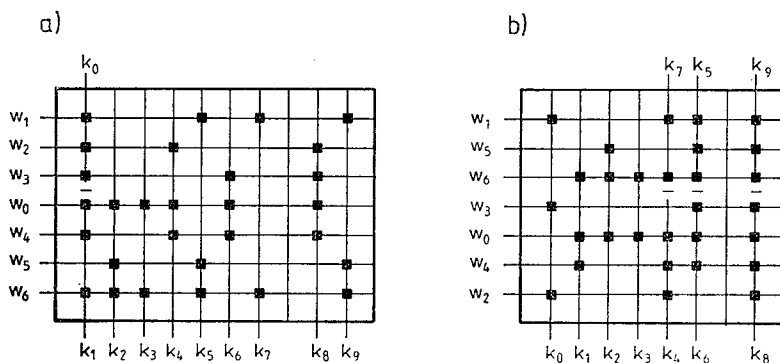
Rys. 12. Matryca PLA 1 po złożeniu przy użyciu a), b): a) algorytmu heurystycznego; b) algorytmu PPN

Rozważmy teraz inny przykład prostej matrycy logicznej — PLA2, której macierz połączeń i graf rozłącznych kolumn (jedynie dla kolumn wejściowych) jest pokazany na rys. 13.

Algorytm heurystyczny, wykorzystujący strategię selekcji wierzchołków jak w poprzednim przykładzie, generuje w pierwszej iteracji zbiór składalny, zawierający jedną



Rys. 13. Przykładowa matryca PLA 2: a) macierz połączeń; b) graf rozłącznych kolumn (dla zmiennych wejściowych)



Rys. 14. Matryca PLA 2 po złożeniu za pomocą: a) algorytmu heurystycznego; b) algorytmu PPN

parę. Zastosowanie strategii mini-maksowej również prowadzi do zbioru składalnego o jednej parze. W obu przypadkach w kolejnej iteracji żadna z par (k_0, k_1) , (k_0, k_2) oraz (k_0, k_3) nie da się połączyć w zbiór składalny z żadną inną parą. W wyniku złożenia otrzymujemy macrycę PLA²⁻, przedstawioną na rys. 14a.

Zastosowanie algorytmu PPN daje w konstrukcji najdłuższe ścieżki o długości 3. Jedną z nich jest ścieżka prowadząca od k_7 do k_6 : (k_7, k_4, k_5, k_6) . Wydzielony z tej ścieżki zbiór składalny $\{(k_7, k_4), (k_5, k_6)\}$ można zrealizować w postaci złożonej macrycy PLA²⁻⁻ jak na rys. 14b.

ZAKOŃCZENIE

W pracy przedstawiono podstawowe idee i algorytmy kompresji macryc PLA metodą składania. Przedstawiono dwie, istotnie różniące się w podejściu do foldingu, metody.

Pierwsza z nich, metoda heurystyczna, zakłada konstruowanie zbioru składalnego przez stopniowe dołączania kolejnych par składalnych i testowanie jego realizowalności. Skuteczność tego algorytmu jest zależna od procedury selekcji par składalnych, dołączanych do już utworzonego zbioru.

Druga, dokładna metoda (PPN), polega na konstruowaniu najdłuższych ścieżek w grafie rozłącznych kolumn, a następnie wydzielaniu z nich zbiorów składalnych i badaniu ich realizowalności. W odróżnieniu od pierwszej, metoda ta pozwala wyznaczać jako pierwsze znalezione rozwiązanie bardzo dobre, nie wykluczając najlepszych.

Ze względu na przeglądowy charakter pracy ograniczono się do opisu problemu i sformułowania jego modelu formalnego oraz szkicowego przedstawienia sposobów jego rozwiązania — wybranych metod i algorytmów. Z tej przyczyny pominięto w rozważaniach tak ważne zagadnienie, jak analiza efektywności i złożoności obliczeniowej opisanych szkicowo algorytmów, ich testowanie i porównywanie itp. (patrz [4], [5], [6]).

BIBLIOGRAFIA

1. A. V. Aho, J. E. Hopcroft, J. D. Ullman: *Data Structures and Algorithms*. Reading, MA: Addison—Wesley, 1983
2. M. N. Biswas: *Foldable Compatibility Matrix for the Folding of Progr. Logic Arrays*. Int. J. Electronics 1986, vol. 61, No 1
3. W. Grass: *A Depth First Search Branch and Bound Algorithms for Optimal Folding of PLA's*. 19-th DAC 1982
4. G. D. Hachtel, A. R. Newton, A. L. Sangiovanni-Vincentelli: *An Algorithm for Optimal PLA's Folding*. IEEE Yrans. on CAD. Vol. CAD-1, Jan. 1982
5. S. Y. Hwang, R. W. Dutton, T. Blank: *A Best First Search Algorithm for Optimal PLA's Folding*. IEEE Trans. on CAD, Vol. CAD-5, No 3, July 1986
6. J. L. Lewandowski, C. L. Liu: *A Branch and Bound Algorithm for Optimal Folding of PLA's*. 21st DAC 1984
7. L. Posa: *Hamiltonian circuits in random graphs*. Discrete Mathematics, vol. 14, 1976

K. JASIŃSKI

COMPACTION OF PROGRAMMABLE LOGIC ARRAYS BY FOLDING

Summary

A topological compaction method, known as folding, for reducing the physical area occupied by Programmable Logic Arrays (PLAs) is considered. A transformation of the folding formulation problem into a graph- and set — theoretical problem with more convenient solution procedures is given. Two different algorithms for PLA are presented. Several test problems and a comparison of the results are given.

K. JASIŃSKI

RÉDUCTION DES MATRICES LOGIQUES DU TYPE PLA
PAR MÉTHODE D'ASSEMBLAGE

Résumé

Le problème de la synthèse logique des systèmes, réalisés dans les Matrices Logiques Programmées (PLA), à l'étape de la conception de leur structure topologique, fait l'objet de l'article. On y a présenté la méthode d'optimisation de la matrice PLA, dite assemblage (en angl. folding), dont le but est de réduire la surface de silicium occupée par PLA. On a porté surtout attention à présenter la description théorique-graphique du problème et on a formulé les différentes méthodes de sa résolution. On a désigné deux modes de résolution — comme les plus importants à savoir: heuristique et systématique du type branch and bound. Pour les modes de résolution choisis on a présenté une description schématique des algorithmes et de la stratégie ainsi que quelques exemples de leur application.

K. JASIŃSKI

REDUKTION LOGISCHER PLA-MATRIZEN MITTELS ZUSAMMENFALTEN

Zusammenfassung

Im Beitrag wurden die Grundideen und Algorithmen der kommission-programmierten logischen PLA-Matrizen mittels Zusammenfalten (englisch: folding) dargestellt. Das Problem wurde formuliert, und dessen formales Modell angegeben, Lösungsstrategie skizziert, gewählte Methoden und Ergänzungen der Algorithmen nebst Schemen und einer Spezifikation von Grundprozeduren beschrieben. Insbesondere wurden zwei — wegen angewandter Strategie unterschiedliche — Methoden geschildert und miteinander verglichen: die angenäherte heuristische 'branch and bound' — Methode sowie die genaue systematische 'best—first—search' — Methode.

К. ЯСИНЬСКИ

РЕДУКЦИЯ ПРОГРАММИРУЕМЫХ ЛОГИЧЕСКИХ МАТРИЦ ТИПА ПЛА МЕТОДОМ СОСТАВЛЕНИЯ

Резюме

Представлены основные идеи и алгоритмы уменьшения программируемых логических матриц ПЛА методом составления. Поставлен вопрос и приведена его формальная модель. Предложены наброски решения, описаны избранные методы реализации алгоритмов вместе со схемами и спецификацией основных процедур. В особенности избраны и сравнены два различных метода: приблизительный и систематический.

Sprężone równania kinetyczne laserów złączowych

WŁODZIMIERZ NAKWASKI

Instytut Fizyki, Politechnika Łódzka

Otrzymano 1988.06.15

Autoryzowano do druku 1989.03.15

W niniejszej pracy przedstawiono matematyczny opis dynamiki pola elektromagnetycznego i pola nośników w laserach złączowych w postaci sprzężonych równań kinetycznych odpowiednio dla gęstości fotonów i koncentracji nośników. Równania te zostały sformułowane dla dwóch przypadków, tj. dla jednomodowego lasera szerokokontaktowego oraz dla wielomodowego planarnego lasera paskowego.

1. WPROWADZENIE

Lasery złączowe cechują się bardzo skomplikowanymi własnościami dynamicznymi. Dobitnie świadczy o tym kształt odpowiedzi optycznej lasera na prostokątny impuls prądu zasilania, wykazującej charakterystyczne opóźnienie oraz zwykle zanikające, głębokie oscylacje. Kształt ten spowodowany jest procesem akumulacji nośników w obszarze czynnym, w trakcie którego koncentracja wzbudzonych elektronów osiąga swą wartość progową, oraz wzajemnym oddziaływaniem pól promieniowania i nośników, mającym charakter rezonansowy.

Zachowanie obu pól opisuje się zwykle za pomocą charakterystycznych stałych czasowych, tj. czasu życia nośników oraz czasu życia fotonów. Energia skupiona w rezonatorze i obszarze czynnym lasera złączowego jest więc przekazywana pomiędzy oboma polami z częstotliwością zależną od obu powyższych stałych czasowych. Obrazu dopełniają fluktuacje czasowe i przestrzenne koncentracji nośników, gęstości fotonów i temperatury, mogące wywołać efekty rezonansowe często trudne do przewidzenia.

Własności dynamiczne laserów złączowych nabrały w ostatnich latach szczególnego znaczenia. Związane jest to z powszechnym stosowaniem tych przyrządów w nowoczesnych układach światłowodowej telekomunikacji optycznej [1–11] o ogromnych szybkościach przekazywania informacji. Przekraczają już one obecnie wartość 1 Gbit/sec [12–15], a w najbliższej przyszłości mogą osiągnąć wartości sięgające nawet 30 Gbit/sec [16]. Zlikwidowanie opóźnienia impulsu laserowego, efektywne tłumienie oscylacji relaksacyjnych oraz ograniczenie do minimum szumów we wiązce promieniowania stały się więc obecnie zagadnieniami pierwszorzędnej wagi.

Własności dynamiczne lasera złączowego opisuje się zwykle za pomocą równań kinetycznych. Ze względu na kwantowy charakter oddziaływań pomiędzy poszczególnymi modami oraz statystyczne własności kwantowych fluktuacji pól fotonów i nośników ładunku, powinny to być równania kinetyczne sformułowane zgodnie z zasadami mechaniki kwantowej [17]. Jednakże zwykle, gdy interesują nas głównie czasowe zmiany koncentracji nośników i gęstości fotonów, korzysta się ze znacznie prostszych, klasycznych równań kinetycznych [18, 19].

2. LASER SZEROKOKONTAKTOWY

Rozpatrzmy wpieryw jednomodowy laser szerokokontaktowy. Równania kinetyczne dla koncentracji elektronów $n(t)$ i gęstości fotonów $N(t)$ można wówczas przedstawić w następującej postaci [20, 21]:

$$\frac{dn}{dt} = \frac{j}{ed_A} - \frac{R_{SP}}{a_i} - \frac{c}{n_G} g \Gamma N \quad (1)$$

$$\frac{dN}{dt} = \frac{c}{n_G} g \Gamma N - \frac{N}{t_{PH}} + b \Gamma R_{SP}, \quad (2)$$

w której oznaczenia są następujące: j jest gęstością prądu wpływającego do obszaru czynnego, d_A — grubością tego obszaru, e — ładunkiem elementarnym, R_{SP} — całkowitą szybkością emisji spontanicznej, a_i — wewnętrzną sprawnością kwantową, c — szybkością światła w próżni, n_G — grupowym współczynnikiem załamania, g — wzmacnieniem na jednostkę długości, Γ — współczynnikiem wypełnienia, t_{PH} — czasem życia fotonów, a b — współczynnikiem emisji spontanicznej.

Pierwszy wyraz po prawej stronie równania (1) opisuje wzrost koncentracji elektronów w obszarze czynnym w wyniku ich wstrzykiwania do tego obszaru przez złącze $p-n$, natomiast dwa następne wyrazy uwzględniają zmniejszanie się tej koncentracji z powodu odpowiednio rekombinacji spontanicznej i niepromienistej oraz rekombinacji stymulowanej. Z kolei po prawej stronie równania (2), pierwszy wyraz opisuje wzrost gęstości fotonów skupionych w modzie promieniowania w wyniku emisji stymulowanej, drugi wyraz uwzględnia straty krawędziowe związane z emisją części promieniowania na zewnątrz przyrządu oraz straty wewnętrzne, tj. straty wewnątrz obszaru czynnego (z wyłączeniem absorpcji międzypasmowej, uwzględnionej już w pierwszym wyrazie), a trzeci wyraz symbolizuje udział części promieniowania spontanicznego (wyrażonej współczynnikiem b) w modzie promieniowania laserowego.

Postać wyrazu opisującego emisję spontaniczną zależy w znacznym stopniu od domieszkowania obszaru czynnego, tj.

$$\frac{R_{SP}}{a_i} = \begin{cases} \frac{B_R n^2}{a_i} & (\text{ślabe domieszkowanie}) \\ n \left(\frac{1}{t_R} + \frac{1}{t_{NR}} \right) = \frac{n}{a_i t_R} = \frac{n}{t_S} & (\text{silne domieszkowanie}) \end{cases} \quad (3)$$

gdzie B_R jest współczynnikiem rekombinacji promienistej, a t_R , t_{NR} i t_S są odpowiednio

promienistym czasem życia, niepromienistym czasem życia i całkowitym czasem życia nośników. Można też w miejsce zależności (3) stosować ogólny wzór [20]:

$$\frac{R_{SP}}{a_i} = \frac{n}{t_{EF}}, \quad (4)$$

w którym zastępczy czas życia t_{EF} elektronu zdefiniowano w następującej postaci

$$\frac{1}{t_{EF}} = \frac{1}{a_i} \left(\frac{\partial R_{SP}}{\partial n} \right)_{TH}, \quad (5)$$

gdzie indeks „TH” oznacza, że wielkość w nawiasach należy określić dla progu akcji laserowej.

Grupowy współczynnik załamania n_G można zapisać następująco [22]:

$$n_G = n_R - \lambda(dn_R/d\lambda), \quad (6)$$

gdzie n_R jest współczynnikiem załamania, a λ — długością fali promieniowania.

Wzmocnienie lokalne g wiąże się zwykle z koncentracją nośników mniejszościowych n wstrzykniętych do obszaru czynnego w sposób następujący

$$g = a_G n - a_0, \quad (7)$$

gdzie a_G i a_0 są parametrami materiałowymi i mogą mieć inne wartości dla laserów różnych typów. Niekiedy stosuje się w miejsce wzoru (7) inne zależności, np.

$$g = a_c n^C \quad (8)$$

gdzie $C = 3$ [19], lub $C = 2,5$ [23].

Pamiętając, że wartość współczynnika b jest bardzo niska ($b = 10^{-5} - 10^{-4}$ [20, 21, 24—26]), można pomijając ostatni wyraz we wzorze (2) wyznaczyć dla stanu ustalonego wartość czasu życia fotonu [18]

$$\frac{1}{t_{PH}} = \frac{c}{n_G} \Gamma g_{TH} = \frac{c}{n_G} (\alpha_i + \alpha_{end}) \quad j > j_{TH}, \quad (9)$$

gdzie α_i i α_{end} są współczynnikami odpowiednio strat wewnętrznych i strat krawędziowych. Wówczas przybliżone rozwiązania stacjonarne równań (1—2) przedstawiają się następująco [19, 26]:

$$N_{ST} = (t_{PH}/t_s) n_{TH} [(j/j_{TH}) - 1], \quad (10)$$

$$j_{TH} = (ed_A n_{TH})/t_s, \quad (11)$$

gdzie j_{TH} jest gęstością prądu progowego, $g_{TH} = g(n = n_{TH})$, a n_{TH} jest progową wartością koncentracji elektronów w obszarze czynnym lasera.

Znacznie jednak dokładniejszy jest wzór uwzględniający udział promieniowania spontanicznego w modzie laserowym [18]:

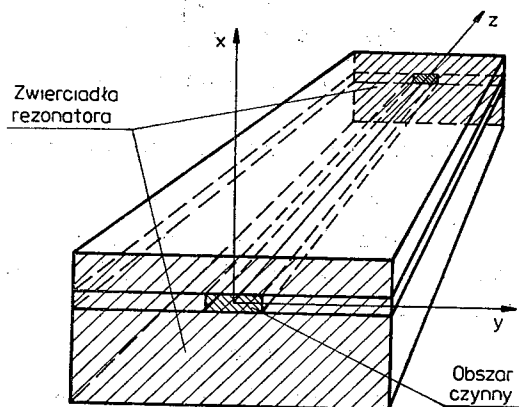
$$N_{ST} = \frac{b R_{SP}}{(c/n_G)(g_{TH} - g)}. \quad (12)$$

Małe zmiany wzmocnienia lokalnego w pobliżu jego wartości progowej powodują, jak widać, znaczne zmiany gęstości fotonów w laserującym modzie, tym mniejsze jednakże, im większy jest współczynnik b [25, 27].

3. LASER PASKOWY

Rozpatrzmy teraz znacznie ogólniejszy przypadek, tj. planarny laser o geometrii paskowej, zwany w dalszej części pracy w skrócie laserem paskowym.

Układ współrzędnych dobieramy w taki sposób (patrz rys.), aby oś z pokrywała się z osią podłużną lasera, oś x — z jego osią poprzeczną, a oś y leżała w obszarze czynnym



Rys. Struktura analizowanego planarnego lasera paskowego

równoległe do płaszczyzny złącza i jednocześnie równoległe do płaszczyzn obu zwierciadeł rezonatora. Wówczas równania kinetyczne, uwzględniające już wielomodowość promieniowania, dają się zapisać następująco [21]:

$$\frac{\partial n(y, t)}{\partial t} = \frac{j(y, t)}{ed_A} - \frac{n(y, t)}{t_s} + D_A \frac{\partial^2 n(y, t)}{\partial y^2} - \frac{c}{n_G} \Gamma g[n(y, t)] \sum_m I_m(y, t) N_m(t) \quad (13)$$

$$\frac{dN_m(t)}{dt} = \frac{c}{n_G} [G_m(t) - \alpha_s - \alpha_{end}] N_m(t) + b_m \frac{n_{AV, m}}{t_s}, \quad (14)$$

gdzie m jest rzędem modu bocznego, tj. ilością miejsc zerowych w rozkładzie natężenia promieniowania danego modu wzdłuż osi y (Rodzaje modów, tj. mody podłużne, różniące się liczbą połówek fali danego modu wzdłuż osi z , mody poprzeczne, o różnej ilości miejsc zerowych rozkładu natężenia promieniowania wzdłuż osi x , oraz powyższe mody boczne, zdefiniowano w monografii [28]).

α_s symbolizuje straty promieniowania związane z jego rozpraszaniem na niejednorodnościach materiału obszaru czynnego oraz na zaburzeniach grubości warstwy czynnej.

W pierwszym wyrazie po prawej stronie równania (13) uwzględniono rozptyw prądu, pominięto natomiast wpływ ewentualnej zmiany grubości obszaru czynnego d_A . Równania kinetyczne uwzględniające tę różnicę można znaleźć na przykład w artykułach [29—31] i w pracach tam cytowanych.

Trzeci wyraz równania (13) opisuje dyfuzję nośników (zarówno elektronów, jak i dziur) w obszarze czynnym. Ambipolarna stała dyfuzji D_A ma w tym przypadku postać następującą [32, 33]:

$$D_A = \frac{\mu_H \mu_E}{\mu_H + \mu_E} \left(\frac{D_H}{\mu_H} + \frac{D_E}{\mu_E} \right), \quad (15)$$

gdzie μ_H i μ_E są ruchliwościami, a D_H i D_E — stałymi dyfuzji, odpowiednio dziur i elektronów. Korzystając z danych opublikowanych w pracy [34], można w przypadku laserów GaAs/(AlGa)As uprościć powyższy wzór do postaci [33]:

$$D_A = 3\mu_H(k_B T/e), \quad (16)$$

gdzie k_B jest stałą Boltzmanna, a T — temperaturą w skali bezwzględnej.

Przez $I_m(y, t)$ oznaczono znormalizowany rozkład natężenia promieniowania skupionego w modzie bocznym rzędu m , tj.

$$I_m(y, t) = \frac{|E_m(y, t)|^2}{(1/S) \int_{-\infty}^{\infty} |E_m(y, t)|^2 dy}, \quad (17)$$

gdzie S jest szerokością paskowego obszaru czynnego, a E_m — rozwiązaniem równania falowego (patrz [32]) dla modu bocznego rzędu m .

Wzmocnienie modowe tego modu, tj. G_m , można wyznaczyć z ogólnego związku:

$$G_m(t) = 2Im\{\gamma_m(t)\}, \quad (18)$$

gdzie γ_m jest zespoloną stałą propagacji (patrz [35]), lub, znając współczynniki strat promieniowania poza obszarem czynnym α_{out} i w obszarze czynnym α_A , stosując wzór [21]:

$$G_m(t) = \frac{1}{n_{R,ef}} \left\{ \frac{n_{R2}}{S} \int_{-\infty}^{\infty} \Gamma(g[n(y, t) - \alpha_A n(y, t)]) I_m(y, t) dy - n_{R1}(1 - \Gamma) \alpha_{out} \right\}, \quad (19)$$

gdzie $n_{R,ef}$ jest efektywnym współczynnikiem załamania (patrz [35]), a n_{R1} i n_{R2} są współczynnikami załamania materiałów odpowiednio warstwy ograniczającej i warstwy czynnej.

Straty promieniowania w obszarze czynnym są głównie spowodowane przez absorpcję na swobodnych nośnikach, a ogólnie można je wyrazić w następującej postaci:

$$\alpha_A(n) = a_A n + b_A, \quad (20)$$

gdzie parametry strat dla laserów GaAs/(AlGa)As są w przybliżeniu równe [36, 37]:

$$a_A = 10^{-21} \text{ m}^2$$

$$b_A = 1,5 \cdot 10^3 \text{ m}^{-1}$$

Uśrednioną z punktu widzenia modu bocznego rzędu m koncentrację elektronów $n_{AV,m}$ definiuje się następująco [21]:

$$n_{AV,m} = (1/S) \int_{-\infty}^{\infty} n(y, t) I_m(y, t) dy, \quad (21)$$

a moc promieniowania P_m skupionego w modzie m i emitowanego na zewnątrz rezonatora przez jedno z jego zwierciadeł jest równa:

$$P_m(t) = 0,5 \hbar \omega \alpha_{end}(c/n_G)(SLd_A)N_m(t) \quad (22)$$

gdzie ω jest pulsacją promieniowania, a $SLd_A = V_A$ przedstawia objętość obszaru czynnego.

Jeżeli chcemy w naszej analizie uwzględnić współzawodnictwo modów podłużnych, należy rozpatrzyć różne mody podłużne rzędu p o różnych długościach fali promieniowania:

$$\lambda_P = \lambda_0 + p\Delta\lambda_{FP} \quad \text{z} \quad \Delta\lambda_{FP} = (\lambda_0)^2/2n_G L. \quad (23) (24)$$

Wówczas wzmocnienie lokalne g będzie zależne od rzędu p modu podłużnego [20]:

$$g_P = g(\lambda_P) = a_G n \{1 - [2\{\lambda(t) - \lambda_P\}/\Delta\lambda_G]^2\} - a_0 \quad (25)$$

i w ostatnim wyrazie we wzorze (13) należy sumowanie przeprowadzić nie tylko po wszystkich modach bocznych m , ale i po wszystkich modach podłużnych p . We wzorze (25) $\Delta\lambda_G$ jest szerokością spektralną krzywej wzmocnienia, przez $\lambda(t)$ oznaczono długość fali promieniowania odpowiadającej maksimum krzywej wzmocnienia [20]:

$$\lambda(t) = \lambda_0 + k\{1 - [n(t)/n_{TH}]\} + \Delta\lambda \quad (26)$$

$\Delta\lambda$ jest wyznaczoną dla progu akcji laserowej odległością na charakterystyce widmowej maksimum krzywej wzmocnienia od wierzchołka odpowiadającego najbliższemu modowi Febry'ego-Perota, λ_0 — długością fali promieniowania skupionego w tym modzie, a k jest stałą wiążącą przesunięcie maksimum krzywej wzmocnienia z koncentracją nośników mniejszościowych (dla laserów $\text{InGa}(\text{AsP})/\text{InP}$ emitujących promieniowane o długości fali $1.55 \mu\text{m}$ — $k = 106.4 \text{ nm}$ [20]).

4. PODSUMOWANIE

W pracy przedstawiono układ dwóch równań kinetycznych stanowiących podstawę do opisu dynamicznych własności lasera złączowego poprzez modelowanie czasowych zmian gęstości fotonów w jego rezonatorze oraz koncentracji nośników mniejszościowych (elektronów) w jego obszarze czynnym. Ze względu na wzajemne oddziaływanie obu pól, tj. pola elektromagnetycznego i pola nośników, powyższy układ równań jest układem sprzężonym.

Równania kinetyczne zostały sformułowane dla dwóch krańcowych przypadków — jednomodowego lasera szerokokontaktowego oraz wielomodowego planarnego lasera paskowego. Uwzględniono współzawodnictwo zarówno różnych modów bocznych, jak i różnych modów podłużnych.

W następnych pracach zostaną przedstawione rozwiązania powyższych równań kinetycznych, tzn. zajmiemy się tzw. efektami włączeniowymi w laserach złączowych, ich szumem promienistym oraz metodami poprawy selektywności spektralnej tych przyrządów.

BIBLIOGRAFIA

1. D. Botez, G.J. Herskowitz: *Components for optical communications systems: a review*. Proc. IEEE, vol. 68, No. 6, 1980, 689—731
2. M. Nakamura, S. Tsuji: *Single-mode semiconductor injection lasers for optical fiber communications (invited paper)*. IEEE J. Quantum Electron., vol. QE-17, No 6, 1981, 994—1005

3. D. H. Newman, M. R. Matthews, I. Garrett: *Sources for optical fibre communications*. Telecommunication J., vol. 48, No 11, 1981, 673—680
4. G. H. Olsen, M. Ettenberg: *Long-wavelength semiconductor diode lasers for optical communications*. RCA Engineer, vol. 27, No 4, 1982, 38—44
5. D. J. Channin: *Emitters for fiberoptic communications*, Laser Focus Magazine, vol. 18, No 11, 1982, 105—113
6. K. Kobayashi: *Semiconductor optical devices for intercity optical fiber transmission systems*. IEEE Communication Society, „Global Telecommunications Conference”, November 26—29, 1984, Atlanta, GA, USA, 1152—1156
7. D. J. Channin: *Optoelectronic performance issues in fiber — optic communications*. RCA Review, vol. 46, No 6, 1985, 240—269
8. J. C. Dymert: *III — V compound semiconductors for communications*, Canadian J. Physics, vol. 63, No 6, 1985, 651—663
9. D. H. Newman, S. Ritchie: *Sources and detectors for optical fibre communications applications: the first 20 years*. IEE Proc. — J (Optoelectronics), vol. 133, No 3, 1986, 213—229
10. S. Yamamoto, H. Sakaguchi: Nunokawa, Y. Iwamoto: *Studies of design for 1.55 μm optical fiber submarine cable systems*. J. Optical Communications, vol. 7, No 1, 1986, 3—11
11. R. W. Dixon, N. K. Dutta: *Lightwave device technology*, AT&T Techn. J., vol. 66, No 1, 1987, 73—83
12. S. Tell, S. T. Eng: *Optical fiber communication experiment at 5 Gbit/sec*, Appl. Opt., vol. 20, 1981, 3853—3858
13. R. A. Linke, B. L. Kasper, J. C. Campbell, A. G. Dentai, I. P. Kaminow: *120 km lightwave transmission experiment at 1 Gb/s using a new long-wavelength avalanche photodetector*. Electron. Lett., vol. 20, 1984, 498—499
14. B. L. Kasper, R. A. Linke, K. L. Walker, L. G. Cohen, T. L. Koch, T. J. Bridges, E. C. Burkhardt, R. A. Logan, R. W. Dawson, J. C. Campbell: *A 130 km transmission experiment at 2 Gb/s using silica-core fiber and a vapor phase transported DFB laser*. Proc. 10th Euro. Conf. Opt. Fiber Commun., Stuttgart, RFN, 1984, PD-6
15. S. K. Korotky, G. Eisenstein, A. H. Gnauck, B. L. Kasper, J. J. Veselka, R. C. Alferness, L. L. Bulil, C. A. Burrus, T. C. D. Huo, L. W. Stulz, K. C. Nelson, L. G. Cohen, R. W. Dawson, J. C. Campbell: *4 Gb/s transmission experiment over 117 km of optical fiber using a Ti:LiNbO₃ external modulator*. Conf. Optical Fiber Commun., San Diego, CA, USA, February 1985
16. R. C. Goodfellow, B. T. Debney, G. J. Rees, J. Buus: *Optoelectronic components for multigigabit systems*, IEEE Trans. Electron Devices, vol. ED-32, No 12, 1985, 2562—2571
17. H. Haug: *Quantum-mechanical rate equations in semiconductor lasers*, Physical Review, vol. 184, No 2, 1969, 338—348
18. M. J. Adams, P. T. Landsberg: *The theory of the injection laser*. [in] GaAs Lasers, C. H. Gooch (Ed.), London, Wiley, 1969
19. M. J. Adams: *Rate equations and transient phenomena in semiconductor lasers*, Opto — electronics, vol. 5, 1973, 201—215
20. M. J. Adams, M. Osiński: *Longitudinal mode competition in semiconductor lasers. Rate equations revisited*. IEE Proc. — I (Solid State and Electron Devices), vol. 129, No 6, 1982, 271—274
21. J. Buus: *Models of the static and dynamic behaviour of strope geometry lasers*. IEEE J. Quantum Electron., vol. QE-19, No 6, 1983, 953—960
22. J. Buus, M. J. Adams: *Phase and group indices for double heterostructure lasers*. Solid — State and Electron Devices, vol. 3, No 6, 1979, 189—195
23. C. J. Hwang: *Properties of spontaneous and stimulated emission in GaAs junction lasers. II. Temperature dependence of threshold current and excitation dependence of superradiance spectra*. Physical Review B, vol. 2, No 10, 1970, 4126—4134
24. K. Petermann: *Calculated spontaneous emission factor for double-heterostructure injection lasers with gain-induced waveguiding*. IEEE J. Quantum Electron., vol. QE-15, No 7, 1979, 566—570

25. G. Arnold, P. Russer, K. Petermann: *Modulation of laser diodes*. [in] *Semiconductor Devices for Optical Communication*, H. Kressel (Ed.), Berlin, Springer 1980, 213—242
26. P. Russer, G. Arnold: *Direct modulation of semiconductor injection lasers*. IEEE Trans. Microwave Theory and Techniques, vol. MTT-30, No 11, 1982, 1809—1821
27. P. M. Boers, M. T. Vlaardingerbroek, M. Danielsen: *Dynamic behaviour of semiconductor lasers*. Electron. Lett., vol. 11, No 10, 1975, 206—208
28. B. Mroziewicz, M. Bugajski, W. Nakwaski: *Lasery Półprzewodnikowe*. Warszawa: PWN, 1985, s. 140
29. W. Nakwaski: *Zasady modelowania zjawisk fizycznych w laserach złączowych. I. Modelowanie zjawisk elektrycznych*. Rozprawy Elektrotechniczne, 1988, vol. 34, z. 3, 959—964
30. W. Nakwaski: *Zasady modelowania zjawisk fizycznych w laserach złączowych. II. Modelowanie zjawisk optycznych*. Rozprawy Elektrotechniczne, 1988, vol. 34, z. 3, 965—970
31. W. Nakwaski: *Zasady modelowania zjawisk fizycznych w laserach złączowych. III. Porównanie modeli*. Rozprawy Elektrotechniczne, 1988, vol. 34, z. 3, 971—981
32. H. Statz, C. L. Tang, J. M. Lavine: *Spectral output of semiconductor lasers*. J. Appl. Phys. vol. 35, No 9, 1964, 2581—2585
33. J. Buus, M. Danielsen: *Carrier diffusion and higher order transversal modes in spectral dynamics of the semiconductor laser*, IEEE J. Quantum Electron., vol. QE-13, No 8, 1977, 669—674
34. K. A. Shore, M. J. Adams: *The effects of carrier degeneracy on transport properties of the double-heterostructure injection laser*. Appl. Phys., vol. 9, 1976, 161—164
35. W. Nakwaski: *Efekt falowodowy w laserach paskowych*. Rozprawy Elektrotechniczne, 1989
36. E. Pinkas, B. I. Miller, I. Hayashi, P. W. Foy: *Additional data on the effect of doping on the lasing characteristic of GaAs — Al_xGa_{1-x}As double-heterostructure lasers*. IEEE J. Quantum Electron., vol. QE-9, No 2, 1973, 281—282
37. J. Buus: *A theoretical investigation of semiconductor lasers*. PhD Thesis, Electromagnetics Institute, Technical University of Denmark, Lyngby 1979, s. 69

W. NAKWASKI

COUPLED RATE EQUATIONS FOR LASER DIODES

Summary

A mathematical description of the dynamics of the electromagnetic field as well as of the carrier field in laser diodes is presented in the form of coupled rate equations for photon density and carrier concentration, respectively. The equations are formulated for two cases, i.e. for a single-mode broad-contact laser diode and for a many-mode planar stripe-geometry laser diode.

W. NAKWASKI

EQUATIONS CINÉTIQUES ACCOUPLÉES DES LASERS CONJUGUÉS

Résumé

Dans l'étude on a présenté la description mathématique de la dynamique d'un champ électromagnétique et celle d'un champ des porteurs dans les lasers conjugués sous forme d'équations cinétiques accouplées convenablement à la densité des photons et à la concentration des porteurs. Ces équations ont été formulées pour deux cas, à savoir pour un laser unimodulaire à contacts larges et pour un laser polymodulaire longitudinal à procédé planar.

W. NAKWASKI

KONJUNGIERTE GESCHWINDIGKEITSGLEICHUNGEN FÜR LASERDIODEN

Zusammenfassung

Im vorliegenden Beitrag wurde eine mathematische Beschreibung der Dynamik eines elektromechanischen Feldes und eines Elektronenfeldes in Laserdioden in Form von konjugierten Geschwindigkeitsgleichungen der Photonendichtheit und der Elektronenkonzentration entsprechend dargestellt. Diese Gleichungen wurden für zwei Fälle formuliert, d.h. für eine breitkontaktige Einmoduslaserdiode und für eine Mehrmodusplanarlaserdioden mit Streifenstruktur.

B. НАКВАСКИ

СОПРЯЖЁННЫЕ КИНЕТИЧЕСКИЕ УРАВНЕНИЯ ЛАЗЕРНЫХ ДИОДОВ

Резюме

Представлено математическое описание динамики электромагнитного поля и поля электронов в лазерных диодах в виде сопряжённых кинетических уравнений соответственно для плотности фотонов и концентрации электронов. Эти уравнения сформулированы для двух случаев, а именно для одномодового лазерного диода с широким контактом и для многомодового планарного лазерного диода с полосковой структурой.

Zagadnienie pomiaru odchyłki częstotliwości i koncepcje stosownych przetworników optoelektronicznych

JANUSZ MINDYKOWSKI

Instytut Elektroenergetyki Okrętowej, Wyższa Szkoła Morska

Otrzymano 1989.03.17

Autoryzowano do druku 1989.04.10

W artykule przedstawiono zagadnienie pomiaru odchyłki częstotliwości, rozumianej jako różnica między wartością częstotliwości mierzonej i jej wartością powinna. Przedstawiono koncepcje stosownych przetworników optoelektronicznych na tle znanych rozwiązań. W pierwszej części rozpatrzono możliwości wykorzystania tradycyjnie stosowanych hercomierzy do pomiaru odchyłki częstotliwości. Omówiono hercomierze oparte na przetwornikach elektromechanicznych oraz na przetwornikach częstotliwości na sygnał stałoprądowy. Podano główne cechy metrologiczne tych rozwiązań oraz omówiono ich zastosowanie dla celów indykacji i sterowania. W drugiej części pracy przedstawiono otrzymane doświadczalnie charakterystyki statyczne nastawnego transoptora fotorezystorowego zasilanego bezpośrednio ze źródła prądowego bądź też poprzez odpowiedni układ przetwarzający. Układ zawiera prostownik liniowy, aktywny filtr dolnoprzepustowy oraz przetwornik typu „napięcie na prąd”. Omówiono dwie wersje takiego układu, z wejściem o charakterze rezystancyjnym i z wejściem o charakterze pojemnościowym oraz ich zastosowanie do budowy transoptorowego przetwornika odchyłki częstotliwości od wartości znamionowej. Przedstawiono podstawy teoretyczne rozważanego przetwornika. Przeprowadzono analizę teoretyczną stałej przetwornika oraz błędu nieliniowości w zależności od współczynnika konstrukcyjnego n . Oszacowano jego główne właściwości metrologiczne. W końcowej części artykułu przedstawiono wyniki badań doświadczalnych oraz wpływające z nich wnioski. Omówiono wpływ wybranego punktu pracy transoptorów oraz współczynnika konstrukcyjnego n na stałą przetwarzania C^* i błąd nieliniowości δ_L przetwornika. Wyniki doświadczeń porównano z wynikami analiz teoretycznych. Sformułowano warunki konstrukcyjne dla osiągnięcia określonych parametrów technicznych.

1. WSTĘP

W warunkach ruchowych bardzo często zachodzi potrzeba porównania aktualnej częstotliwości w odniesieniu do wartości powinna. Zagadnienie to stanowi pewną analogię do zagadnienia pomiaru napięcia woltomierzem o podziałce typu „łupy napięciowej”. Ważną informacją dla użytkownika może być odchyłka częstotliwości od częstotliwości znamionowej, stanowiącej stałą cechę konstrukcyjną danego urządzenia. W innych przypadkach chodzi natomiast o względną różnicę dwóch częstotliwości. Pierwsze z zagadnień wiąże się z oceną stanu pracy urządzeń technicznych, natomiast drugie dotyczy synchroni-

zacji generatora z pracującą siecią niesztynną. Tego typu zastosowania występują głównie na statkach. Pod kątem tak zróżnicowanych potrzeb wskazane jest dokonanie przeglądu znanych rozwiązań pomiarowych oraz przedstawienie propozycji odnoszących się do zasad działania i właściwości metrologicznych potrzebnej nowej aparatury.

Bezwzględną odchyłkę częstotliwości od wartości znamionowej zdefiniujemy jako:

$$\Delta f = f - f_n \quad (1a)$$

gdzie:

f — wartość częstotliwości mierzonej

f_n — częstotliwość znamionowa

Zależność (1a) dla pulsacji przyjmuje postać:

$$\Delta \omega = \omega - \omega_n. \quad (1b)$$

Natomiast względną odchyłkę częstotliwości od wartości znamionowej określa wyrażenie:

$$\varrho = \frac{f - f_n}{f_n} = \frac{\omega - \omega_n}{\omega_n}. \quad (2)$$

Wielkość odchyłki częstotliwości od jej wartości znamionowej (często zadanej) może być w procesie eksploatacji urządzeń technicznych wykorzystana dla celów indykacji lub sterowania. W pierwszym przypadku możliwe jest wykorzystanie powszechnie znanych hercomierzy po ich stosownym przeskalowaniu. Warunkiem jest tu, by częstotliwość odniesienia była wartością stałą. Natomiast w drugim przypadku, tj. gdy częstotliwość odniesienia nie jest stała (np. sieć okrętowa) wydaje się celowe skonstruowanie przyrządu mierzącego bezpośrednio odchyłkę częstotliwości. Jego sygnał wyjściowy mógłby być wykorzystany dla celów sterowania i kontroli, np. w samoczynnych synchronizatorach. Stosownych rozwiązań poszukiwano w oparciu o znane układy mostkowe z transoptorami [11, 12, 13, 14, 17] pracującymi jako czujniki „ $I/\Delta G$ ”. Wybrano układy analogowe, kierując się głównie względami ekonomicznymi i stosunkowo niskimi wymaganiami stawianymi przez przepisy Polskiego Rejestru Statków [3, 4, 5]. Przy normalnie stosowanej częstotliwości znamionowej $f_n = 50$ Hz lub 60 Hz zakres pomiarowy powinien wynosić:

$$-0,1 \leq \varrho \leq +0,1$$

Zwykle żądana jest klasa dokładności 0,5. Rozumie się przez to, że dopuszczalny uchyb w dowolnym punkcie zakresu pomiarowego nie przekracza wartości

$$\Delta f = \pm 0,5 \frac{f_n}{100}$$

W pierwszej części niniejszej pracy przeprowadzono analizę możliwości wykorzystania znanych hercomierzy pod kątem wykorzystania wyniku pomiaru odchyłki częstotliwości dla celów indykacji oraz sterowania. Następnie przedstawiono teoretyczne podstawy działania nowych przyrządów o podobnym przeznaczeniu opartych na zastosowaniu przetworników transoptorowych „ $I/\Delta G$ ”. Dokonano też próby oszacowania dokładności takich przyrządów. Omówiono również wymagania konstrukcyjne i zakres zastosowań proponowanych rozwiązań.

2. HERCOMIERZE ANALOGOWE I MOŻLIWOŚCI ICH WYKORZYSTANIA DO POMIARU ODCHYLEŃ CZĘSTOTLIWOŚCI

Znane rozwiązania hercomierzy analogowych można podzielić na dwie grupy:

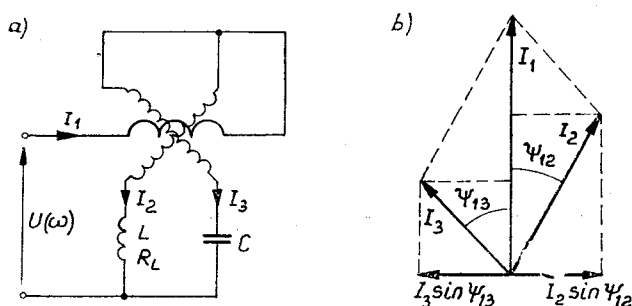
- hercomierze oparte na przetwornikach elektromechanicznych,
- hercomierze oparte na przetwornikach częstotliwości na sygnał stałoprądowy.

2.1. HERCOMIERZE BAZUJĄCE NA PRZETWORNIKACH ELEKTROMECHANICZNYCH

Do najbardziej rozpowszechnionych przyrządów do pomiaru częstotliwości opartych na przetwornikach elektromechanicznych należą hercomierze ilorazowe i wibracyjne.

Hercomierze ilorazowe

Znane hercomierze ilorazowe wykonywane są w oparciu o elektrodynamiczne, elektromagnetyczne i magnetoelektryczne ustroje pomiarowe. Na rys. 1 podano zasadniczy schemat hercomierza ilorazowego elektrodynamicznego [1, 2, 6].



Rys. 1. Hercomierz ilorazowy elektrodynamiczny, a) układ, b) wykres wektorowy prądów

Przetworniki elektrodynamiczne ilorazowe są tak skonstruowane, że spełnione są warunki [6]:

$$\frac{dL_{12}}{d\alpha} = c \sin \alpha \quad (3)$$

$$\frac{dL_{13}}{d\alpha} = c \cos \alpha$$

gdzie:

α — odchylenie organu ruchomego

c — stała konstrukcyjna.

Drogą odpowiednich przekształceń można otrzymać:

$$\operatorname{tg} \alpha = \frac{I_3 \cos \psi_{13}}{I_2 \cos \psi_{12}} \quad (4)$$

Przy uwzględnieniu, że $I_3 \sin \psi_{13} = I_2 \sin \psi_{12}$ (rys. 1b) zależność (4) można napisać w formie:

$$\operatorname{tg} \alpha = \frac{\operatorname{tg} \psi_{12}}{\operatorname{tg} \psi_{13}} \quad (4a)$$

Zakładając jednakowe indukcyjności L_c i rezystancje R_c obu cewek ruchomych oraz pomijając impedancję cewki nieruchomej i przesunięcie fazowe między napięciem U i prądem I_1 , po wykonaniu przekształceń otrzymuje się [6]:

$$\operatorname{tg} \alpha = \frac{R_c}{R_c + R_L} \frac{\omega(L + L_c)}{\omega L_c - \frac{1}{\omega C}} \quad (4b)$$

Odchylenie α jest więc nieliniową funkcją kontrolowanej częstotliwości, przy czym dla przetwornika ferrodynamicznego zależność (4b) może mieć nieco zmienioną postać. W zależności od rozwiązania układowego w gałęziach cewek ruchomych można spotkać elementy typu $R-L$, $R-C$ lub $L-C$ i stosownie do rodzaju rozwiązania modyfikacji ulegnie równanie (4b).

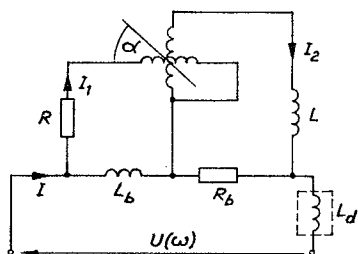
Na rys. 2 przedstawiono podstawowy schemat hercomierza ilorazowego elektromagnetycznego [6, 22]. Pomijając impedancję cewek przetwornika oraz rezystancje cewek L i L_b , dla $I_1 \ll I$ oraz $I_2 \ll I$ można napisać:

$$\begin{aligned} I_1 &= \frac{\omega L_b}{R} I \\ I_2 &= \frac{R_b}{\omega L} I. \end{aligned} \quad (5)$$

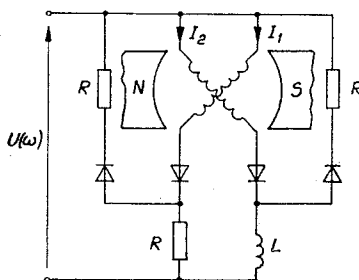
Biorąc pod uwagę, iż dla tego miernika $\alpha = \varphi \left(\frac{I_1}{I_2} \right)$, po uwzględnieniu zależności (5) otrzymamy

$$\alpha = \varphi \left(\frac{L_b L}{R_b R} \cdot \omega^2 \right). \quad (6)$$

Niekiedy szeregowo z napięciem o częstotliwości kontrolowanej włącza się dławik [22], którego indukcyjność L_d ma na celu zmniejszenie wpływu wyższych harmonicznych na wskazania hercomierza oraz może służyć do lepszego dopasowania ustroju pomiarowego do napięcia zasilającego przyrząd. Na rys. 3 przedstawiono podstawowy schemat herco-



Rys. 2. Hercomierz ilorazowy elektromagnetyczny



Rys. 3. Hercomierz ilorazowy magnetoelektryczny

mierza ilorazowego magnetoelektrycznego [2]. Pomijając impedancję cewek przetwornika, rezystancję cewki L oraz rezystancję diod można zapisać:

$$I_1 = \frac{U}{R} \quad I_2 = \frac{U}{\omega L}. \quad (7)$$

Stosunek prądów w cewkach wynosi:

$$\frac{I_1}{I_2} = \frac{\omega L}{R} \quad (8)$$

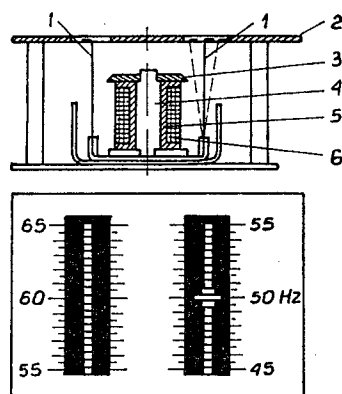
a odchylenie organu ruchomego

$$\alpha = \varphi \left(\frac{\omega L}{R} \right) \quad (9)$$

Zaletą tego rodzaju hercomierzy, obok niezależności wskazań od zmian napięcia charakterystycznej dla wszystkich logometrów, jest niewielki wpływ temperatury, pod warunkiem, że parametry zastosowanych diod prostownikowych są identyczne. Hercomierze ilorazowe charakteryzują się niewielkim zakresem pomiarowym, dużym poborem mocy, znacznym błędem pomiaru przekraczającym 1% oraz złożoną na ogół postacią charakterystyki $\alpha = \varphi(\omega)$.

Hercomierze wibracyjne

W hercomierzach wibracyjnych elementem wskazującym są języczki wykonane z blachy stalowej i pobudzane do drgań mechanicznych. Źródłem tych drgań są elektromagnesy wzbudzone napięciem o mierzonej częstotliwości. Przykładowe rozwiązanie konstrukcyjne takiego miernika przedstawiono na rys. 4.



Rys. 4. Hercomierz wibracyjny, 1 — języczki, 2 — skala, 3 — nakładka, 4 — rdzeń, 5 — korpus cewki, 6 — uzwojenie elektromagnesu [22]

Hercomierz taki ma dwa szeregi języczków, dzięki czemu może być wykorzystany do pomiaru częstotliwości zarówno w sieci o częstotliwości 50 Hz, jak i 60 Hz. Dla zapewnienia jednoznaczności wskazań, parametry poszczególnych języczków dobiera się tak, aby ich częstotliwość drgań swobodnych odpowiadała różnym częstotliwościom siły pobudzającej. Gdy częstotliwość własna języczka jest dokładnie równa częstotliwości siły pobudzającej

zachodzi rezonans mechaniczny i amplituda drgań osiąga wówczas maksimum. Dla wartości leżącej w połowie przedziału między częstotliwościami swobodnymi dwu sąsiednich języczków, będą one drgać z jednakową amplitudą, jednak wyraźnie mniejszą od rezonansowej. Przy jeszcze większej rozbieżności amplituda drgań danego języczka staje się bardzo niewielka.

Warto zauważyć, że w przypadku innego rozwiązania konstrukcyjnego, polegającego na przyporządkowaniu każdemu szeregowi języczków niezależnego układu elektromagnetycznego wzbudzających, można otrzymać przyrząd dwusystemowy, kontrolujący jednocześnie częstotliwość dwóch generatorów lub generatora i pracującej sieci elektroenergetycznej.

Do podstawowych wad hercomierzy wibracyjnych zaliczyć można:

- wpływ kontrolowanego napięcia na częstotliwość rezonansową języczków [22]
- stosunkowo mały zakres pomiarowy (brak wskazań przy częstotliwościach sygnału mierzonego leżącego poza zakresem pomiarowym),
- mały kąt widoczności wskazania, szczególnie przy niekorzystnych warunkach oświetleniowych,
- zmiana wskazań spowodowana zmęczeniem materiału sprężystego i zmianą masy (np. złuszczenie się farby),
- wpływ obcych drgań mechanicznych na wskazania.

2.2. HERCOMIERZE OPARTE NA PRZETWORNIKACH CZĘSTOTLIWOŚCI NA SYGNAŁ STAŁOPRĄDOWY

W nowszych rozwiązaniach konstrukcyjnych do pomiaru częstotliwości stosuje się mierniki magnetoelektryczne wyposażone w przetwornik częstotliwości na napięcie lub prąd. Przetwornikiem takim może być:

- mostek prądu przemiennego,
- układ ładowania i/lub rozładowania kondensatora,
- układ bazujący na dwóch różnicowo połączonych obwodach rezonansowych,
- układ komparacyjny, oparty na wzmacniaczu operacyjnym, porównujący odpowiednio uformowany i zróżniczkowany sygnał napięciowy o częstotliwości badanej z napięciem odniesienia. Przetworniki te będą omówione niżej.

Mostki do pomiaru częstotliwości

Biorąc pod uwagę zakres częstotliwości i wymaganą na ogół dokładność w pomiarach związanych z energetyką, jako miernik odchyłki częstotliwości najbardziej nadaje się nierównoważony mostek Robinsona [6, 7]. Jego schemat przedstawiono na rys. 5.

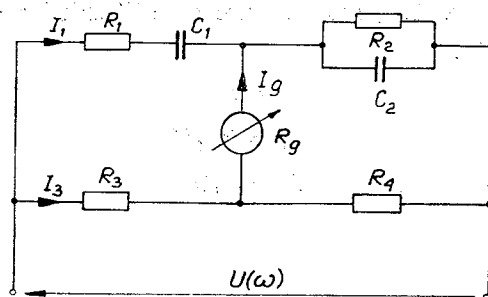
Dla mostka Robinsona w stanie równowagi słuszna jest zależność:

$$\left(R_1 + \frac{1}{j\omega C_1}\right) R_4 = \frac{R_2 R_3}{1 + j\omega R_2 C_2}. \quad (10)$$

A zatem warunki równowagi przyjmują postać:

$$\omega^2 R_1 C_1 R_2 C_2 = 1 \quad (11a)$$

$$\frac{C_2}{C_1} = \frac{R_3}{R_4} - \frac{R_1}{R_2}. \quad (11b)$$



Rys. 5. Układ niezrównoważonego mostka Robinsona

Ze względu na typizację elementów stosuje się zwykle $R_1 = R_2 = R$ oraz $C_1 = C_2 = C$, stąd $R_3 = 2R_4$. Mostek jest w równowadze dla pulsacji

$$\omega_0 = \frac{1}{RC}. \quad (12)$$

W niektórych rozwiązaniach konstrukcyjnych rezystancja gałęzi pomiarowej R_g może osiągać bardzo duże wartości. Mamy wówczas przetwornik częstotliwości z miernikiem typu napięciowego, a równanie niezrównoważonego mostka Robinsona, zgodnie z wynikami podanymi w [7] przyjmuje następującą postać:

$$U_g = I_g R_g = UR_g \frac{(1+j\eta)^2 - 2j\eta}{3R(1+j\eta) + (3R_g + 2R_4)[(1+j\eta)^2 + j\eta]}, \quad (13)$$

gdzie:

I_g — wartość prądu wyjściowego,

U — napięcie zasilające mostek (napięcie o częstotliwości badanej),

R_g — rezystancja przyrządu pomiarowego,

R, R_4 — rezystancje odpowiednich gałęzi mostka,

$\eta = \frac{\omega}{\omega_0}$ — względna pulsacja

ω — pulsacja mierzona.

W krańcowym przypadku, gdy $R_g = \infty$ uzyskuje się po przekształceniach

$$U_g = U \frac{1 - \eta^2}{3[(1+j\eta)^2 + j\eta]}, \quad (13a)$$

a po wprowadzeniu współczynnika zmian częstotliwości mostka

$$X = \frac{3\eta}{1 - \eta^2}$$

dostaje się

$$U_g = \frac{U}{3} \frac{1}{1 + jX}. \quad (13b)$$

Na podstawie wyrażenia (13b) można łatwo określić składowe: czynną, bierną i pozorną napięcia pomiarowego:

$$\begin{aligned} U_{gR} &= k_{UR} U \\ U_{gX} &= k_{UX} U \\ U_g &= k_U U \end{aligned} \quad (13c)$$

gdzie:

$$\begin{aligned} k_{UR} &= \frac{1}{3} \frac{1}{1+X^2}, \\ k_{UX} &= \frac{1}{3} \frac{-X}{1+X^2}, \quad k_U = \frac{1}{3} \frac{1}{\sqrt{1+X^2}}. \end{aligned}$$

Najlepszą charakterystykę mierniczą otrzymuje się dla składowej U_g [7]. Z zależności $U_g = f(\omega)$ wynika, iż wartości odpowiednich składowych napięcia pomiarowego dla rozważanego przetwornika są zależne nie tylko od wartości napięcia zasilającego U , lecz także od stosunku częstotliwości ω napięcia zasilającego do częstotliwości własnej ω_0 , dla której mostek został skonstruowany. Warto zaznaczyć, iż błąd pomiaru tymi mostkami nie przekracza 0,5%, o ile napięcie zasilające ma kształt sinusoidalny. Błąd ten rośnie, gdy napięcie zasilające zawiera wyższe harmoniczne. Dla zwiększenia dokładności pomiaru należy wówczas stosować selektywny miernik wyjściowy.

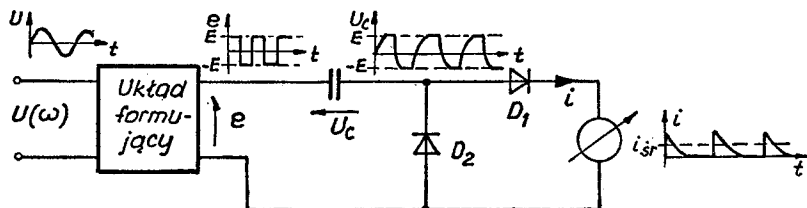
Hercomierze integracyjne

Zasada działania hercomierza integracyjnego polega na cyklicznym ładowaniu i rozładowaniu kondensatora z częstotliwością mierzoną i pomiarze prądu rozładowania, ładowania lub ładowania i rozładowania kondensatora.

Hercomierze integracyjne są wykonywane w dwóch podstawowych odmianach [2, 6]: z przełączaniem elektromechanicznym i z przełączaniem elektronicznym. Hercomierze według pierwszej wersji charakteryzują się liniową podziałką, zakresem pomiarowym (10 ÷ 200 Hz) ograniczonym bezwładnością przełącznika i mają niezbyt dużą dokładność (1 ÷ 3%).

Układ hercomierza integracyjnego z przełączaniem elektronicznym przedstawiono na rys. 6.

Napięcie o częstotliwości mierzonej jest przetwarzane w układzie formującym na napięcie o przebiegu prostokątnym i częstotliwości równej częstotliwości mierzonej.



Rys. 6. Hercomierz integracyjny z przełączaniem elektronicznym

W czasie cyklu ładowania i rozładowania napięcie na kondensatorze C zmienia się od $-E$ do $+E$ i odwrotnie. Za pomocą diod D_1 i D_2 oraz miernika magnetoelektrycznego mierzony jest prąd rozładowania kondensatora, o wartości średniej określonej zależnością

$$i_{sr} = 2 \cdot E \cdot C \cdot f, \quad (14)$$

gdzie f — częstotliwość badanego napięcia.

Odchylenie organu ruchomego miernika α wyraża zależność

$$\alpha = S_i \cdot i_{sr} \quad (15)$$

gdzie S_i — czułość miernika.

Zależności (14) i (15) są słuszne przy założeniu, że stałe czasowe ładowania τ_1 i rozładowania τ_2 kondensatora C są znacznie mniejsze od okresu napięcia prostokątnego T . Stałe te są równe:

$$\begin{aligned} \tau_1 &= (R_w + R_D + R_m)C \\ \tau_2 &= (R_w + R_D)C, \end{aligned} \quad (16)$$

gdzie:

R_w — rezystancja wyjściowa układu formującego

R_D — rezystancja diody

R_m — rezystancja miernika.

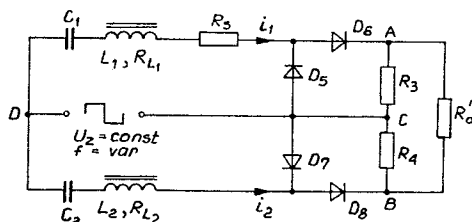
Błąd pomiaru hercomierza integracyjnego zależy głównie od dokładności kształtowania impulsów prostokątnych, stałości amplitudy napięcia prostokątnego, stałości pojemności oraz klasy użytego miernika magnetoelektrycznego.

Błąd ten nie przekracza na ogół 2%, przy czym istnieje możliwość zmniejszenia tej wartości znacznie poniżej 1% — co wiąże się jednak z koniecznością wykorzystania w układzie formującym przerzutników, wymagających dodatkowego stabilizowanego napięcia zasilania.

Ze względu na prostą budowę, szeroki zakres mierzonych częstotliwości, stosunkowo dobrą dokładność pomiaru i proporcjonalną zależność wskazań od częstotliwości hercomierze integracyjne są chętnie stosowane.

Rezonansowy przetwornik częstotliwości napięć sieciowych

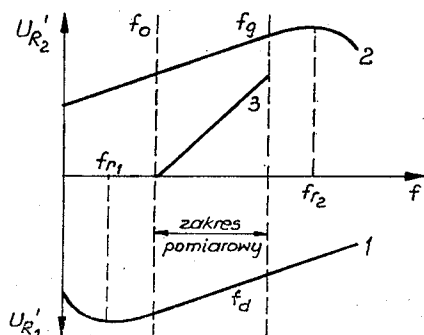
Przetwornik częstotliwości przetwarza częstotliwość napięć sieciowych na proporcjonalną wartość natężenia prądu stałego. Układ elektryczny przetwornika zawiera następujące elementy: transformator wejściowy, stabilizator napięcia przemiennego, dwa szeregowo obwody rezonansowe z diodami prostownikowymi, filtr prądu wyjściowego i opornik wyrównawczy. Uproszczony schemat przetwornika [21] przedstawiono na rys. 7.



Rys. 7. Uproszczony schemat pomiarowego przetwornika częstotliwości

Zasada pomiaru (przetwarzania) omawianego przetwornika polega na wykorzystaniu zależności prądów dwóch obwodów rezonansowych od częstotliwości napięcia przyłożonego do tych obwodów. Wzrost częstotliwości leżącej w zakresie pomiarowym przetwornika tj. $f_{r1} < f < f_{r2}$ (rys. 8) powoduje wzrost prądu I_2 w drugim obwodzie rezonansowym przetwornika i zmniejszenie prądu I_1 w jego pierwszym obwodzie rezonansowym. Różnica jednopółkowo wyprostowanych prądów obu obwodów rezonansowych, a ściślej różnica spadków napięć na opornikach R_3 i R_4 wywołanych przez te prądy jest miarą częstotliwości mierzonej. Ilustruje to rys. 8, na którym przedstawiono krzywe rezonansowe obu obwodów RLC i wypadkową różnicy rzędnych tych krzywych w zakresie pomiarowym częstotliwości.

Równanie charakterystyki przetwarzania omawianego przetwornika częstotliwości,



Rys. 8. Krzywe rezonansowe napięć U'_{R1} i U'_{R2} obwodów rezonansowych przetwornika częstotliwości, 1 — krzywa dla pierwszego obwodu rezonansowego, 2 — krzywa dla drugiego obwodu rezonansowego, 3 — krzywa wypadkowa będąca różnicą rzędnych krzywych 1 i 2, f_{r1} — częstotliwość rezonansowa pierwszego obwodu rezonansowego, f_{r2} — częstotliwość rezonansowa drugiego obwodu rezonansowego, f_d — dolna częstotliwość graniczna zakresu pomiarowego, f_g — górna częstotliwość graniczna zakresu pomiarowego

w którym zmienna niezależna f przybiera wartości od f_d do f_g a zmienna zależna ΔU_{AB} — od 0 do ΔU_{ABn} ma następującą postać [21]:

$$\Delta U_{AB} = \frac{U_{ODm}}{\pi} \cdot R' \cdot \frac{\sqrt{R_1^2 + \left(\omega L_1 - \frac{1}{\omega C_1}\right)^2} - \sqrt{R_2^2 + \left(\omega L_2 - \frac{1}{\omega C_2}\right)^2}}{\sqrt{R_1^2 + \left(\omega L_1 - \frac{1}{\omega C_1}\right)^2} \cdot \sqrt{R_2^2 + \left(\omega L_2 - \frac{1}{\omega C_2}\right)^2}}, \quad (17)$$

gdzie:

U_{ODm} — amplituda harmonicznego podstawowego napięcia zasilającego

$R' = R'_1 = R'_2$ — rezystancje otrzymane przy przekształceniu trójkąta ABC złożonego z rezystancji R_3 , R_4 i R_0 na równoważną gwiazdę o rezystancjach: $R'_1 R'_2, R'_{10}$.

$R_1 = R_s + R'_1 + R_{L1}$

$R_2 = R'_2 + R_{L2}$

Przy odpowiednio zaprojektowanych obwodach rezonansowych, można uzyskać praktycznie liniową zależność napięcia wyjściowego od częstotliwości mierzonej otrzymując

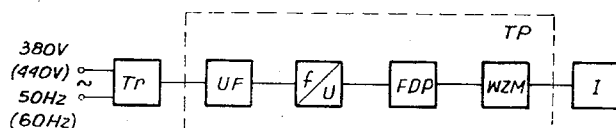
niedokładność przetwarzania $\pm 1\%$ przy nieliniowości charakterystyki przetwarzania $\pm 0,6\%$ [21]. Warto podkreślić, iż na skutek zastosowania stabilizatora napięcia przemiennego, o prostokątnym napięciu wyjściowym U_Z zasilającym układ pomiarowy (rys. 8), przedstawiony przetwornik częstotliwości nie jest wrażliwy ani na wahania napięcia sieciowego, ani na kształt krzywej tego napięcia. Nadto, transformator wejściowy izoluje galwanicznie układ pomiarowy przetwornika od sieci.

Hercomierze elektroniczne analogowe

Istnieje wiele rozwiązań hercomierzy elektronicznych. Jedną z wersji analogowych, jedno- i dwusystemowy hercomierz elektroniczny, opracowano w Zespole Metrologii Eksploatacyjnej Statku Instytutu Elektroenergetyki Okrętowej WSM w Gdyni [3, 4, 5].

Podstawowym elementem struktury przyrządu jest tor pomiarowy, przedstawiony na rys. 9.

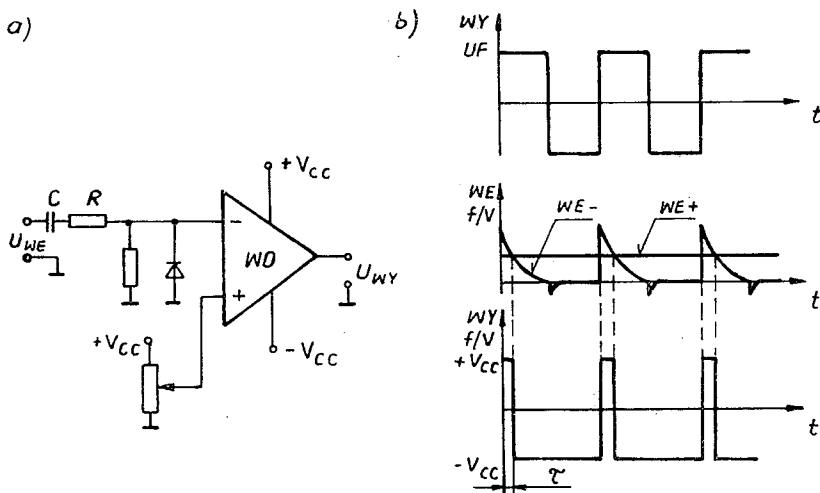
Sinusoidalny sygnał wejściowy o amplitudzie 380 V/440 V obniżony za pomocą transformatora Tr i uformowany do postaci fali prostokątnej w układzie formującym UF , steruje



Rys. 9. Schemat blokowy hercomierza elektronicznego analogowego (dla jednego toru pomiarowego)

przetwornikiem częstotliwości na napięcie f/U . Układ ten wytwarza sygnał napięciowy, którego wartość średnia jest proporcjonalna do częstotliwości badanego przebiegu.

Przetwornik f/U zbudowano na bazie wzmacniacza operacyjnego. Schemat przetwornika przedstawia rys. 10a, zaś zależności czasowe sygnałów — rys. 10b. Układ różniczkujący



Rys. 10. Przetwornik częstotliwość/napięcie oparty na wzmacniaczu operacyjnym, a) schemat połączeń b) zależności czasowe sygnałów

jący RC formuje sygnał podawany na jedno z wejść wzmacniacza U_{WE} . Na drugie wejście podawane jest stałe napięcie odniesienia $+V_{cc}$. Na wyjściu przetwornika pojawia się ciąg impulsów prostokątnych o okresie $T = \frac{1}{f}$ i czasie trwania τ . Wartość średnią napięcia wyjściowego określa zależność:

$$U_{WY} = f \cdot \tau \cdot (U_1 - U_2) + U_2 \quad (18)$$

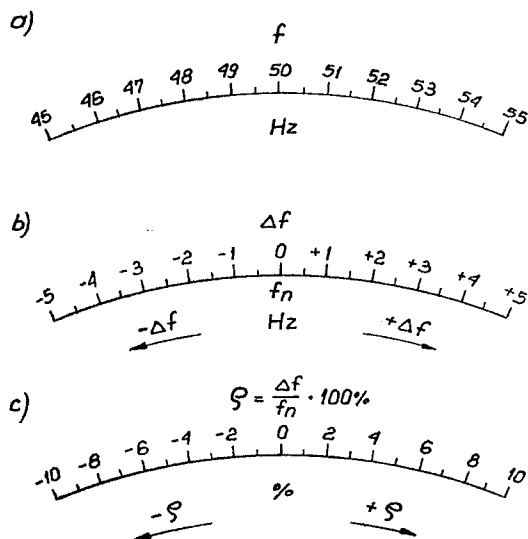
gdzie U_1 , U_2 — wartości napięć nasycenia wzmacniacza operacyjnego.

Filtr dolnoprzepustowy FDP, następujący po przetworniku f/U , eliminuje z przebiegu składowe o częstotliwościach zbliżonych do 50 Hz oraz ich wyższe harmoniczne. Wzmacniacz prądu stałego WZM wzmacnia wartość sygnału napięciowego z filtru dolnoprzepustowego FDP, dostosowując jego poziom do potrzeb elektroluminescencyjnego układu ekspozycji wyniku I .

Wartość mierzonej częstotliwości wskazywana jest w zespole 13 diod elektroluminescencyjnych. Poszczególne diody wskazują częstotliwości odległe od siebie o 0,5 Hz, a różnica o 0,25 Hz wskazywana jest przez świecenie dwóch sąsiednich diod. Dwie dodatkowe diody skrajne służą do sygnalizowania przekroczeń nominalnego zakresu pomiarowego. Przyrząd wykonano w klasie 0,5.

2.3 HERCOMIERZE ANALOGOWE A POMIAR ODCHYLENIA CZĘSTOTLIWOŚCI

Powszechnie znane hercomierze analogowe różnych typów są w zasadzie miernikami o obustronnie ograniczonym zakresie. Przez zmianę skalowania tych mierników można uzyskać miernik odchyłki częstotliwości od wartości znamionowej, co przedstawiono na



Rys. 11. Zmiana podziałki mierników częstotliwości w celu uzyskania miernika odchyłki częstotliwości, a) klasycznego hercomierza wskazówkowego, b) skala miernika bezwzględnej odchyłki częstotliwości od wartości znamionowej, c) skala miernika względnej odchyłki częstotliwości od wartości znamionowej

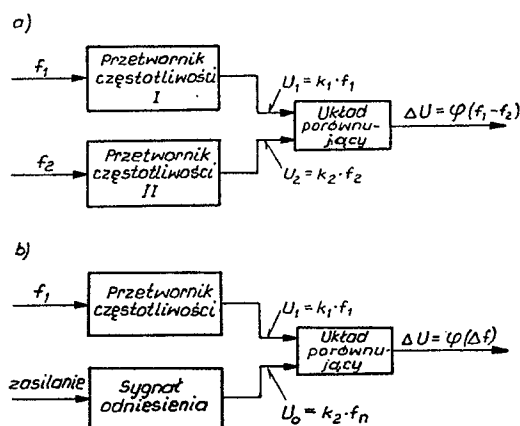
rys. 11. W tym przypadku wartość odchyłki częstotliwości zostaje przetworzona na odchylenie wskazówki przyrządu analogowego lub ekspozycję wyniku w postaci cyfrowej. Tak sformułowane przekształcenie tradycyjnego hercomierza pozwala każdy z przedstawionych typów przyrządu traktować jako potencjalny miernik odchyłki częstotliwości od wartości znamionowej dla potrzeb indykacji. Natomiast w przypadku wykorzystania odchyłki częstotliwości od wartości znamionowej dla potrzeb sterowania konieczne jest przetworzenie wartości Δf na odpowiadającą jej wartość prądu czy napięcia, a więc uzyskanie określonego sygnału analogowego lub cyfrowego. Do takich zastosowań nie nadają się hercomierze ilorazowe i wibracyjne.

W oparciu o przetworniki częstotliwości na sygnał stałoprądowy istnieje natomiast możliwość budowy miernika odchyłki częstotliwości od wartości znamionowej (lub różnicy dwóch częstotliwości) z wyjściem analogowym.

Schemat blokowy takiego miernika jest zależny od stałości częstotliwości powinna. W przypadku, gdy częstotliwość ta nie jest stała, konieczne jest wykorzystanie dwóch przetworników częstotliwości na sygnał stałoprądowy oraz członu porównującego.

Jeden z przetworników będzie kontrolował częstotliwość badanego napięcia, drugi częstotliwość powinna sieci, a skonstruowany miernik będzie mierzył różnicę obu częstotliwości. Natomiast w sytuacji, gdy częstotliwość powinna ma wartość stałą, drugi z przetworników częstotliwości można zastąpić sygnałem napięcia lub prądu odniesienia, odpowiadającym tej wartości. Przyrząd mierzy wówczas odchyłkę częstotliwości od wartości znamionowej.

Odpowiednie schematy blokowe przedstawiono na rys. 12. Schematy blokowe przedstawione na rys. 12 bazują na przetwornikach f/U i dotyczą pomiarów bezwzględnej różnicy dwóch częstotliwości oraz bezwzględnej odchyłki częstotliwości od wartości znamionowej. Analogiczne rozwiązania można podać w oparciu o przetworniki f/I . Nadto, dodając układ dzielący i wykorzystując sygnał odpowiednio U_2 bądź U_0 można łatwo uzyskać przetworniki względnej różnicy dwóch częstotliwości bądź też względnej odchyłki częstotliwości od wartości znamionowej.



Rys. 12. Schemat blokowy przetwornika, a) różnicy dwóch częstotliwości, dla sieci o częstotliwości powinna podlegającej wahaniom, b) odchyłki częstotliwości od wartości znamionowej, dla sieci o stałej wartości częstotliwości powinna

Przedstawione wcześniej znane rozwiązania przetworników częstotliwości na sygnał stałoprądowy posiadają szereg sygnalizowanych wad. Mierniki odchyłki częstotliwości lub różnicy dwóch częstotliwości wykonane w oparciu o te przetworniki byłyby układami o znacznej złożoności technicznej i niewielkiej niezawodności. Wydaje się więc celowym poszukiwanie nowego typu przetwornika. Ma on mierzyć bezpośrednio odchyłkę częstotliwości od wartości powinnej i mieć charakterystykę zbliżoną do proporcjonalnej. Sygnał wyjściowy takiego przetwornika winien być odseparowany galwanicznie od sieci. W dalszych rozważaniach będą brane pod uwagę układy analogowe, w których wykorzystano transoptorowe elementy przetwarzające typu „ $I/\Delta G$ ”.

3. TRANSOPTOROWE ELEMENTY PRZETWARZAJĄCE „ $I/\Delta G$ ” I KONCEPCJA ICH WYKORZYSTANIA W PRZETWORNIKACH ODCHYLKI CZĘSTOTLIWOŚCI

3.1. PODSTAWOWE CECHY NASTAWNEGO TRANSOPTORA FOTOREZYSTOROWEGO

Poniżej rozpatrzono transoptor jako człon analogowego przetwornika pomiarowego, a więc przypisano mu funkcję imitancji sterowanej prądem diody elektroluminescencyjnej. Ograniczono się do transoptorów o wyjściu fotorezystorowym, jako mogących działać niezależnie od polaryzacji napięcia przyłożonego do fotodetektora, co znacznie poszerza zakres zastosowań rozważanych przetworników. Nadto biorąc pod uwagę stosunkowo duże dopuszczalne prądy obciążenia fotorezystorów [19], w układach z takimi elementami możliwe jest uzyskanie korzystnej czułości napięciowej sygnału wyjściowego względem wartości mierzonej [8, 14].

W nastawnych transoptorach fotorezystorowych [8, 10] o regulowanej odległości „ a ” między diodą elektroluminescencyjną a fotorezystorem, ich przewodność wyjściowa w przypadku ogólnym jest funkcją czterech zmiennych:

$$G_T = f(I, a, \vartheta, U), \quad (19)$$

gdzie:

I — prąd sterujący transoptora

a — odległość między diodą elektroluminescencyjną a fotodetektorem

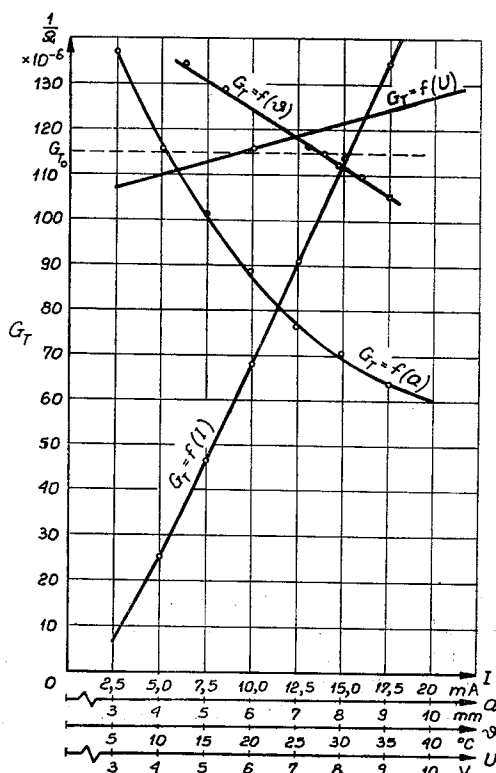
ϑ — temperatura otoczenia

U — napięcie przyłożone do fotodetektora.

W wyniku badań wstępnych [8] postanowiono traktować prąd sterujący I jako wielkość sterującą przewodność wyjściową transoptora, odległość „ a ” — jako parametr regulacyjny wyznaczający punkt pracy transoptora, zaś wielkości U oraz ϑ — jako zakłócenia. Do badań wykorzystano transoptory zbudowane w oparciu o diody elektroluminescencyjne CQYP40 i fotorezystory RPP 131. Wyniki badań wstępnych przedstawiono na rys. 13. Do wyznaczenia równania (19) posłużono się jedną z metod planowania doświadczenia, opartą na rachunku prawdopodobieństwa, a mianowicie analizę czynnikową [8].

W rozważanym przypadku najogólniejsza postać zależności (19) może być przedstawiona jako:

$$\begin{aligned} G_T = & G_0 + G_1 I + G_2 a + G_3 \vartheta + G_4 U + G_{11} I^2 + \\ & + G_{22} a^2 + G_{33} \vartheta^2 + G_{44} U^2 + G_{12} I a + G_{13} I \vartheta + \\ & + G_{14} I U + G_{23} a \vartheta + G_{24} a U + G_{34} \vartheta U. \end{aligned} \quad (20)$$



Rys. 13. Zależność $G_T = f(I)$, $G_T = f(a)$, $G_T = f(T)$ i $G_T = f(U)$ przy pozostałych trzech wielkościach stałych: $I_0 = 15$ mA, $a_0 = 4$ mm, $T_0 = 26^\circ\text{C}$, $U_0 = 6$ V

Na podstawie cyklu eksperymentów częściowych przyjęto $a = 5$ mm = const i szukano opisu właściwości statycznych transoptora w postaci równania:

$$G_T = G_0 + G_1 I + G_3 \vartheta + G_4 U + G_{33} \vartheta^2 + G_{44} U^2 + G_{13} I\vartheta + G_{14} IU. \quad (21)$$

Współczynniki równania (21) wyznaczono metodą analizy regresyjnej na podstawie wyników kilku serii pomiarów, wykonanych w układzie przedstawionym na rys. 14a) i b)

Zasadnicza różnica między zasilaniem według układów 14b) i 14c) polega na tym, że w pierwszym przypadku wielkością wejściową dla badanego elementu transoptorowego jest prąd I , natomiast w drugim — napięcie U_0 . Przetwornik U/I jest opisany liniowym równaniem przetwarzania, tj. $I = k_w U_0$ (dla zbudowanego modelu $k_w \approx 0,01$ mA/mV), i zapewnia prądowe sterowanie diody elektroluminescencyjnej w zakresie od 0 do I_{dop} . Spośród rozważanych modeli matematycznych właściwości statycznych nastawnego transoptora fotorezystorowego [9] na szczególną uwagę zasługują dwa z nich, które prowadzą do następujących wyników:

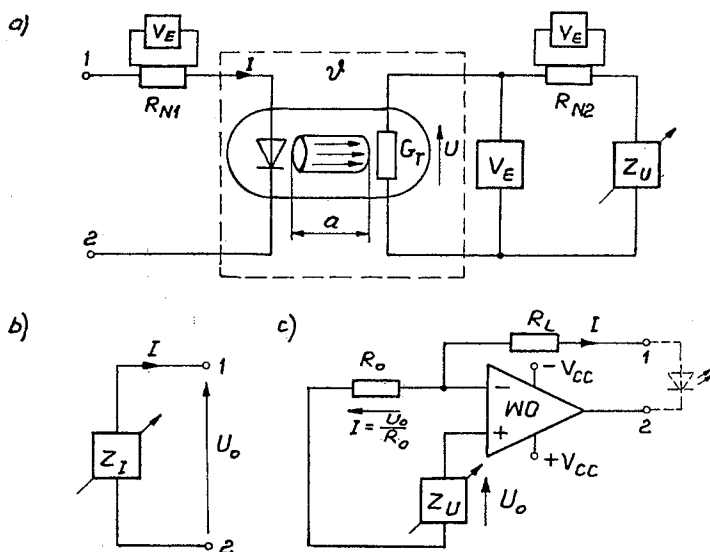
$$G'_T = A' \cdot I + BU \quad (22a)$$

$$G''_T = A'' \cdot I, \quad (22b)$$

gdzie:

$$A' = A'' = A = G_1 + G_{13} \vartheta$$

$$B = G_4$$



Rys. 14. Układ pomiarowy do wyznaczania właściwości statycznych transoptora fotorezystorowego, a) układ badanego transoptora, b) zasilanie badanego transoptora ze źródła prądowego, c) zasilanie badanego transoptora z przetwornika U/I , Z_I — zasilacz stabilizowany prądowy, Z_U — zasilacz stabilizowany napięciowy, V_E — woltomierze elektroniczne, R_N — rezystory wzorcowe, R_o — rezystor dopasowujący, R_L — rezystancja obciążenia, WO — wzmacniacz operacyjny, $\pm V_{cc}$ — napięcie zasilające wzmacniacz operacyjny

Szczegółowe dane o jakości aproksymat (22a) i (22b) w porównaniu z innymi rozważanymi modelami oraz o wartościach współczynników A i B można znaleźć w pracach [8, 9].

Proponowane modele matematyczne właściwości statycznych transoptora fotorezystorowego są słuszne dla niewielkiego względnego zakresu zmian odległości „ a ”. Parametr „ a ” pozwala wybrać punkt pracy danego transoptora, jak również dopasować charakterystyki współpracujących ze sobą transoptorów, przy czym przewodność G_T jest w przybliżeniu odwrotnie proporcjonalna do $\sqrt{a}$.

Przewodność rozważanego transoptora zależy przy tych założeniach głównie od prądu sterującego, zaś zmiany temperatury i napięcia zakłócają pomiar. W układach pomiarowych z takimi elementami należy więc stosować połączenia różnicowe czujników, najlepiej umieszczając je we wspólnej obudowie, ekranującej termicznie.

Z otrzymanych rezultatów pomiarowych wynika, iż dla rozważanego transoptora przewodność „ciemna” G_0 jest pomijalna. Należy podkreślić, iż powyższe rozważania są słuszne dla prądów sterujących większych od prądu progowego diody elektroluminescencyjnej i w zastosowaniach praktycznych trzeba ten fakt uwzględnić.

Do badań laboratoryjnych wykorzystano zasilanie badanych transoptorów wg rys. 14b, natomiast do konstrukcji przetworników odchyłki częstotliwości, z uwagi na większą elastyczność i posiadane zalety — układ wg rys. 14c. Wyznaczona doświadczalnie w układzie wg rys. 14a) i c) średnia wartość współczynnika A dla badanej serii transoptorów jest równa $6,7 \frac{\mu S}{mA}$, zaś wartość prądu progowego diody jest rzędu 2,5 mA.