POLSKA AKADEMIA NAUK
KOMITET ELEKTROTECHNIKI

526

ROZPRAWY ELEKTROTECHNICZNE

KWARTALNIK

SPIS TREŚCI DO TOMU XX — 1974

PAŃSTWOWE WYDAWNICTWO NAUKOWE
WARSZAWA

POLSKA AKADEMIA NAUK
KOMITET ELEKTROTECHNIKI

SPIS TREŚCI DO TOMU XX

CONTENTS

TABLE DES MATIÈRES

INHALT

СОДЕРЖАНИЕ

POLSKA AKADEMIA NAUK
KOMITET ELEKTROTECHNIKI

ROZPRAWY ELEKTROTECHNICZNE

TOM XX · ZESZYT 1

KWARTALNIK

PAŃSTWOWE WYDAWNICTWO NAUKOWE
WARSZAWA 1974

KOMITET REDAKCYJNY

Redaktor Naczelny

PROF. DR WITOLD NOWICKI

Członkowie

PROF. TADEUSZ CHOLEWICKI, PROF. EUGENIUSZ JEZIERSKI,
PROF. DR WITOLD ROSIŃSKI, PROF. DR STANISŁAW RYŻKO

Sekretarz

JANINA MATHEISEL

ADRES REDAKCJI

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

PAŃSTWOWE WYDAWNICTWO NAUKOWE
Warszawa, Miodowa 10

Nakład 700 (544+156)	Oddano do składania 15. VIII. 1973 r.
Ark. wyd. 19,25 ark. druk. 15,75	Podpisano do druku w grudniu 1973 r.
Papier druk. sat. kl. III, 80 g, 70×100	Druk ukończono w styczniu 1974 r.
Zamówienie nr 1193/73 R—8	Cena zł 40.—

Drukarnia im. Rewolucji Październikowej. Warszawa

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 1, ss. 3-13

621.372.44:621.375:621.372.57

Zastosowanie metody rewersyjnej do analizy nieliniowych obwodów elektrycznych

ZYGMUNT RYDZEWSKI (ŁÓDŹ)

Instytut Transformatorów, Maszyn i Aparatów Elektrycznych Politechniki Łódzkiej

Otrzymano 14.5.1953

W pracy podano rozwiązania nieliniowych równań różniczkowych obwodów szeregowych z liniowymi i nieliniowymi indukcyjnościami, uzyskane metodą rewersyjną.

Dla obwodu prądu stałego nieliniową funkcję aproksymowano wielomianem o nieparzystych potęgach do piątej włącznie.

W przypadku obwodu prądu przemiennego otrzymano rozwiązanie $i(t)$ dla dowolnie wybranej fazy napięcia wymuszającego przy wykorzystaniu wielomianu aproksymującego trzeciego stopnia.

1. WSTĘP

Rewersyjna metoda rozwiązywania pewnego typu nieliniowych zwyczajnych równań różniczkowych, przedstawiana przez kilku autorów [1], [2], [3], [4], może być stosowana przy analizie nieliniowych obwodów elektrycznych.

Idea metody polega na zastąpieniu nieliniowego zwyczajnego równania różniczkowego określonego typu przez układ równań liniowych rozwiązywanych za pomocą rachunku operatorowego.

2. TEORIA METODY

Rozważmy nieliniowe zwyczajne równanie różniczkowe:

$$w_1 y + w_2 y^2 + w_3 y^3 + \ldots + w_n y^n = kV(t), \tag{1}$$

przy założeniu $w_1 \neq 0$, w którym:

t — zmienna niezależna,

$y(t)$ — nieznane rozwiązanie — odpowiedź układu,

k — stała,

$V(t)$ — znana funkcja czasu — funkcja wymuszająca,

$w_1 \ldots w_n$ — ogólne operatory lub funkcje operatora $D = \frac{d}{dt}$.

Przewidujemy rozwiązanie równania (1) $y(t)$ o postaci:

$$y = y_1 + y_2 + \ldots + y_n = A_1 k + A_2 k^2 + \ldots + A_n k^n. \tag{2}$$

Podstawiając (2) do (1) i porównując po obu stronach wyrażenie o tej samej potędze k otrzymujemy układ liniowych równań różniczkowych.

Pierwsze z nich wynika z porównania wyrażeń

$$w_1 y_1 = kV(t);$$

podstawiając $y_1 = A_1 k$ mamy

$$w_1 A_1 = V(t); \tag{3}$$

a zatem

$$A_1 = \frac{V(t)}{w_1}. \tag{4}$$

Drugie z równań liniowych ma postać:

$$w_1 A_2 k^2 + w_2 A_1^2 k^2 = 0,$$

więc

$$w_1 y_2 = -w_2 y_1^2; \tag{5}$$

ostatecznie

$$A_2 = -\frac{w_2}{w_1} A_1^2. \tag{6}$$

W analogiczny sposób można wyliczyć współczynniki $A_3 \ldots A_n$. Poniżej podano współczynniki do szóstego włącznie

$$A_3 = -\frac{1}{w_1}(2w_2 A_1 A_2 + w_3 A_1^3), \tag{7}$$

$$A_4 = -\frac{1}{w_1}[w_2(A_2^2 + 2A_1 A_3) + 3w_3 A_1^2 A_2 + w_4 A_1^4], \tag{8}$$

$$A_5 = -\frac{1}{w_1}[2w_2(A_1 A_4 + A_2 A_3) + 3w_3(A_1 A_2^2 + A_1^2 A_3) + 4w_4 A_1^3 A_2 + w_5 A_1^5], \tag{9}$$

$$A_6 = -\frac{1}{w_1}[w_2(A_2^3 + 2A_1 A_5 + 2A_2 A_4) + w_3(A_2^3 + 3A_1^2 A_4 + 6A_1 A_2 A_3) + 2w_4(2A_1^2 A_2^2 + 2A_1^3 A_3) + 5w_5 A_1^4 A_2 + w_6 A_1^6]. \tag{10}$$

Pierwsze trzynaście wyrazów stosowanych w metodzie rewersyjnej podali Van Orstrand [2] oraz G. Cohn i B. Saltzberg [4].

Mając wyznaczone współczynniki $A_1, A_2, \ldots, A_n$ można na podstawie (2) wyrazić rozwiązanie równania (1) wykorzystując na przykład przekształcenie Laplace'a-Carsona.

Rozwiązanie równania (1) jest więc kombinacją wielomianów i funkcji wykładniczych.

3. ZASTOSOWANIE METODY REWERSYJNEJ DO BADANIA NIELINIOWYCH OBWODÓW ELEKTRYCZNYCH

3.1. Elektryczny nieliniowy obwód prądu stałego

Rozpatrzmy nieliniowy szeregowy obwód podany na rys. 1. Zawiera on nieliniową indukcyjność $\mathscr{L}$.

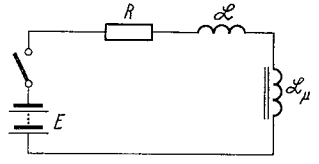
Równanie różniczkowe tego obwodu ma postać

$$\mathscr{L}\frac{di}{dt}+Ri+z\frac{d\varphi}{dt}=E, \qquad i(t)=0 \quad \text{dla} \quad t=0, \tag{11}$$

przy czym:

E — siła elektromotoryczna ogniwa,
$\mathscr{L}$ — indukcyjność cewki liniowej,
$\mathscr{L}_\mu$ — cewka z rdzeniem ferromagnetycznym,
R — rezystancja,
z — liczba zwojów cewki z rdzeniem.

Aproksymacja nieliniowej funkcji $\psi = f(i)$ elementu $\mathscr{L}_\mu$ — zgodnie z założeniami metody — powinna być przyjęta w postaci wielomianu. Funkcja $\psi = f(i)$ przedstawia, pomi-



Rys. 1. Badany obwód prądu stałego

jając pętlę histerezy, jednowartościową charakterystykę magnesowania spełniającą warunek symetrii $f(\xi) = -f(\xi)$; założono też, że pozostałość magnetyczna jest równa zeru. Można ją więc aproksymować za pomocą wielomianu o potęgach nieparzystych:

$$\psi = z\varphi = L_1 i - L_3 i^3 + L_5 i^5. \tag{12}$$

Wprowadzając operator D do równania (11) oraz uwzględniając wyrażenie (12) otrzymuje się:

$$\mathscr{L}Di + Ri + L_1 Di - L_3 Di^3 + L_5 Di^5 = E. \tag{13}$$

Po wprowadzeniu oznaczenia $\frac{R}{\mathscr{L}+L_1} = b$ równanie (13) przyjmuje postać

$$(D+b)i - \frac{L_3 \cdot b}{R} Di^3 + \frac{L_5 b}{R} Di^5 = \frac{Eb}{R}. \tag{14}$$

Jeżeli zgodnie z (1) przyjąć:

$$w_1 = D+b, \qquad w_4 = 0,$$

$$w_2 = 0, \qquad w_5 = \frac{L_5 b}{R} D,$$

$$w_3 = \frac{-L_3 b}{R} D, \qquad k = 1$$

oraz

$$V(t) = \frac{Eb}{R}, \tag{15}$$

to wykorzystując związki (3), (5), (7), (8) i (9) otrzymamy układ liniowy równań różniczkowych.

Z zależności (3)

$$(D+b) i_1 = \frac{Eb}{R}. \tag{16}$$

Zachowując warunek początkowy $i_1(t) = 0$ dla $t = 0$ i wprowadzając przekształcenie Laplace'a-Carsona otrzymujemy:

$$(p+b) i_1(p) = \frac{Eb}{R}. \tag{17}$$

Z odwrotnego przekształcenia Laplace'a-Carsona otrzymuje się [6]

$$i_1(t) = \frac{E}{R}(1-e^{-bt}). \tag{18}$$

Z zależności (5) wynika, że następne liniowe równanie różniczkowe ma postać

$$i_2(t) = 0, \quad \text{gdyż} \quad w_2 = 0.$$

Z wyrażenia (7) otrzymuje się

$$w_1 i_3 = -(2w_2 A_1 A_2 + w_3 A_1^3). \tag{19}$$

Ponieważ $w_2 = 0$, więc

$$w_1 i_3 = -w_3 A_1^3.$$

Wykorzystując zależność (15) dochodzi się do równania

$$(D+b) i_3 = \frac{L_3 b}{R} D i_1^3 = \frac{L_3 b}{R} D \frac{E}{R}(1-e^{-bt})^3 =$$

$$= L_3 b \frac{E^3}{R^4} D(1-3e^{-bt}+3^{-2bt}-e^{-3bt}). \tag{20}$$

Stosując zaś przekształcenie Laplace'a-Carsona

$$i_3(p) = \frac{L_3 b E^3}{R^4} \frac{p}{p+b}\left(1 - \frac{3p}{p+b} + \frac{3p}{p+2b} - \frac{p}{p+3b}\right). \tag{21}$$

Z odwrotnego przekształcenia Laplace'a-Carsona otrzymuje się po uproszczeniach rozwiązanie $i_3(t)$

$$i_3(t) = \frac{L_3 bE^3}{R^4}(-4{,}5e^{-bt}+3bte^{-bt}+6e^{-2bt}-1{,}5e^{-3bt}). \tag{22}$$

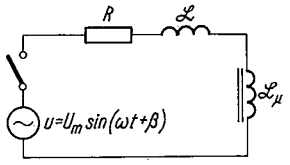
Podobnie wyznacza się następne wyrazy równania (14) i otrzymuje poszukiwane jego rozwiązanie przy założonej aproksymacji (12) funkcji nieliniowej.

Rozwiązanie $i(t)$ po uporządkowaniu ma postać

$$\begin{aligned} i(t) = i_1(t)+i_3(t)+i_5(t) &= \frac{E}{R}(1-e^{-bt})+ \\ &+\frac{L_3 b^3 E}{R^4}(-4{,}5e^{-bt}+3bte^{-bt}+6e^{-2bt}-1{,}5e^{-3bt})+ \\ &+9\frac{L_5 b^2 E^5}{R^7}(-2{,}625e^{-bt}+2{,}5bte^{-bt}-0{,}5b^2t^2e^{-bt}+8e^{-2bt}+ \\ &-4bte^{-2bt}-8{,}75e^{-3bt}+4e^{-4bt}+1{,}5bte^{-3bt}-0{,}625e^{-5bt})+ \\ &-\frac{L_5 bE^5}{R^6}(-10{,}41e^{-bt}+5bte^{-bt}+20e^{-2bt}-15e^{-3bt}+\frac{20}{3}e^{-4bt}-\frac{5}{4}e^{-5bt}). \end{aligned} \tag{23}$$

3.2. Elektryczny nieliniowy obwód prądu przemiennego

Rozważania dotyczące nieliniowych obwodów prądu przemiennego ograniczono także do obwodu szeregowego (rys. 2).



Rys. 2. Badany obwód prądu przemiennego

Równanie różniczkowe obwodu ma postać

$$\mathscr{L}\frac{di}{dt}+Ri+z\frac{d\varphi}{dt} = U_m \sin(\omega t+\beta), \tag{24}$$

gdzie:

$\mathscr{L}$ — indukcyjność,
R — rezystancja,
$\mathscr{L}_\mu$ — cewka z rdzeniem ferromagnetycznym,
z — liczba zwojów cewki z rdzeniem ferromagnetycznym,
φ — strumień magnetyczny cewki,
U — wartość maksymalna sinusoidalnego napięcia wymuszającego,
β — kąt przesunięcia fazowego.

Założono przy tym, że:

1° $i(t) = 0$ dla $t = 0$,

2° rdzeń dławika nie posiada magnetyzmu szczątkowego.

Charakterystykę nieliniowego elementu α_μ dławika z rdzeniem ferromagnetycznym aproksymowano jednowartościową funkcją o postaci

$$\psi = z\varphi = L_1 i - L_3 i^3. \tag{25}$$

Podstawiając (25) do (24) i wprowadzając operator $D = \frac{d}{dt}$ otrzymuje się równanie obwodu przy przyjętej aproksymacji (25)

$$\mathscr{L} Di + Ri + L_1 Di - L_3 Di^3 = U_m \sin(\omega t + \beta). \tag{26}$$

Jeżeli wprowadzić oznaczenie $\frac{R}{\mathscr{L} + L_1} = b$, to równanie (25) przybiera postać

$$(D+b)i - \frac{bL_3}{R} Di^3 = \frac{U_m b}{R} \sin(\omega t + \beta). \tag{27}$$

Porównując odpowiadające sobie składniki równań (27) i (1) otrzymuje się następujące związki:

$$w_1 = D + b, \qquad w_3 = -\frac{bL_3}{R} D,$$

$$w_2 = 0, \qquad k = 1,$$

$$V(t) = \frac{U_m b}{R} \sin(\omega t + \beta). \tag{28}$$

Rozwiązanie równania (27) poszukuje się w postaci (2).

Zgodnie z zasadami metody rewersyjnej pierwsze z liniowych równań składowych wynika z zależności (3)

$$(D+b) A_1 = \frac{U_m b}{R} \sin(\omega t + \beta). \tag{29}$$

Po przekształceniu zaś

$$(D+b) A_1 = \frac{U_m b}{R} (\sin\omega t \cos\beta + \cos\omega t \cdot \sin\beta). \tag{30}$$

Ponieważ współczynnik $k = 1$, więc z równania (2) wynika, że $A_1 = i_1(t)$ przy warunku początkowym $i_1(t) = 0$ dla $t = 0$.

Jeżeli uwzględnić zależności (4) i (28) oraz zastosować przekształcenie Laplace'a-Carsona, to

$$i_1(p) = \frac{U_m b}{R} \cdot \frac{1}{p+b} \frac{p}{\omega^2 + p^2} \cos\beta + \frac{p^2}{\omega^2 + p^2} \sin\beta = \frac{U_m b}{R} \frac{p\omega\cos\beta + p^2 \sin\beta}{(p+b)(p^2+\omega^2)}. \tag{31}$$

Przekształcenie odwrotne $\mathscr{C}^{-1}$ wyrażenia (31) ma postać

$$i_1(t) = \frac{U_m b}{R} \frac{\omega\cos\beta - b\sin\beta}{\omega^2+b^2} e^{-bt} + \frac{\omega\cos\beta - j\omega\sin\beta}{2(-\omega^2-j\omega b)} e^{-j\omega t} +$$

$$+ \frac{\omega\cos\beta + j\omega\sin\beta}{2(-\omega^2+j\omega b)} e^{-j\omega t}. \tag{32}$$

Wprowadzając do (32) oznaczenia: $Z = \sqrt{\omega^2+b^2}$ oraz $\operatorname{tg}\Theta = \frac{\omega}{b}$ otrzymano:

$$i_1(t) = \frac{U_m b}{RZ} [\sin(\Theta-\beta)e^{-bt} + \sin(\omega t+\beta-\Theta)]. \tag{33}$$

Ponieważ $z_2 = 0$ — por. (28) — więc kolejnym składowym równaniem liniowym zgodnie z (7) oraz (27) jest

$$(D+b)A_3 = \frac{U_m^3 L_3}{Z_3(\mathscr{L}+L_1)^4} D[\sin(\Theta-\beta)e^{-bt} + \sin(\omega t+\beta-\Theta)]. \tag{34}$$

Ponieważ współczynnik $k = 1$, więc z równania (2) wynika, że $A_3 = i_3(t)$ przy warunku początkowym $i_3(t) = 0$ dla $t = 0$.

Wprowadzając nowe oznaczenia można prąd $i_3(t)$ wyrazić równaniem:

$$i_3(t) = \frac{U_m^3 L_3}{Z^3(\mathscr{L}+L_1)^4} \frac{D}{(D+b)} \cdot Q(t), \tag{35}$$

które po przekształceniu operatorowym ma postać

$$i_3(p) = \frac{U_m^3 L^3}{Z^3(\mathscr{L}+L_1)^4} \frac{p}{(p+b)} \cdot Q(p), \tag{36}$$

przy czym $Q(p) = \mathscr{C}[Q(t)]$.

Jeżeli do wyrażenia (36) zastosować wzór na transformatę całki Duhamela, wtedy przekształcenie odwrotne wyraża się splotem funkcji [6]

$$i_3(t) = \frac{U_m^3 L_3}{Z^3(\mathscr{L}+L_1)^4} \left[\frac{d}{dt} Q(t) * e^{-bt}\right], \tag{37}$$

gdyż $\mathscr{C}^{-1}\left[\frac{p}{p+b}\right] = e^{-bt}$.

W wyrażeniu (37) należy teraz zastosować wzór na pochodną splotu funkcji [6]

$$i_3(t) = \frac{U_m^3 L_3}{Z^3(\mathscr{L}+L_1)^4} \frac{d}{dt} \int_0^t Q(u)e^{-b(t-u)}du =$$

$$= \frac{U_m^3 L_3}{Z_3(\mathscr{L}+L_1)} \frac{d}{dt} \int_0^t e^{-b(t-u)}[\sin(\Theta-\beta)e^{-bu} + \sin(\omega t+\beta-\Theta)]^3 du. \tag{38}$$

Występującą w zależności (38) całkę można przedstawić w postaci sumy następujących czterech składników:

$$I_1 = e^{-bt}\sin^3(\Theta-\beta)\int_0^t e^{-2bu}du,$$

$$I_2 = 3e^{-bt}\sin^2(\Theta-\beta)\int_0^t e^{-bu}\sin(\omega u+\beta-\Theta)du,$$

$$I_3 = 3e^{-bt}\sin(\Theta-\beta)\int_0^t \sin^2(\omega u+\beta-\Theta)du,$$

$$I_4 = e^{-bt}\int_0^t e^{bu}\sin^3(\omega u+\beta-\Theta)du. \tag{39}$$

Po rozwiązaniu tych całek [7], a następnie zróżniczkowaniu względem czasu, otrzymuje się po uporządkowaniu poszukiwane rozwiązanie $i(t)$:

$$\begin{aligned} i(t) = i_1(t)+i_3(t) = {} & \frac{U_m b}{RZ}[\sin(\Theta-\beta)e^{-bt}+\sin(\omega t+\beta-\Theta)]+ \\ & + \frac{U_m^3 L_3}{Z^3(\mathscr{L}+L_1)^4}\left\{\left[\frac{\sin^3(\Theta-\beta)}{2}(3e^{-3bt}-e^{-bt})+\right.\right. \\ & - \frac{3\sin^2(\Theta-\beta)}{Z^2}\{[b^2\sin(\beta-\Theta)+b\omega\cos(\beta-\Theta)]e^{-bt}+ \\ & - [2b^2\sin(\omega t+\beta-\Theta)+\omega^2\sin(\omega t+\beta-\Theta)+b\omega\cos(\omega t+\beta-\Theta)]e^{-2bt}\}+ \\ & + \frac{3\sin(\Theta-\beta)}{4\omega}\cdot[b\sin 2(\omega t+\beta-\Theta)-2\omega\cos 2(\omega t+\beta-\Theta)]e^{-bt}+ \\ & + \frac{3\sin(\Theta-\beta)}{2}[1-bt]e^{-bt}-\frac{3b}{4\omega}\sin(\Theta-\beta)\sin 2(\beta-\Theta)e^{-bt}+ \\ & + \frac{\omega}{b^2+9\omega^2}[3b\sin^2(\omega t+\beta-\Theta)\cos(\omega t+\beta-\Theta)+ \\ & + 3\omega\sin^3(\omega t+\beta-\Theta)-6\omega\sin(\omega t+\beta-\Theta)\cdot\cos^2(\omega t+\beta-\Theta)]+ \\ & + \frac{6\omega^3}{Z^2(b^2+9\omega^2)}[b\cos(\omega t+\beta-\Theta)+\omega\sin(\omega t+\beta-\Theta)]+ \\ & - \frac{b\omega}{b^2+9\omega^2}\sin^2(\beta-\Theta)\cdot\left[3\cos(\beta-\Theta)-\frac{b}{\omega}\sin(\beta-\Theta)\right]e^{-bt}+ \\ & \left.\left. + \frac{6\omega^3 b}{Z^2(b^2+9\omega^2)}\left[\frac{b}{\omega}\sin(\beta-\Theta)-\cos(\beta-\Theta)\right]e^{-bt}\right]\right\}. \end{aligned} \tag{40}$$

Z zależności (40) otrzymuje się rozwiązanie $i(t)$ równania (25) także wówczas, gdy funkcja wymuszająca ma postać $u = U_m \sin\omega t$, tj. przy $\beta = 0$:

$$
\begin{aligned}
i(t) = {} & \frac{U_m b}{RZ}\left[\sin\Theta e^{-bt} + \sin(\omega t - \Theta)\right] + \frac{U_m^3 L_3}{Z^3(\mathscr{L} + L_1)^4}\left\{\left\{\frac{\sin^3\Theta}{2}(3e^{-3bt} - e^{-bt}) + \right.\right. \\
& + \frac{3\sin^2\Theta}{Z^2}(b^2\sin\Theta - b\omega\cos\Theta)e^{-bt} + \\
& + \frac{3\sin^2\Theta}{Z^2}\left[2b^2\sin(\omega t - \Theta) + \omega^2\sin(\omega t - \Theta) + b\omega\cos(\omega t - Q)\right]e^{-2bt} + \\
& + 3\sin\Theta\left\{\left[-\frac{1}{2}\cos^2(\omega t - \Theta) + \frac{1}{2}\sin^2(\omega t - \Theta) + \right.\right. \\
& \left.\left. + \frac{b}{2\omega}\cos(\omega t - \Theta)\sin(\omega t - \Theta)\right]e^{-bt} + \frac{b}{2\omega}\sin\Theta\cos\Theta e^{-bt}\right\} + \\
& + \frac{3}{2}\sin\Theta(1 - bt)e^{-bt} + \frac{\omega}{b^2 + 9\omega^2}\left[3b\sin^2(\omega t - \Theta)\cos(\omega t - \Theta) + \right. \\
& + 3\omega\sin^3(\omega t - \Theta) - 6\omega\sin(\omega t - \Theta)\cdot\cos^2(\omega t - \Theta)\Big] + \\
& + \frac{6\omega^3}{Z^2(b^2 + 9\omega^2)}\left[b\cos(\omega t - \Theta) + \omega\sin(\omega t - \Theta)\right] + \\
& - \frac{b}{b^2 + 9\omega^2}\sin^2\Theta\left[3\cos\Theta + \frac{b}{\omega}\sin\Theta\right]e^{-bt} + \\
& \left.\left. + \frac{6\omega^3 b}{Z^2(b^2 + 9\omega^2)}\left[-\frac{b}{\omega}\sin\Theta - \cos\Theta\right]e^{-bt}\right\}\right\}. \qquad (41)
\end{aligned}
$$

4. ZAKOŃCZENIE

Oprócz poszukiwania rozwiązań analitycznych w ramach tej pracy wykonano także badania eksperymentalne w celu sprawdzenia dokładności obliczeń. Na podstawie uzyskanych wyników można sformułować następujące wnioski:

— dokładność obliczeń w porównaniu z pomiarami zależy głównie od dokładności aproksymacji, charakterystyki elementu nieliniowego;

— zależności analityczne (23), (40), (41) są skomplikowane; omawiana metoda może być praktycznie stosowana przy użyciu elektronicznych maszyn cyfrowych;

— metoda rewersyjna pozwala na rozwiązanie obwodów nieliniowych, w których oprócz rozważanych w tej pracy elementów, występuje także pojemność połączona szeregowo;

— rozwiązanie analityczne (40) równania (24) umożliwia analizę rozważanych obwodów nieliniowych prądu przemiennego przy dowolnie wybranej fazie napięcia wymuszającego;

— zastosowanie wielomianu aproksymującego piątego stopnia w przypadku obwodu prądu stałego, pozwala na analizę tych obwodów przy dużej nieliniowości charakterystyki $\psi = f(i)$.

LITERATURA CYTOWANA W TEKŚCIE

1. L. A. Pipes, The Reversion Method for Solving Nonlinear Differential Equations, Journal of Applied Physics, Vol. 23, No 2, 1952.
2. Van Ostrand, Phil. Mag. 19, s. 366, 1910.
3. G. I. Cohn, B. Salzberg, Solution of Nonlinear Differential Equations by the Reversion Method, Journal of Applied Physics, Vol. 24, No 2, 1953.
4. G. I. Cohn, B. Saltzberg, Solution of Nonlinear Differential Equations by the Reversion Method Through the First Thirteen Terms, Journal of Applied Physics, Vol. 25, No 2, 1954.
5. P. J. Nowacki, Obliczanie nieliniowych obwodów elektrycznych i magnetycznych, PWN, Warszawa 1959.
6. J. Osiowski, Zarys rachunku operatorowego, WNT, Warszawa 1965.
7. I. M. Ryżik, I. S. Gradsztajn, Tablicy integrałow, sum, rjadow i proizwedenii, Izd. Nauka, Moskwa 1962.

Z. RYDZEWSKI

APPLICATION OF THE METHOD FOR THE INVESTIGATION OF NONLINEAR ELECTRICAL CIRCUITS

Summary

In the paper are given solutions of nonlinear differential equations concerning series circuits which contain linear and nonlinear inductances. In order to obtain these solutions the reversion method has been used. For a *d*—*c* circuit the nonlinear functions have been approximated by aid of polynoms with odd powers (including the 5-th power). In the case of *a*—*c* circuit there has been obtained the solution $i(t)$ for an optional phase of forcing voltage by aid of third degree approximating polynom.

Z. RYDZEWSKI

APPLICATION DE LA MÉTHODE DE REVERSION DANS L'ANALYSE DES CIRCUITS ÉLÉCTRIQUES NONLINEARES

Résumé

On a présenté dans cet article les solutions des équations différentielles nonlinéaires des circuits électriques series à inductance linéaire et nonlineaire. Pour obtenir ces solutions la méthode de reversion a été utilisée.

Pour les circuits de courant continu la fonction nonlineaire a été approximée à l'aide des polynômes à puissances impaires (la 5-me puissance inclusivement).

Dans le cas des circuits de courant alternatif les solutions $i(t)$ ont été obtenues pour une phase quelconque de tension forcante à l'aide d'une approximation polynômique de troisième degré.

Z. RYDZEWSKI

ANWENDUNG DER REVERSIONS-METHODE IN UNTERSUCHUNG VON NICHTLINEAREN ELEKTRISCHEN KREISEN

Zusammenfassung

In der Arbeit sind Lösungen nichtlinearer Differentialgleichungen für Kreise in Reihenschaltung mit linearen und nichtlinearen Induktivitäten angeführt. Diese Lösungen wurden nach der Reversions-Methode ermittelt. Für den Gleichstromkreis wurde die nichtlineare Funktion bei Anwendung von Polynomen in ungeraden Potenzen (bis 5-ter Potenz) approximiert.

Im Fall eines Wechselstromkreises erhielt man die Lösung $i(t)$ für beliebig gewählte Phase der Zwangsspannung mit Hilfe von Approximations-Polynomen in dritter Potenz.

З. РЫДЗЕВСКИ

ПРИМЕНЕНИЕ РЕВЕРСИВНОГО МЕТОДА К ИССЛЕДОВАНИЮ НЕЛИНЕЙНЫХ ЭЛЕКТРИЧЕСКИХ ЦЕПЕЙ

Р е з ю м е

В статье приводятся решения нелинейных дифференциальных уравнений для последовательных цепей с линейными и нелинейными индуктивностями, полученных реверсивным методом.

Для цепи постоянного тока нелинейная функция аппроксимировалась полиномом с нечетными показателями степени до пятой включительно.

В случае цепи переменного тока было получено решение $i(t)$ для произвольно выбранной фазы вынуждающего напряжения при использовании аппроксимационного полинома третьей степени.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 1, ss. 15-40

621.317.66

Metody wyznaczania obszarów sprawności

JAN ZABRODZKI (WARSZAWA), STANISŁAW BUDKOWSKI (WARSZAWA)

Instytut Maszyn Matematycznych Politechniki Warszawskiej

Otrzymano 26.5.1973

W pracy omówiono metody wyznaczania obszarów sprawności. Rozważane metody podzielono na symulacyjne i eksperymentalne. W każdej z tych grup wyróżniono metody przeszukiwania przestrzeni parametrów: systematyczne i losowe oraz metody bezpośredniego wyznaczania brzegu obszaru sprawności.

1. WSTĘP

W nowoczesnych metodach projektowania i kontroli obiektów istotną rolę odgrywają tzw. obszary sprawności, czyli zbiory punktów w przestrzeni parametrów, dla których spełnione są warunki gwarantujące poprawność działania obiektów. Znajomość obszaru sprawności pozwala z jednej strony na wybór optymalnego punktu pracy projektowanego obiektu z uwzględnieniem aspektów niezawodnościowych (rozrzuty wartości elementów, starzenie się elementów, różne warunki pracy) i problemu maksymalnego uzysku produkcyjnego (rozrzuty technologiczne i produkcyjne), a z drugiej strony na dokładną kontrolę eksploatowanego obiektu [36].

Niezbędnym warunkiem dla pełnego wykorzystania możliwości oferowanych przez koncepcję obszarów sprawności jest opracowanie metod ich wyznaczania. Omówieniu tego problemu poświęcone jest niniejsze opracowanie.

W pracy, po wprowadzeniu podstawowych pojęć związanych z obszarami sprawności, omówione zostały różne metody wyznaczania tych obszarów. Rozważane metody podzielono na symulacyjne i eksperymentalne. W każdej z tych grup wyróżniono metody przeszukiwania przestrzeni parametrów systematyczne i losowe oraz metody bezpośredniego wyznaczania brzegu obszaru sprawności.

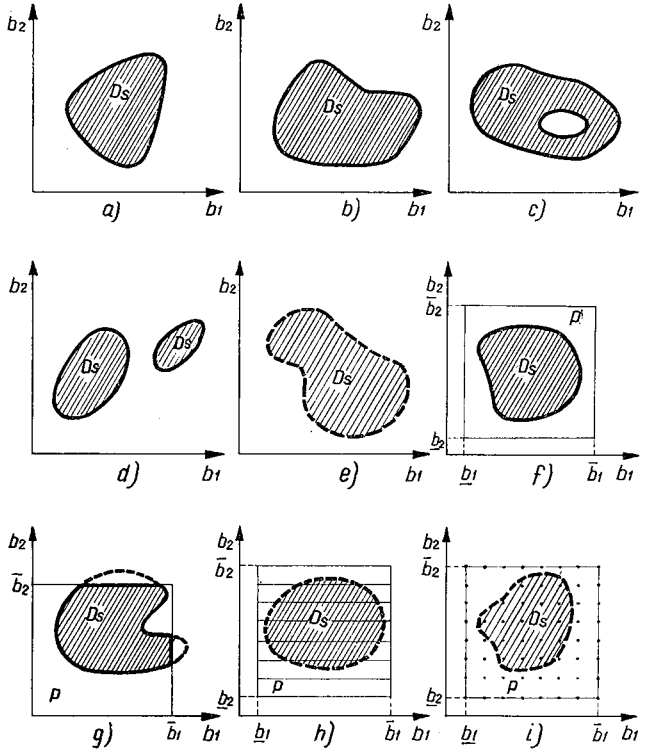
W pracy podano możliwie pełny przegląd metod wyznaczania obszarów sprawności, co umożliwia ich porównanie oraz wybór metody najlepszej dla konkretnego przypadku. Niektóre z omawianych metod zostały opracowane i były praktycznie wykorzystane przez autorów. Część rozważanych metod była opracowana i stosowana przez innych autorów. Pozostałe metody są potencjalnie możliwymi do wykorzystania — są to bądź nowe metody, bądź metody znane w innych dziedzinach, zaadoptowane do celów wyznaczania obszarów sprawności.

2. OBSZAR SPRAWNOŚCI

Dowolny obiekt można scharakteryzować podając szereg istotnych w rozważanym zagadnieniu parametrów oraz wiążących je zależności o postaci:

$$y_j = f_j(b_1, b_2, \ldots b_i, \ldots, b_n) \quad j = 1, 2, \ldots, m, \tag{1.1}$$

gdzie y_j są parametrami wyjściowymi reprezentującymi istotne cechy użytkowe obiektu, a b_i są parametrami charakteryzującymi elementy składowe obiektu, wielkości oddziaływujące na wejścia obiektu oraz warunki otoczenia, w których obiekt pracuje.



Rys. 1. Przykłady płaskich obszarów sprawności

a) obszar wypukły, b) obszar wklęsły, c) obszar dwuspójny, d) obszar niespójny, e) otwarty obszar sprawności, f) ograniczona przestrzeń parametrów P, g) obszar sprawności z ograniczeniem, h) przestrzeń parametrów i obszar sprawności gdy jeden parametr przyjmuje wartości dyskretne, i) dyskretny obszar sprawności

Pojęcie sprawności obiektu[1)] jest ściśle związane z wymaganiami jakie narzuca się na jego cechy. Zbiór tych wymagań, zwany też układem więzów, jest zestawem warunków (kryteriów) sprawności przyjmujących postać:

$$\underline{y}_j \leqslant y_j = f_j(b_1, b_2, \ldots, b_n) \leqslant \overline{y}_j, \quad j = 1, 2, \ldots, m, \tag{1.2}$$

gdzie $\underline{y}_j$ i $\overline{y}_j$ oznaczają odpowiednio kres dolny i górny dopuszczalnych wartości parametrów y_j.

[1)] Obiekt jest sprawny, gdy jest zdolny do pełnienia swych funkcji lub pełni swe funkcje zgodnie z wymaganiami. W przeciwnym wypadku obiekt jest niesprawny [23].

Można zauważyć, że przy określonych warunkach sprawności obiekt dla pewnych kombinacji wartości parametrów b_i jest sprawny a dla innych nie. Określoną kombinację wartości parametrów b_i (tzw. punkt pracy obiektu) można interpretować jako punkt w n-wymiarowej przestrzeni kartezjańskiej, nazywanej dalej przestrzenią parametrów P. Zbiór punktów, dla których obiekt spełnia odpowiednie wymagania można interpretować jako n-wymiarowy obszar w przestrzeni parametrów zwany obszarem sprawności.

Zależnie od postaci warunków sprawności obszar sprawności może być domknięty (np. rys. 1a, b) albo otwarty (rys. 1e)[2].

Zależnie od rodzaju funkcji f_j oraz od przyjętych warunków sprawności, obszar sprawności może być jednospójny wypukły (rys. 1a), wklęsły (rys. 1b), albo wielospójny (rys. 1c). Mogą wystąpić również obszary niespójne (rys. 1d).

Ponieważ ze względów technologicznych lub konstrukcyjnych parametry obiektu (przynajmniej jeden) mogą przyjmować wartości jedynie w pewnym skończonym przedziale, przestrzeń parametrów P jest przestrzenią ograniczoną (rys. 1f). Często zakres dopuszczalnych wartości parametrów jest ograniczany dodatkowo ze względu na konieczność zabezpieczania obiektu przed uszkodzeniem (np. wartość napięcia kolektor-emiter w tranzystorze nie może przekroczyć wartości napięcia przebicia). Ograniczenie przestrzeni parametrów może prowadzić do ograniczenia obszaru sprawności. Przykład obszaru sprawności z ograniczeniem pokazany jest na rys. 1g.

Jeżeli przedziały wartości parametrów są przedziałami ciągłymi, to przestrzeń parametrów będąca iloczynem kartezjańskim tych przedziałów ma strukturę ciągłą. Jeżeli jeden lub więcej parametrów może przyjmować jedynie wartości dyskretne z określonego przedziału (np. ze względu na normalizację), przestrzeń parametrów ma budowę nieciągłą. W przypadku, gdy wszystkie istotne parametry mogą przyjmować jedynie wartości dyskretne z odpowiednich przedziałów, przestrzeń parametrów ma strukturę dyskretną. Na rys. 1h pokazano płaski obszar sprawności, w przypadku gdy jeden z parametrów przyjmuje wartości dyskretne, a na rys. 1i obszar sprawności o budowie dyskretnej.

3. METODY SYMULACYJNE WYZNACZANIA OBSZARU SPRAWNOŚCI

Metody symulacyjne wyznaczania obszaru sprawności można podzielić na metody oparte o przeszukiwanie przestrzeni parametrów i metody bezpośredniego wyznaczania brzegu obszaru sprawności.

3.1. Wyznaczanie obszaru sprawności metodami przeszukiwania przestrzeni parametrów

Wyznaczanie obszaru sprawności polega na odpowiednim wyborze kolejnych punktów przestrzeni parametrów, bądź jej fragmentu i sprawdzaniu, czy dla odpowiadających tym punktom wartości parametrów spełnione są nierówności (1.2). Jeżeli tak, to uważa się, że

[2] W dalszym ciągu pracy mówiąc o obszarze sprawności będziemy mieli na myśli obszar domknięty, co w niczym nie ogranicza ogólności rozważań.

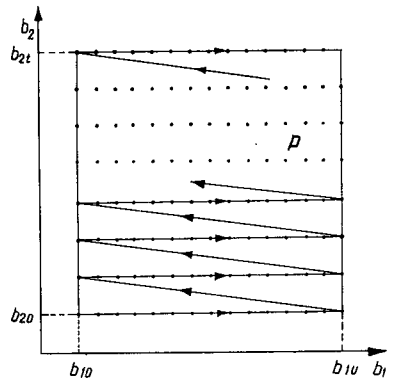
dany punkt należy do obszaru sprawności. Ze względu na dużą liczbę wykonywanych obliczeń metody te wymagają korzystania z pomocy maszyn cyfrowych.

Poszczególne metody przeszukiwania przestrzeni parametrów różnią się sposobem wyboru kolejnych punktów. Wyróżnia się dwie grupy metod: metody systematyczne i metody liczb losowych i złotych.

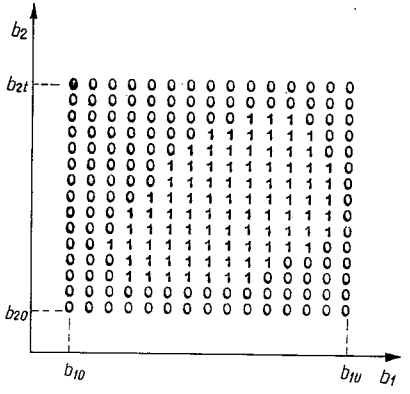
2.1.1. *Metody systematyczne*

W metodach systematycznych wyznaczania obszaru sprawności konieczna jest dyskretyzacja przestrzeni parametrów. Uzyskane punkty dyskretne są wybierane w odpowiedniej kolejności.

Najprostszy sposób przeszukiwania przestrzeni parametrów polega na kolejnym wybieraniu wszystkich punktów dyskretnej przestrzeni parametrów, tak jak to pokazano na rys. 2. Taki sposób przeszukiwania był wykorzystywany przez A. Sielickiego do wyznaczania obszarów sprawności układów logicznych budowanych z elementów dyskretnych [26, 27]. Wyniki obliczeń wyprowadzane były z maszyny (UMC-1) w postaci dwuwymia-



Rys. 2. Przykład systematycznego przeszukiwania przestrzeni parametrów



Rys. 3. Przykład macierzy sprawności

rowych tablic tzw. macierzy sprawności, których elementami są jedynki, w przypadku gdy odpowiedni punkt należy do obszaru sprawności, i zera, w przypadku przeciwnym. Przykład macierzy sprawności pokazany jest na rys. 3. Można podać oczywiście wiele innych wariantów systematycznego przeszukiwania parametrów (np. metoda wzrastających odchyłek [2], w której wychodząc ze środkowego punktu badanej przestrzeni parametrów wybiera się kolejne punkty tak, że tworzą one linię łamaną obiegającą punkt początkowy podobnie jak spirala). Warianty te z punktu widzenia wyznaczania obszaru sprawności nie wnoszą nic nowego i stąd nie będą one rozważane dokładniej.

Natomiast znacznie ciekawsze są metody przeszukiwania systematycznego, które pozwalają ograniczyć liczbę sprawdzanych punktów. W pracy [30] podany jest algorytm (tzw. algorytm *B*), w którym wstępnie zawęża się przestrzeń parametrów do prostokąta zbudo-

wanego w ten sposób, że jego elementami są między innymi wszystkie punkty obszaru sprawności (rys. 4). Kroki algorytmu są następujące:

a) rozwiązując kolejne pary równań spośród równań

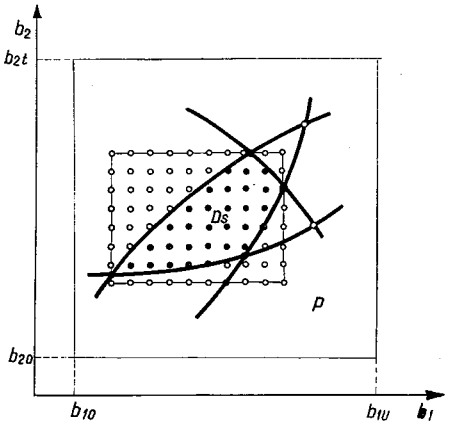
$$\begin{aligned} \underline{y}_j &= f_j(b_1, b_2 \ \ldots, b_n), \\ \overline{y}_j &= f_j(b_1, b_2, \ldots, b_n), \end{aligned} \qquad j = 1, 2, \ldots, m, \tag{1.3}$$

wyznacza się współrzędne punktów przecięcia się krzywych brzegowych obszaru sprawności;

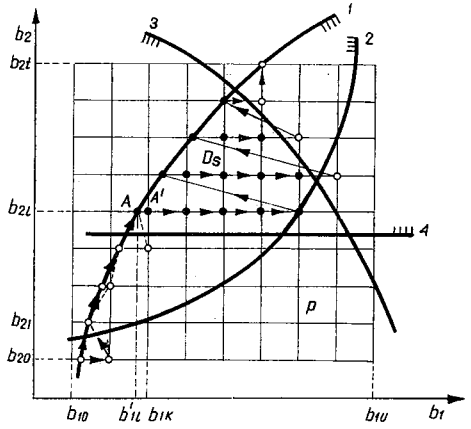
b) spośród punktów wyznaczonych w kroku a) wybiera się te, które należą do brzegu obszaru sprawności;

c) w przestrzeni parametrów wyznacza się najmniejszy prostokąt mający tę własność, że obejmuje wszystkie punkty znalezione w kroku b);

d) sprawdza się kolejno punkty należące do prostokąta wyznaczonego w kroku c) i wyprowadza się wyniki w postaci macierzy sprawności.



Rys. 4. Ilustracja algorytmu *B*



Rys. 5. Ilustracja algorytmu *D*

Algorytm jest słuszny przy założeniu, że funkcje (1.3) opisujące krzywe brzegowe są funkcjami monotonicznymi.

W niektórych wypadkach można uzyskać dalsze zmniejszenie czasu obliczeń stosując algorytm, którego działanie ilustruje rys. 5 (algorytm *D* w pracy [30]). Kolejne kroki algorytmu są następujące:

a) dla wartości b_{2l}, ($l = 0, 1, 2, \ldots, t$) parametru b_2, z jednego z równań (1.3) wyznacza się wartość b'_{1l} parametru b_1, uzyskując współrzędne punktu leżącego na odpowiedniej krzywej brzegowej (krzywa 1 na rys. 5);

b) sprawdza się, czy punkt otrzymany w kroku a) należy do obszaru sprawności (nierówności (1.2)). Jeżeli tak, przechodzi się do kroku c). Jeżeli nie, powtarza się etapy a) i b) dla kolejnej wartości $b_{2(l+1)}$;

c) otrzymaną wartość b'_{1l} zastępuje się najbliższą większą wartością b_{1k} ($k = 0, 1, \ldots, \ldots, u$), (punkty A i A' na rys. 1.5);

d) sprawdza się czy punkt o współrzędnych (b_{2l}, $b_{1(k+1)}$) spełnia nierówności (1.2). Jeżeli tak, to krok ten jest powtarzany. Proces zostaje przerwany z chwilą gdy kolejny punkt

nie spełnia nierówności (1.2), bądź gdy $k = u$. Wtedy wartość parametru b_{2l} zwiększa się do kolejnej wartości $b_{2(l+1)}$ i o ile $(l+1) \leqslant t$ powraca się do kroku a) i kontynuuje się obliczenia dla współrzędnej $b_{2(l+1)}$. Współrzędne punktów spełniających nierówności (1.2) są drukowane.

Algorytm może być stosowany jedynie wtedy, gdy z góry znany jest charakter funkcji (1.3) opisujących krzywe brzegowe. Poprawna praca algorytmu zależy od właściwego wyboru krzywej brzegowej w kroku a). Na przykład algorytm nie będzie działał prawidłowo gdy wybierze się krzywe 2 albo 3 z rys. 5. Poza tym, sam algorytm zależy od tego, czy obszar sprawności jest domknięty czy otwarty. W przypadku otwartego obszaru sprawności należy wykonywać dodatkowe kroki, tak jak to zaznaczono na rys. 5 linią przerywaną (po kroku b) zawsze trzeba przejść do kroków c) i d)).

3.1.2. *Metody liczb losowych i złotych*

W metodach liczb losowych współrzędne kolejnych punktów przestrzeni parametrów wybierane są w ten sposób, że stanowią kolejne wartości ciągu liczb pseudolosowych generowanych w sposób programowy (na maszynie cyfrowej) [38] lub układowy [11].

Znana jest również metoda [6, 7, 9], w której do wybierania kolejnych punktów przestrzeni parametrów wykorzystuje się liczby złote.

Algorytm generacji współrzędnych kolejnych punktów próbkowych oparty jest o własności liczb niewymiernych (w szczególności jest nią liczba złota z, której wartość wynika ze złotego podziału odcinka i można ją określić z równania $1 : z = z : (1-z)$. Tworząc ciąg o postaci

$$Z_h = h \cdot z - E(h \cdot z); \quad \text{E} — \text{funkcja część całkowita}, \quad h = 1, 2, \ldots,$$

pokazano [32], że liczby Z_h rozmieszczone są w przybliżeniu równomiernie na odcinku [0, 1]. Natomiast rozwijając dwójkowo liczby Z_h i tworząc z tego rozwinięcia N nowych rozwinięć w ten sposób, że każde z nich powstaje z rozwinięcia Z_h przez wybranie co N-tych bitów, uzyskuje się N-współrzędnych liczby w kostce N-wymiarowej $[0, 1]^N$. Pokazano [10], że rozmieszczenie tak określonych punktów jest w przybliżeniu równomierne w tej kostce.

Liczby złote nie spełniają statystycznych testów pseudolosowości [24] ale posiadają tę cenną własność, że gwarantują lepszą równomierność rozmieszczenia w kostce $[0, 1]^N$ względem liczb pseudolosowych (właśnie m.in. z tego względu nie spełniają niektórych testów pseudolosowości).

3.2. Metody wyznaczania brzegu obszaru sprawności

3.2.1. *Metoda graficzna*

Jedna z możliwych metod rozwiązania układu nierówności (1.2) polega na sprowadzeniu zagadnienia do rozwiązania układu równości

$$\begin{aligned} \underline{y}_j &= f_j(b_1, b_2, \ldots, b_n), \\ \overline{y}_j &= f_j(b_1, b_2, \ldots, b_n), \end{aligned} \qquad j = 1, 2, \ldots, m, \tag{1.3}$$

czyli do wyznaczenia hiperpowierzchni ograniczających obszar sprawności. Metodę tę w przypadku dwuwymiarowych obszarów sprawności wykorzystuje wielu autorów [4, 6, 9, 12, 15, 18, 30]. A. Sielicki nazywa ją metodą graficzną [30].

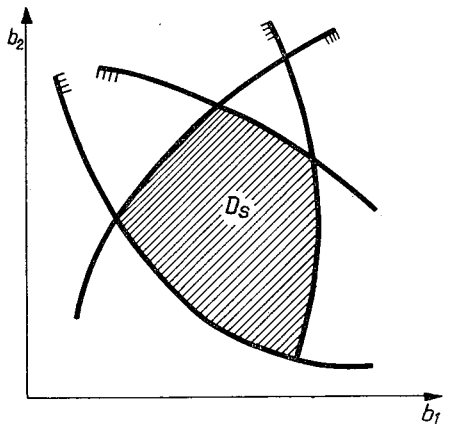
Kolejne etapy rozwiązania w metodzie graficznej są następujące:

a) zredukowanie liczby zmiennych w warunkach sprawności (1.2) do dwóch;

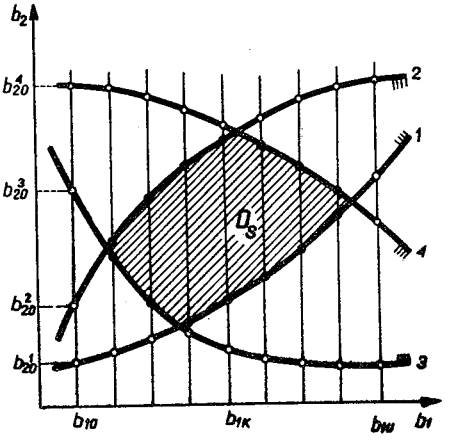
b) wyznaczenie z równań (1.3) krzywych brzegowych, z których każda dzieli przestrzeń parametrów na dwa rozłączne i dopełniające się zbiory;

c) określenie zredukowanego obszaru sprawności poprzez graficzne wyznaczenie koniunkcji zbiorów wyznaczonych w punkcie b).

Metodę graficzną ilustruje rys. 6.



Rys. 6. Ilustracja metody graficznej



Rys. 7. Ilustracja algorytmu *C*

Znany jest algorytm pozwalający wyznaczyć obszar sprawności metodą graficzną przy wykorzystaniu maszyny cyfrowej (algorytm C w pracy [30]). Kolejne kroki algorytmu są następujące:

a) koduje się znaki nierówności we wszystkich warunkach sprawności (1.2);

b) przedział zmian wartości jednego z parametrów (np. b_1) dzieli się na u części. Uzyskane wartości b_{1k} ($k = 0, 1, \ldots, u$) podstawia się do równań (1.3), co pozwala wyznaczyć wartości drugiego parametru b_{2k}^p, także punkty o współrzędnych (b_{1k}, b_{2k}^p) leżą na krzywych brzegowych p;

c) spośród punktów uzyskanych w kroku b) wybiera się te, które należą do brzegu sprawności. Wyboru dokonuje się w oparciu o badanie kodów znaków nierówności (1.2) i porównanie współrzędnych b_{2k}^p. Współrzędne wybranych punktów drukuje się. Algorytm ilustruje rys. 7.

3.2.2. *Metoda obchodzenia brzegu obszaru sprawności*[3)]

W metodzie obchodzenia brzegu obszaru sprawności przeszukuje się w sposób systematyczny kolejne punkty przestrzeni parametrów (bądź wybranego fragmentu przestrzeni

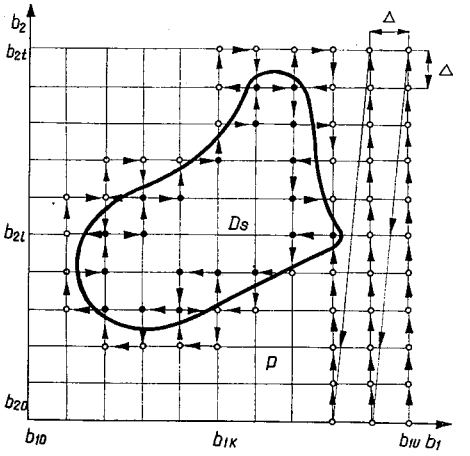
[3)] W pracy [22] metoda obchodzenia brzegu obszaru sprawności była wykorzystana do wyznaczania granic obszarów stabilności rozważanego obiektu.

parametrów) dotąd, dopóki nie zostanie znaleziony pierwszy punkt, dla którego spełnione są warunki sprawności (1.2). Następne punkty są wybierane jedynie z najbliższego „otoczenia" brzegu sprawności, zgodnie z następującym algorytmem (rys. 8).

Wartości parametrów b_{1k} i b_{2l} są zmieniane kolejno. W każdym kroku wartość parametru b_1 albo b_2 jest zwiększana albo zmniejszana o stałą wielkość Δ, przy czym

$$\Delta = b_{1(k+1)} - b_{1k} = b_{2(l+1)} - b_{2l}.$$

Jeżeli w czasie zmiany wartości parametru zostanie przekroczony brzeg obszaru sprawności i nowo otrzymany punkt znajduje się wewnątrz obszaru sprawności, to w następnych krokach (aż do ponownego wyjścia poza obszar sprawności) wartości parametrów zmieniane są tak, żeby punkt pracy obiektu poruszał się przeciwnie do ruchu wskazówek



Rys. 8. Ilustracja metody obchodzenia brzegu obszaru sprawności

zegara. Jeżeli natomiast po przekroczeniu brzegu obszaru sprawności punkt znalazł się poza obszarem, to w następnych krokach (aż do ponownego wejścia do obszaru sprawności) wartości parametrów zmieniane są zgodnie z kierunkiem ruchu wskazówek zegara. Przy każdym przekroczeniu brzegu obszaru drukuje się współrzędne punktu leżącego wewnątrz obszaru. Proces kończy się po wyznaczeniu całego obszaru sprawności.

W przypadku obszaru sprawności z ograniczeniem, punkt pracy porusza się wzdłuż linii ograniczającej w kierunku ruchu wskazówek zegara tak długo, aż ponownie zostanie przekroczony brzeg obszaru sprawności.

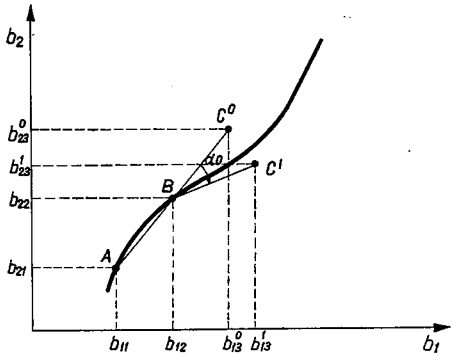
3.2.3. *Metoda wyznaczania kolejnych punktów brzegu obszaru sprawności*

Metoda wyznaczania kolejnych punktów brzegu obszaru sprawności podana została przez W. Esipowa w pracy [14]. W metodzie tej na podstawie znajomości współrzędnych dwóch punktów brzegu obszaru sprawności $A(b_{11}, b_{21})$ i $B(b_{12}, b_{22})$ wyznacza się przybliżone współrzędne trzeciego punktu $C^0(b^0_{13}, b^0_{23})$, zgodnie z regułą

$$b^0_{13} = 2b_{12} - b_{11},$$
$$b^0_{23} = 2b_{22} - b_{21}.$$

Geometrycznie oznacza to, że punkt C^0 wyznacza się na prostej łączącej punkty A i B w odległości równej odległości punktów A i B, licząc od punktu B (rys. 9).

Następnie sprawdza się czy punkt C^0 leży na brzegu, wewnątrz czy poza obszarem sprawności. W pierwszym przypadku punkt C^0 jest poszukiwanym punktem C brzegu obszaru i przystępuje się do wyznaczania następnego punktu (na podstawie znajomości punktów B i C).



Rys. 9. Ilustracja metody wyznaczania kolejnych punktów brzegu obszaru sprawności

W przypadku gdy punkt C^0 leży poza obszarem sprawności koryguje się jego współrzędne zgodnie z zależnościami

$$b_{13}^1 = b_{12} + (b_{13}^0 - b_{12})\cos\alpha_0 + (b_{23}^0 - b_{22})\sin\alpha_0,$$
$$b_{23}^1 = b_{22} - (b_{13}^0 - b_{12})\sin\alpha_0 + (b_{23}^0 - b_{22})\cos\alpha_0,$$

gdzie α_0 jest kątem należącym do przedziału $15° \leqslant \alpha_0 \leqslant 30°$.

Gdy punkt C^0 leży wewnątrz obszaru sprawności koryguje się jego współrzędne zgodnie z zależnościami

$$b_{13}^1 = b_{12} + (b_{13}^0 - b_{12})\cos\alpha_0 - (b_{23}^0 - b_{22})\sin\alpha_0,$$
$$b_{23}^1 = b_{22} + (b_{13}^0 - b_{12})\sin\alpha_0 + (b_{23}^0 - b_{22})\cos\alpha_0.$$

Dla otrzymanego punktu $C^1(b_{13}^1, b_{23}^1)$ sprawdza się czy leży on na brzegu, wewnątrz czy poza obszarem sprawności. Jeżeli punkt C^1 leży po tej samej stronie brzegu obszaru co punkt C^0, to powtarza się procedurę wykorzystaną przy wyznaczaniu punktu C^1, znajdując punkt C^2. Proces kontynuuje się tak długo aż kolejny punkt C^k znajdzie się po drugiej stronie brzegu obszaru sprawności niż punkt C^{k-1}. Wtedy rozpoczynając od punktu C^{k-1} ponownie powtarza się procedurę wyznaczania kolejnego przybliżenia z tym, że zamiast kąta α_0 bierze się kąt $\alpha_1 = \frac{\alpha_0}{3}$. Proces powtarza się tak długo, aż kąt $\alpha_j = \frac{\alpha_{j-1}}{3}$ stanie się mniejszy od ustalonej poprzednio wartości $\alpha_{\min}$ (praktycznie 1°), chyba że wcześniej stwierdzi się, że kolejno wyznaczony punkt C^k należy do brzegu obszaru sprawności, czyli $C^k = C$.

Następnie powtarza się całą procedurę z tym, że dwoma początkowymi punktami dla wyznaczenia następnego punktu brzegu obszaru sprawności są punkty B i C.

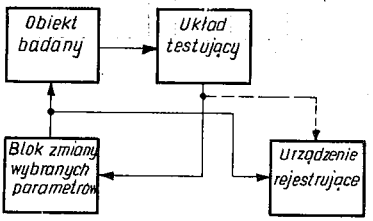
4. METODY EKSPERYMENTALNE WYZNACZANIA OBSZARU SPRAWNOŚCI

4.1. Uwagi wstępne

Opisane powyżej metody symulacyjne wyznaczania obszaru sprawności wymagają znajomości modelu matematycznego obiektu. Określanie takiego modelu jest często trudne. Trudności zwiększają się zresztą w miarę skomplikowania obiektu i zjawisk fizycznych, o które oparte jest jego działanie.

Modele są jedynie pewnym przybliżeniem rzeczywistości i uzyskany obszar sprawności jest tylko pewną aproksymacją obszaru rzeczywistego.

W związku z powyższym duże znaczenie posiadają metody eksperymentalne określenia obszarów sprawności polegające na badaniu rzeczywistych układów. Często zresztą inny jest zakres zastosowań obu metod. Metody symulacyjne stosowane są przy projektowaniu układu, a eksperymentalne przy jego badaniach.



Rys. 10. Schemat blokowy zestawu do wyznaczania obszarów sprawności

Zasadniczym problemem na jaki należy zwrócić uwagę przy porównywaniu metod eksperymentalnych jest to, czy podczas eksperymentu pojawiają się kombinacje wartości parametrów nie należące do obszaru sprawności. Istnienie takich prób podczas eksperymentu może prowadzić dla pewnych obiektów do ich zniszczenia lub też powstania w obiekcie pewnych zmian trwałych.

Ogólny schemat zestawu do eksperymentalnego wyznaczania obszarów sprawności pokazany jest na rys. 10. W skład zestawu wchodzą: blok zmiany wybranych parametrów badanego obiektu, układ testujący i blok rejestracji wyników.

Wśród parametrów b_i ($i = 1, 2, \ldots, n$) badanego obiektu wybiera się te, w stosunku do których należy wyznaczyć obszar sprawności. Wartości pozostałych parametrów są ustalone. Dla chwilowo ustalonej kombinacji wartości parametrów, dla których określa się obszar sprawności, sprawdza się przy pomocy układu testującego, czy parametry wyjściowe badanego obiektu spełniają warunki sprawności. Wynik badania jest rejestrowany i zostaje ustalona kolejna kombinacja wartości parametrów zmiennych.

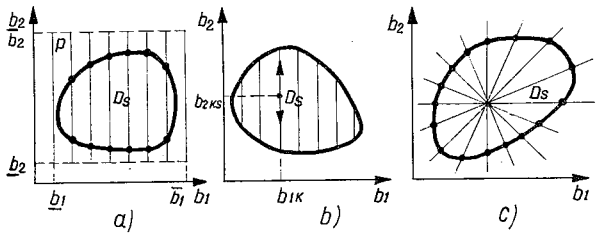
Budowa poszczególnych bloków zestawu zależy od stopnia zautomatyzowania pomiarów, od przyjętego sposobu wybierania kolejnych wartości parametrów zmiennych oraz od wykorzystywanego urządzenia rejestrującego (miernik uniwersalny, oscyloskop, rejestrator XY, dalekopis, drukarka). Z reguły, sposób rejestracji wyników ogranicza liczbę parametrów zmiennych. Na przykład wykorzystując oscyloskop można wyznaczać dwuwymiarowe obszary sprawności (co nie wyklucza możliwości badania trójwymiaro-

wych obszarów sprawności poprzez zdejmowanie odpowiednich przekrojów dwuwymiarowych).

Opisane niżej metody doświadczalnego znajdywania obszarów sprawności obiektu można podzielić, podobnie jak i metody symulacyjne, na dwie grupy: metody przeszukiwania systematycznego i liczb losowych oraz złotych oraz metody bezpośredniego wyznaczania brzegu obszaru sprawności.

4.2. Metody przeszukiwania systematycznego

Najprostsza metoda przeszukiwania systematycznego polega na tym, że wartość jednego spośród wybranych parametrów jest zmieniana skokowo, a wartość drugiego jest zmieniana w sposób ciągły w ustalonych granicach (rys. 11a). Urządzenie testujące sygnalizuje



Rys. 11. Przykłady sposobów przeszukiwania systematycznego

te wartości parametru zmienianego w sposób ciągły, dla których przekraczane są granice obszaru sprawności. Metoda taka była wykorzystana w pracy [17] przy wyznaczaniu tzw. pól zakazanych badanych liczników. Pomiary wykonywane były ręcznie.

Wariant tej metody wykorzystany był w pracy [1] do badania sieci logicznych. Polega on na tym, że dla ustalonej wartości b_{1k} parametru b_1 znajduje się początkową wartość parametru b_{2k_s} tak, żeby punkt pracy o współrzędnych (b_{1k}, b_{2k_s}) znajdował się wewnątrz obszaru sprawności. Następnie, przy ustalonej wartości b_{1k} zmienia się wartość b_{2k_s} (kolejno w obu kierunkach) tak długo, dopóki układ testujący nie zasygnalizuje niepoprawnej pracy badanej sieci. Procedurę powtarza się dla kolejnych wartości b_{ik} parametru b_1 (rys. 11b). Metoda ta jest pewną odmianą tzw. metody badań granicznych (p. pkt 4.2.1).

Do najprostszych metod systematycznych przeszukiwania obszaru sprawności można zaliczyć również tzw. metodę „pęku prostych" [14][4]. W metodzie tej, z dowolnego punktu przestrzeni parametrów prowadzi się pęk prostych. Dla każdej prostej wyznacza się punkty przecięcia z brzegiem obszaru sprawności (rys. 11c).

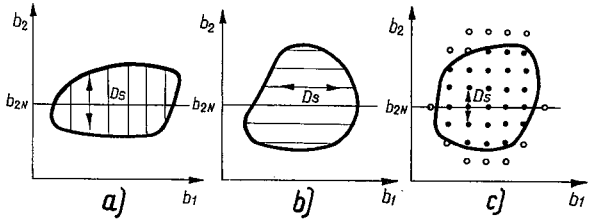
4.2.1. *Metoda badań granicznych*

Metoda badań granicznych opracowana została na początku lat pięćdziesiątych w ramach prac nad niezawodnością wojskowych maszyn cyfrowych [20, 33]. Polega ona na tym, że spośród „istotnych" parametrów b_i $(i = 1, 2, \ldots, n)$, od których zależy sprawność obiektu, wybiera się parametr, który ma największy wpływ na sprawność obiektu (tzw.

[4] W oryginale [14] metoda podana jest w postaci algorytmu na maszynę cyfrową.

parametr charakterystyczny). Następnie, dla tego parametru i niezależnie dla każdego z pozostałych parametrów określa się granice dwuwymiarowych obszarów sprawności.

Możliwe są różne sposoby postępowania. W jednym z nich dla kolejno ustalonych dyskretnych wartości parametru charakterystycznego zmienia się wartość drugiego parametru w sposób ciągły wokół jego wartości nominalnej (pozostałe parametry są ustalone) (rys. 12a). Inny sposób polega na tym, że zmienia się parametr charakterystyczny w sposób ciągły, a drugi parametr w sposób dyskretny (rys. 12b). Można również oba parametry zmieniać w sposób dyskretny (rys. 12c).



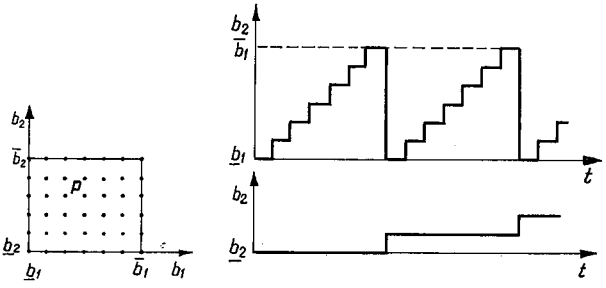
Rys. 12. Wyznaczanie obszaru sprawności metodą badań granicznych
a) parametr charakterystyczny zmienia się w sposób dyskretny, a parametr istotny w sposób ciągły, b) parametr charakterystyczny zmienia się w sposób ciągły, a parametr istotny w sposób dyskretny, c) oba parametry zmieniają się w sposób dyskretny

Wadą metody badań granicznych jest to, że nie pozwala ona uwzględnić funkcjonalnych zależności pomiędzy poszczególnymi parametrami istotnymi. Poza tym, nie zawsze jako parametr charakterystyczny wybierany jest ten, który ma największy wpływ na sprawność obiektu. Często o wyborze decydują dodatkowe czynniki (na przykład łatwość jego regulacji). Przy badaniach układów elektronicznych, jako parametr charakterystyczny najczęściej wybierane jest jedno z napięć zasilających.

4.2.2. *Metoda generatorów schodkowych*

W metodzie generatorów schodkowych zmianami dwóch wybranych parametrów sterują generatory schodkowe synchronizowane odpowiednimi ciągami zegarowymi. Zasadę wybierania kolejnych punktów przestrzeni parametrów ilustruje rys. 13.

Niżej omówione są dwa urządzenia realizujące metodę generatorów schodkowych: urządzenie do badań matrycowych i urządzenie do badania pamięci ferrytowych.

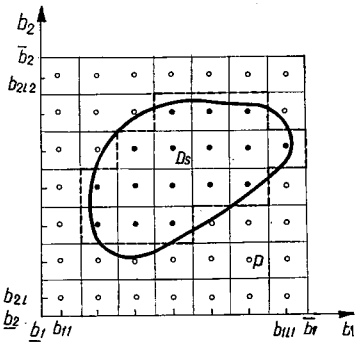


Rys. 13. Zasada przeszukiwania w przypadku wykorzystania generatorów schodkowych

W urządzeniu do badań matrycowych przy badaniu obiektu uwzględnia się jednocześnie wszystkie parametry istotne z punktu widzenia sprawności obiektu. Dopuszczalne przedziały zmian wartości istotnych parametrów $(\underline{b}_i, \bar{b}_i)$ dzieli się na l_i części (kwantów). Każda część reprezentowana jest przez wartość średnią. Określona kombinacja wartości średnich nazywana jest „sytuacją". Ilość możliwych sytuacji wynosi

$$N = \prod_{i=1}^{n} l_i.$$

Wybierając kolejne sytuacje przeszukuje się przestrzeń parametrów. Dla każdej sytuacji sprawdza się warunki sprawności obiektów. Przyjmuje się, że jeżeli dla danej wartości średniej obiekt jest niesprawny, to jest on również niesprawny dla wszystkich wartości należących do odcinka reprezentowanego przez tę wartość średnią (rys. 14).



Rys. 14. Wyznaczanie obszaru sprawności metodą matrycową

Opisane w pracy [34] urządzenie, jest urządzeniem elektromechanicznym. Zapewnia ono automatyczne wybieranie kolejnych sytuacji, sprawdzanie sprawności badanego obiektu dla każdej sytuacji i rejestrowanie liczby sytuacji, dla których wystąpiła niesprawność a także rodzaju tej niesprawności. Wybieranie kolejnych sytuacji polega na tym, że wyjścia z licznika impulsów zegarowych sterują przekaźnikami włączającymi w badanym obiekcie elementy o odpowiednich wartościach.

Urządzenie umożliwia badanie układów o n istotnych parametrach, przy czym spełnione muszą być dwa warunki

$$\sum_{i=1}^{n} l_i \leqslant 60 \quad \text{i} \quad l_i = 2^{s_i}$$

(s — liczba całkowita). Przy zachowaniu tych warunków urządzenie może badać jednocześnie kilka układów albo jeden układ przy kilku warunkach sprawności. Szybkość przeszukiwania przestrzeni parametrów, zależnie od rodzaju badanego układu, wynosi 20 Hz do 0,25 Hz.

Uproszczony model urządzenia do badań matrycowych opracowano w Katedrze Konstrukcji Maszyn Cyfrowych Politechniki Wrocławskiej [28]. Jest to automatyczny rejestrator macierzy sprawności przeznaczony do badania tranzystorowego układu negacji. Przyrząd umożliwia przeprowadzenie badań w dwóch układach współrzędnych: a) dla parametrów zewnętrznych — napięcia polaryzującego bazę (E_b) i napięcia kolektoro-

wego (E_k), b) dla parametrów wewnętrznych — szeregowego rezystora w bazie tranzystora (R_s) i rezystora polaryzacji bazy tranzystora (R_b). Wyniki badań są drukowane na dalekopisie w postaci macierzy sprawności (p. pkt 3.2.1). Możliwe jest przeprowadzanie badań dla 24 wartości $R_s(E_b)$ i 26 wartości $R_b(E_k)$.

Wartości parametrów zmieniane są skokowo. Do sterowania zmianami wartości parametrów użyto rozdzielaczy stykowych współpracujących z odpowiednimi dzielnikami napięcia oraz z zespołem rezystorów o różnych wartościach. Szybkość pracy urządzenia limitowana jest przez szybkość dalekopisu.

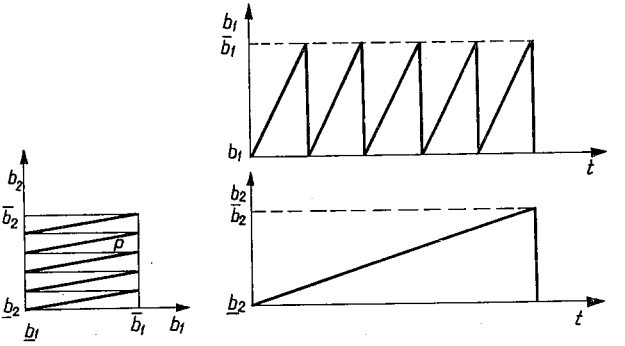
Urządzenie do badania pamięci ferrytowych (tzw. *Memory Exerciser-Plotter*) opracowane w 1965 r. [35] służy do wyznaczania dwuwymiarowych obszarów sprawności pamięci ferrytowej. Istotnymi parametrami są w tym wypadku prądy odczytu, zapisu i zakazu oraz próg wzmacniacza odczytu. Wyjścia generatorów schodkowych przyjmują 32 różne wartości, przy czym istnieje możliwość podwojenia tej liczby dla uzyskania dodatkowych szczegółów przekroju.

W kolejnych punktach przestrzeni parametrów sprawdza się działanie badanej pamięci drogą wielokrotnego zapisywania i odczytywania informacji, a następnie porównywania odczytanych wyników z postacią wpisanej uprzednio informacji. W przypadku wystąpienia niezgodności, uruchomiony zostaje układ sygnalizacji błędu. Wyniki wyświetlane są na ekranie oscyloskopu, którego wejścia X i Y połączone są z wyjściami generatorów schodkowych. Wyjście z układu sygnalizacji błędu połączone jest z wejściem modulacji jasności oscyloskopu. Punkty, dla których pamięć nie pracuje prawidłowo, są rozjaśniane. W efekcie, na ekranie uzyskuje się bezpośrednio obraz dwuwymiarowego obszaru sprawności.

Wyeliminowanie w urządzeniu części elektromechanicznych oraz zastosowanie oscyloskopu w charakterze urządzenia rejestrującego pozwoliło na znaczne skrócenie czasu uzyskania wyniku, w porównaniu z urządzeniami do badań matrycowych.

4.2.3. *Metoda generatorów przebiegów piłokształtnych*

Do przeszukiwania przestrzeni parametrów można wykorzystać generatory przebiegów piłokształtnych w oparciu o znane metody wybierania stosowane w telewizji (rys. 15).



Rys. 15. Przykładowe rozwiązanie przeszukiwania w przypadku stosowania generatorów przebiegów piłokształtnych

4.2.4. *Metoda figur Lissajous*

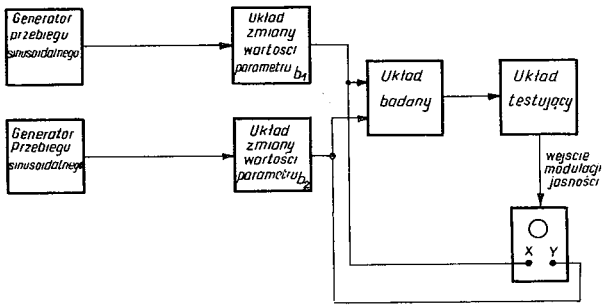
W przypadku wykorzystania generatorów schodkowych lub generatorów przebiegów piłokształtnych zwiększenie liczby parametrów, w funkcji których wyznaczany jest obszar sprawności, wiąże się z koniecznością stosowania generatorów o coraz większych częstotliwościach. Trudności tej można uniknąć wykorzystując do przeszukiwania przestrzeni parametrów własności figur Lissajous[5] — linii zakreślanych przez punkt, którego ruch jest wypadkową dwóch wzajemnie prostopadłych drgań harmonicznych.

Jeżeli wartości parametrów b_1 i b_2 zmieniają się sinusoidalnie z amplitudami odpowiednio $2A = \overline{b}_1 - \underline{b}_1$ i $2B = \overline{b}_2 - \underline{b}_2$ i z częstotliwościami kołowymi ω_1 i ω_2, to równania parametryczne figury Lissajous są następujące:

$$b_1 = A\sin\omega_1 t,$$
$$b_2 = B\sin(\omega_2 t + \varphi),$$

gdzie φ jest początkową różnicą faz między przebiegami.

Figury Lissajous mieszczą się w prostokącie o bokach $2A$ i $2B$. Ich kształt zależy od stosunku częstotliwości $\frac{\omega_1}{\omega_2}$ i od początkowej różnicy faz φ. Przy stosunku częstotliwości wymiernym, ruch punktu jest okresowy i figury Lissajous tworzą krzywą zamkniętą. Przy stosunku częstotliwości niewymiernym, ruch punktu nie jest ściśle okresowy i figury



Rys. 16. Schemat blokowy zestawu do przeszukiwania przestrzeni parametrów metodą figur Lissajous

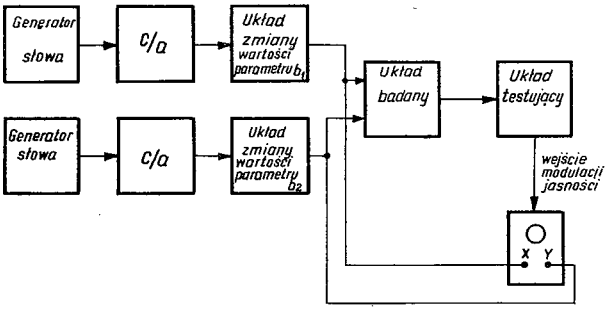
Lissajous są krzywymi wypełniającymi cały prostokąt. Stąd, dokładność przeszukiwania prostokąta przestrzeni parametrów zależy od stosunku częstoliwości, który może być bliski jedności. Wymaga to stosowania dwóch stabilnych generatorów o stałym (regulowanym) stosunku częstotliwości.

Do rejestracji wyników można wykorzystać oscyloskop XY z wejściem modulacji jasności sterowanym przez układ testujący. Możliwe są dwa sposoby wyświetlania wyników. Pierwszy polega na tym, że wyświetlane są jedynie te fragmenty figury Lissajous, dla których spełnione są (nie są spełnione) warunki sprawności. W drugim sposobie, wyświetlane są jedynie te punkty figury, które odpowiadają przejściu przez brzeg obszaru sprawności. Odpowiedni schemat blokowy urządzenia pokazany jest na rys. 16.

[5] pomysł dra hab. K. Bieńkowskiego.

4.3. Metody liczb losowych i złotych

W metodach liczb losowych do ustalenia kolejnych wartości wybranych parametrów wykorzystuje się generatory liczb losowych, których wyjścia bądź bezpośrednio sterują blokiem ustalania wartości parametrów (w przypadku dyskretyzacji przestrzeni parametrów), bądź sterują konwerterami cyfrowo-analogowymi, których wyjścia z kolei są połączone z wejściami bloku zmiany wartości parametrów. Przykładowy schemat blokowy zestawu do wyznaczania obszaru sprawności przy losowym przeszukiwaniu przestrzeni parametrów pokazany jest na rys. 17. Generatory liczb pseudolosowych mogą być w sposób prosty budowane w oparciu o liczniki łańcuchowe [11].



Rys. 17. Schemat blokowy zestawu do losowego przeszukiwania przestrzeni parametrów

Urządzenie wyposażone w generator liczb złotych (p. pkt 3.2.3) zostało skonstruowane w Instytucie Maszyn Matematycznych Politechniki Warszawskiej. Zastosowany generator liczb złotych działa na zasadzie sumowania aktualnej liczby z liczbą złotą, przy czym pomijana jest część całkowita wyniku sumowania [24].

4.4. Metody wyznaczania brzegu obszaru sprawności

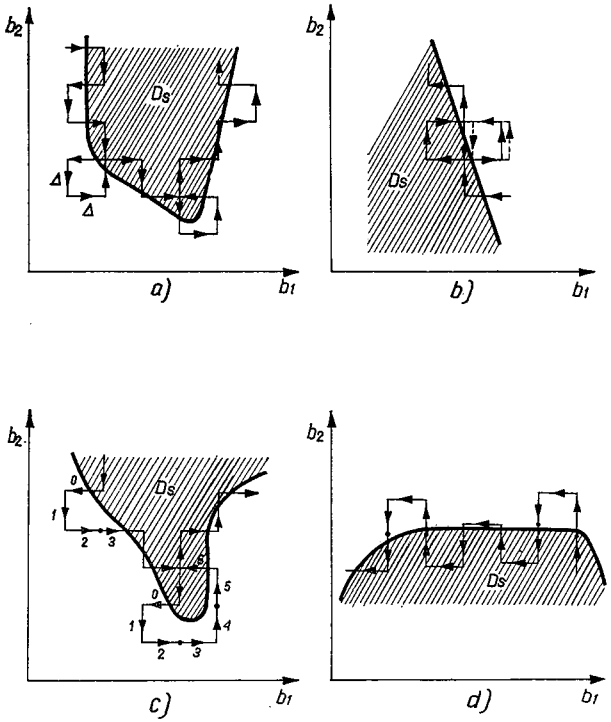
4.4.1. *Metoda obchodzenia brzegu obszaru sprawności*

Do doświadczalnego wyznaczania brzegu obszaru sprawności można wykorzystać metodę obchodzenia brzegu obszaru sprawności wykorzystaną w modelowym urządzeniu do rozpoznawania znaków [14, 25]. Jest ona pewną modyfikacją analitycznego algorytmu wyznaczania brzegu obszaru sprawności opisanego w punkcie 3.2.2.

W metodzie obchodzenia brzegu obszaru sprawności wartości dwóch wybranych parametrów obiektu są zmieniane kolejno. W każdym kroku wartość parametru jest zwiększana lub zmniejszana o ustaloną wielkość Δ. Po zmienieniu wartości parametru o Δ sprawdza się, czy w czasie zmiany przekroczony został brzeg obszaru sprawności. Jeżeli tak i jeżeli nowy punkt pracy znajduje się wewnątrz obszaru sprawności, to następny krok wykonywany jest zgodnie z zasadą obchodzenia brzegu obszaru w kierunku zgodnym

z ruchem wskazówek zegara. Jeżeli natomiast po przekroczeniu brzegu punkt pracy znalazł się poza obszarem sprawności, kierunek obchodzenia brzegu w następnym kroku jest ustalany przeciwnie do ruchu wskazówek zegara (rys. 18a).

W przypadku, gdy w czasie zmiany parametru o wielkość Δ nie został przekroczony brzeg obszaru sprawności, następne kroki, aż do ponownego przekroczenia brzegu obszaru, są wykonywane z zachowaniem ostatniego kierunku obchodzenia brzegu obszaru, z tym jednak, że zmiana zmienianego parametru występuje po każdym nieparzystym kroku.



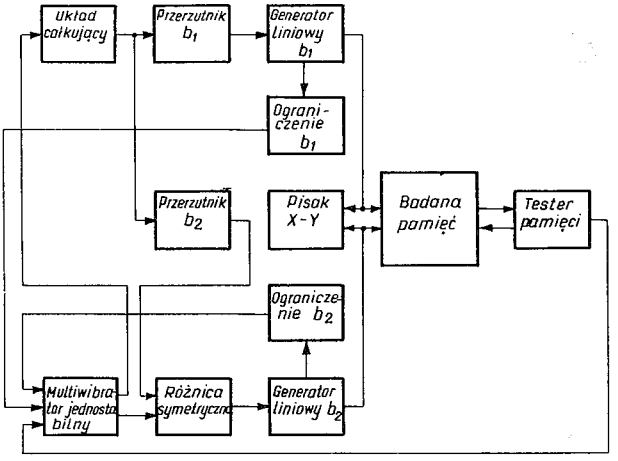
Rys. 18. Metoda obchodzenia brzegu obszaru sprawności
a) algorytm podstawowy, b) uzasadnienie celowości modyfikacji algorytmu, c) algorytm zmodyfikowany, d) algorytm zmodyfikowany w przypadku wyznaczania brzegu obszaru sprawności z ograniczeniem przy obecności zakłóceń

Za krok zerowy uważany jest ostatni krok, w którym został przekroczony brzeg obszaru (rys. 18c). Takie postępowanie ma na celu zabezpieczenie przed wpływem zakłóceń, w wyniku czego zmiany wartości parametru w kolejnych krokach mogą być różne. Jak pokazano na rys. 18b (linia przerywana), może to spowodować nieprawidłową pracę urządzenia.

Mimo wprowadzenia opisanej modyfikacji, w stosunku do analitycznej metody obchodzenia brzegu obszaru sprawności (p. pkt 3.2.2), urządzenie zbudowane w oparciu o omawianą metodę może (w obecności zakłóceń) pracować niejednoznacznie (np. w przypadku obszarów sprawności z ograniczeniem, co ilustruje rys. 18d). Wadą metody jest również to, że w czasie badania obiektu punkt pracy wychodzi poza obszar sprawności.

4.4.2. *Metoda stosowana w urządzeniu* Schmoo plotter

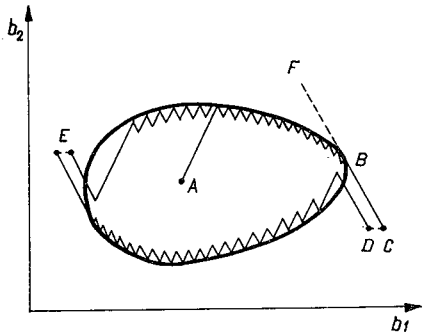
Ciekawą metodę wyznaczania brzegu obszaru sprawności opracowano w laboratorium firmy International Business Machines Corp. (1966). W oparciu o tę metodę zbudowano urządzenie, tzw. *Schmoo plotter*, przeznaczone w zasadzie do badania obszarów sprawności pamięci ferrytowych [16]. Może być ono przystosowane również do badania innych obiektów.



Rys. 19. Schemat blokowy *schmoo plotter'a*

Na rys. 19 podano schemat blokowy urządzenia w przypadku badania pamięci ferrytowej. Działanie urządzenia ilustruje rys. 20.

Dwa generatory przebiegów liniowych sterują zmianami dwóch wybranych parametrów b_1 i b_2 (prądów lub napięć; zmiany innych parametrów wymagają stosowania odpowied-



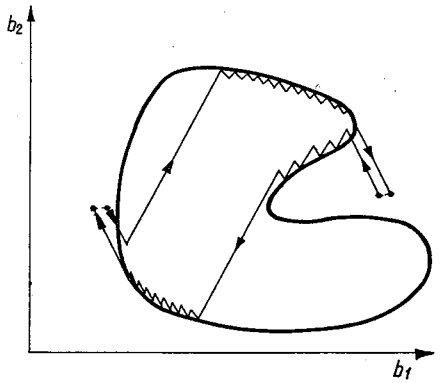
Rys. 20. Zasada wyznaczania granicy obszaru sprawności w urządzeniu *schmoo plotter*

nich konwerterów). Poczynając od początkowego punktu pracy (punkt A na rys. 20) jednocześnie zwiększane są wartości obu parametrów. Wartość parametru b_1 zmienia się znacznie wolniej niż wartość parametru b_2. Wartości parametrów b_1 i b_2 zmieniane są tak długo, aż układ testujący zasygnalizuje, że badana pamięć przestała pracować poprawnie.

nie. Sygnał błędu uruchamia układ multiwibratora jednostabilnego. Zmiana poziomu wyjściowego tego układu powoduje, poprzez układ różnicy symetrycznej, zmianę kierunku zmian wartości parametru b_2. Wartość parametru b_1 jest w dalszym ciągu zwiększana. Układ multiwibratora jednostabilnego wraca do stanu początkowego po czasie wystarczającym dla czterokrotnego sprawdzenia każdego słowa pamięci i ponownie zmienia się kierunek zmian parametru b_2. Punkt pracy zdąża w kierunku brzegu obszaru sprawności.

Proces powtarza się tak długo, aż zostanie osiągnięty najbardziej w prawo wysunięty punkt obszaru sprawności (punkt B na rys. 20) lub gdy zostanie osiągnięta granica dopuszczalnych zmian wartości parametru b_1 (ustalana przez układ ograniczający b_1). Wtedy wartość parametru b_2 (po ostatniej zmianie kierunku) maleje, nie napotykając brzegu obszaru sprawności. Jeżeli czas przebywania punktu pracy poza obszarem sprawności przekroczy dopuszczalną wartość (ustaloną przez układ całkujący), sygnał z układu całkującego powoduje zmianę stanu przerzutników b_1 i b_2 (punkt C na rys. 20). Dodatkowo, wartość parametru b_1 jest zmniejszana skokowo o pewną ustaloną wartość (odcinek CD na rys. 20). Ma to na celu zabezpieczenie przed powrotem punktu pracy po tej samej linii co poprzednio (odcinek CB na rys. 20), co mogłoby doprowadzić do „zgubienia" brzegu obszaru (na rys. 20 punkt pracy poruszałby się po prostej CF).

Poczynając od punktu D punkt pracy porusza się w kierunku obszaru sprawności. Po przekroczeniu brzegu obszaru sprawności rozpoczyna się normalny cykl pracy urządzenia. Wyznaczana jest dolna część krzywej ograniczającej obszar sprawności.



Rys. 21. Przykład obszaru sprawności, którego *schmoo plotter* nie wyznaczy poprawnie

Po osiągnięciu najbardziej w lewo wysuniętego punktu obszaru sprawności (punkt E na rys. 20) lub granicy dopuszczalnych zmian parametru b_1, ponownie zmieniane są stany przerzutników b_1 i b_2 i urządzenie wyznacza górną część krzywej ograniczającej obszar sprawności. Praca urządzenia trwa do chwili, gdy zostanie wyznaczony cały brzeg obszaru sprawności.

Zmiany wartości parametrów b_1 i b_2 rejestrowane są na pisaku XY dając w efekcie zarys brzegu obszaru sprawności. Czas wykreślania brzegu jest rzędu kilku minut.

Dokładna analiza pracy urządzenia pokazuje, że umożliwia ono badanie jedynie wypukłych obszarów sprawności. W przypadku wystąpienia wklęsłego obszaru sprawności, część obszaru może zostać „obcięta". Ilustruje to rys. 21. Poza tym, w niektórych przy-

padkach, nawet dla obszaru wypukłego pewne fragmenty obszaru mogą być pominięte. Wadą metody jest również konieczność wchodzenia głęboko w obszar niesprawności, w chwilach gdy zmienia się kierunek zmian wartości parametru b_1.

4.4.3. *Metoda śledzenia brzegu płaskiego obszaru sprawności*

Każdą z omówionych wyżej metod doświadczalnego wyznaczania obszaru sprawności cechuje ograniczony zakres stosowalności. W większości, w metodach tych wykorzystuje się koncepcję przeszukiwania całej przestrzeni parametrów lub jej wybranego fragmentu, co wiąże się z koniecznością badania obiektu zarówno w warunkach sprawności jak i niesprawności. Czas potrzebny do wyznaczenia obszaru sprawności, nawet w przypadku zautomatyzowania pomiarów, jest znaczny. Dla ominięcia tych trudności celowe wydaje się wykorzystanie metod bezpośredniego wyznaczania brzegu obszaru sprawności.

Jak jednak pokazano, w obu omówionych metodach (p. pkt 4.4.1 i 4.4.2), aczkolwiek w mniejszym stopniu, nadal występuje problem badania obiektu w warunkach sprawności i niesprawności. Dodatkowo, w metodzie stosowanej w urządzeniu *schmoo plotter* powstaje problem ograniczenia zakresu stosowalności do obszarów wypukłych. W metodzie obchodzenia powstaje natomiast zagadnienie zapewnienia jednoznaczności wyznaczania brzegu obszaru.

Niżej przedstawiona jest metoda, nie posiadająca wad znanych rozwiązań — tzw. metoda śledzenia brzegu obszaru sprawności [36, 37]. Metoda ta pozwala wyznaczać brzeg w zasadzie dowolnego obszaru sprawności z zachowaniem warunku, że punkt pracy w czasie badania nie wyjdzie poza obszar sprawności. Jedynym warunkiem wyznaczenia brzegu obszaru sprawności jest znajomość co najmniej jednego punktu należącego do obszaru sprawności.

Metoda śledzenia brzegu obszaru sprawności polega na wykonaniu szeregu kroków według algorytmu ustalonego na podstawie analizy możliwych konfiguracji obszarów sprawności.

Algorytm metody śledzenia brzegu sprawności charakteryzują trzy podstawowe cechy wyróżniające go spośród innych, znanych algorytmów:

1) w czasie badania obiektu w funkcji dwóch wybranych parametrów, wartości tych parametrów są zmieniane kolejno, niezależnie od siebie;

2) w kolejnych krokach wielkości zmian wartości parametrów są nie większe od z góry ustalonych wartości;

3) po każdej zmianie wartości parametru otrzymana wartość jest zmieniana w kierunku przeciwnym o niewielką ustaloną wielkość.

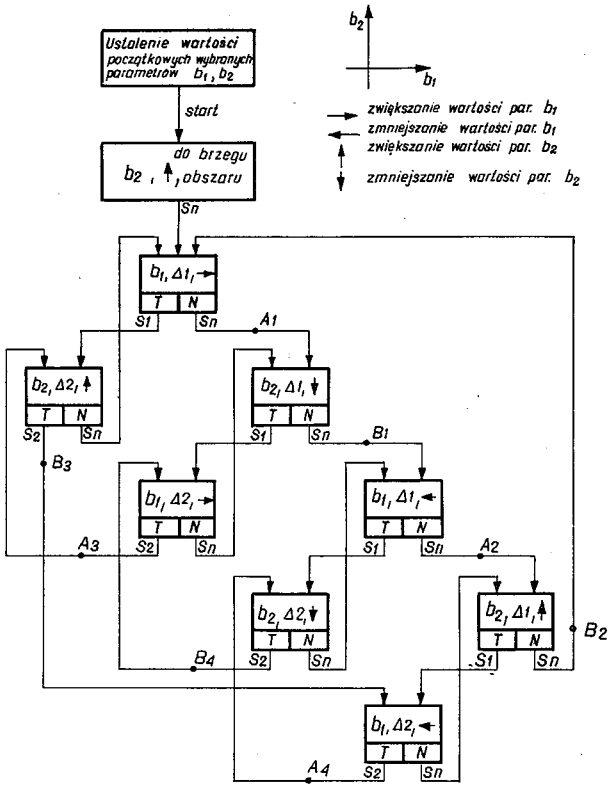
Przed przystąpieniem do wyznaczania brzegu obszaru sprawności ustala się warunki sprawności obiektu, wartości początkowe dwóch wybranych parametrów b_1 i b_2 obiektu, dla których obiekt jest sprawny oraz początkowe kierunki zmian wartości parametrów b_1 i b_2.

W pierwszym kroku zmienia się wartość jednego parametru do chwili, gdy przestanie być spełniony przynajmniej jeden z warunków sprawności. W kolejnych krokach zmienia się na przemian wartości parametrów b_1 i b_2 tak, żeby punkt pracy obiektu nie znalazł

się poza obszarem sprawności, przy czym w czasie gdy zmienia się wartość jednego parametru wartość drugiego jest stała i równa wartości ustalonej w kroku poprzednim.

W drugim i w następnych krokach podejmuje się próbę zmienienia wartości parametru o stałą wielkość $\Delta 1$, albo o stałą wielkość $\Delta 2$ ($\Delta 1 < \Delta 2$). Zmienianie wartości parametru zostaje przerwane z chwilą gdy osiągnięty został brzeg obszaru sprawności (próba nieudana), lub gdy wykonana została zmiana o wartość $\Delta 1$ albo o $\Delta 2$ (próba udana).

W pierwszym kroku po osiągnięciu brzegu obszaru sprawności, a w następnych krokach po udanej albo nieudanej próbie zmienienia wartości parametru o $\Delta 1$ albo o $\Delta 2$, uzyskaną w danym kroku wartość parametru zmienia się w kierunku przeciwnym o stałą wielkość Δ ($\Delta < \Delta 1 < \Delta 2$).



Rys. 22. Algorytm metody śledzenia brzegu obszaru sprawności

(Po każdym kroku występuje nie zaznaczona na schemacie zmiana nowo otrzymanej wartości parametru w kierunku przeciwnym o wartość $\Delta < \Delta 1 < \Delta 2$)

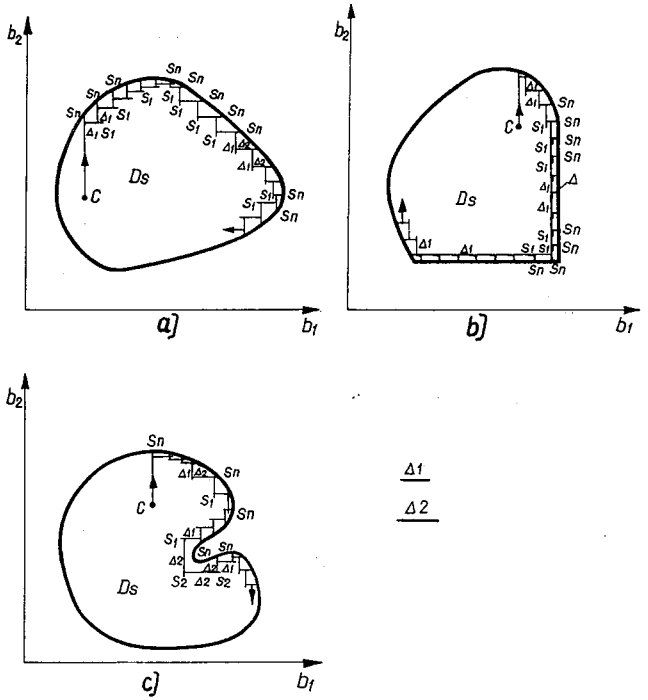
Próbę zmiany wartości wybranego parametru o wielkość $\Delta 1$ podejmuje się wtedy, gdy w poprzednim kroku nie udała się próba zmiany wartości drugiego parametru o wielkość $\Delta 1$ albo o $\Delta 2$.

Próbę zmiany wartości wybranego parametru o wielkość $\Delta 2$ podejmuje się wtedy, gdy w poprzednim kroku udała się próba zmiany wartości drugiego parametru o $\Delta 1$ albo o $\Delta 2$.

Kierunek zmieniania wartości parametru zmienia się na przeciwny wtedy, gdy w dwu kolejnych krokach próby zmiany wartości parametrów o $\Delta 1$ albo o $\Delta 2$ nie były udane, albo gdy w poprzednim kroku udała się próba zmiany wartości parametru o $\Delta 2$. Zmiana kierunku występująca w czasie zmieniania wartości parametru o Δ jest odwracalną zmianą chwilową.

Na rys. 22 podany jest algorytm metody śledzenia brzegu obszaru sprawności w przypadku, gdy początkowe kierunki zmieniania wartości wybranych parametrów b_1 i b_2 odpowiadają zwiększaniu tych wartości.[6)]

Na rys. 23 podano przykłady ilustrujące działanie algorytmu dla wypukłego obszaru sprawności (rys. 23a), dla obszaru z ograniczeniem (rys. 23b) i dla obszaru wklęsłego (rys. 23c). Na wszystkich rysunkach punkt C oznacza początkowy punkt pracy obiektu (jest to punkt o którym wiadomo tylko, że należy do obszaru sprawności). Na rysunkach

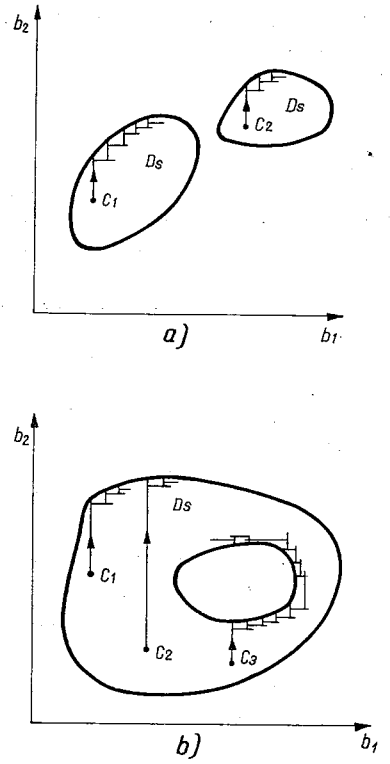


Rys. 23. Przykłady ilustrujące działanie algorytmu w przypadku: a) obszaru wypukłego, b) obszaru z ograniczeniami, c) obszaru wklęsłego

zaznaczono sytuacje występujące po skończeniu kolejnego kroku oraz o jaką wielkość, $\Delta 1$ czy $\Delta 2$, próbowano zmienić wartość parametru w danym kroku. Rysunek 23b, na którym pokazano przypadek obszaru sprawności z ograniczeniami wyjaśnia celowość wprowadzenia po każdej zmianie wartości parametru zmniejszania tej wartości o wielkość Δ.

[6)] Na schemacie blokowym z rys. 22 ze względu na przejrzystość schematu nie zaznaczono kroków odpowiadających automatycznemu cofaniu o Δ.

Opisana metoda wyznaczania brzegu obszaru sprawności metodą śledzenia umożliwia również badanie obszarów niespójnych. Wymaga to jednak wykonania dodatkowych czynności. W przypadku obszarów niespójnych (rys. 24a) należy niezależnie dla każdej z części obszaru wyznaczyć jej brzeg rozpoczynając badanie od punktów C_1, C_2 ... należących do różnych części obszaru.



Rys. 24. Przykłady ilustrujące możliwości badania obszarów sprawności: a) niespójnych, b) wielospójnych, metodą śledzenia brzegu obszaru

W przypadku, gdy istnieje podejrzenie, że badany obszar sprawności jest wielospójny, należy przeprowadzić dodatkowe próby wyznaczenia brzegu obszaru sprawności rozpoczynając badania od różnych punktów początkowych C_1, C_2, ... (należących do obszaru sprawności). Jeżeli w pierwszym kroku, w czasie zwiększania parametru b_2 natrafi się na brzeg „dziury" w obszarze sprawności, to brzeg ten można wyznaczyć zgodnie z opisanym wyżej algorytmem. O tym, że brzeg ten jest brzegiem „dziury" można wnioskować z faktu, że punkt początkowy leży poza wyznaczonym obszarem. Ilustruje to rys. 24b dla przypadku obszaru dwuspójnego.

Wykorzystując opisaną metodę śledzenia brzegu obszaru sprawności można zbudować urządzenie pozwalające na automatyczne wyznaczanie brzegu obszaru sprawności. Model takiego urządzenia opisano w pracy [36].

LITERATURA CYTOWANA W TEKŚCIE

1. W. Bara, A. Sielicki, Określenie optymalnego położenia punktu pracy sieci logicznej, Zeszyty Naukowe Politechniki Wrocławskiej, Łączność V, 1963.
2. M. Barański, A. Sielicki, Metody poszukiwania przestrzeni wymuszeń przy kontroli podzespołów cyfrowych, Prace Naukowe Instytutu Cybernetyki Technicznej Politechniki Wrocławskiej, Studia i Materiały, nr 4, 1970.
3. K. Bieńkowski, Miary jakości elektrycznych układów logicznych z potencjałowym kodowaniem sygnałów, Zeszyty Naukowe Politechniki Warszawskiej, Elektryka 53, 1967.
4. W. Bonngenaar, N. C. de Troye, Worst Case Considerations in Designing Logical Circuits, IEEE Trans. on EC, August, 1965.
5. J. Bromirski, A. Sielicki, Kryteria i algorytmy optymalizacji wtórnej podstawowych układów logicznych, Przegląd Elektrotechniczny, nr 5, 1967.
6. S. Budkowski, O pewnych metodach optymalizacji układów logicznych, rozprawa doktorska, Politechnika Warszawska, 1967.
7. S. Budkowski, Optymalizacja układów logicznych negacji sumy metodą ślepego przeszukiwania przestrzeni zmiennych decyzyjnych, Rozprawy Elektrotechniczne, T. XV, z. 3, 1969.
8. S. Budkowski, O pewnych metodach badania sprawności układów, Archiwum Automatyki i Telemechaniki, T. XVI, z. 3, 1971.
9. S. Budkowski, O pewnych metodach optymalizacji układów logicznych, Archiwum Automatyki i Telemechaniki, T. XIV, z. 4, 1969.
10. S. Budkowski, O pewnej metodzie generacji współrzędnych punktów równomiernie rozmieszczonych w kostce $[0,1]^n$, Algorytmy, nr 12, 1970.
11. S. Budkowski, T. Kończyk, Licznik łańcuchowy jako generator liczb pseudolosowych, Rozprawy Elektrotechniczne, T. XVIII, z. 1, 1972.
12. W. C. Cave, Move Toward Automated Circuit Design With a Method of Vector Analysis That Reduces Network Planning to a Standart Programming Problem, Electronic Design, 17, August 15, 1968.
13. W. C. Cave, Evaluate Digital Circuits Automatically in Terms of Output Signall Level Constraints and Tolerances, to Get Optional Design, Electronic Design, 18, September 1, 1968.
14. W. I. Czernieckij, G A. Diduk, A. A. Potapienko, Matematiczeskije metody i algoritmy issledowania awtomaticzeskich sistem, Energia, 1970.
15. L. A. Delholm, Design and Application of Transistor Switching Circuits, McGraw-Hill, 1968.
16. J. E. Gersbach, The Great Schmoo Plott: Testing Memories Automatically, Electronics, July 25, 1966.
17. J. Gierdalski, J. Połoński, Metoda weryfikacji elementów prototypu, Prace PIT, nr 44, 1964.
18. Ł. M. Goldenberg, Teoria i rascziot impulsnych ustrojstw na poluprowadnikowych priborach, Moskwa 1969.
19. G. H. Gray, Digital computer engineering, Prentice-Hall, 1963.
20. H. F. Heath, R. E. Nienburg, M. M. Astrahan, L. R. Walters, Reliability of an Air Defense Computing System: Component Development, Circuit Design, Marginal Checking and Maintenance Programming, IRE trans. on Electronic Computers, December, 1956.
21. K. A. Joudu, Optimizacja ustrojstw awtomatiki po kritieriju nadiożnosti, Moskwa, 1966.
22. B. M. Kagan, T. M. Ter-Mikaelian, Reszenie inżiniernych zadacz na cifrowych wyczislitielnych maszinach, Energia, 1964.
23. A. Kiliński, Definicje opisowo-analityczne i wartościująco-normatywne podstawowych pojęć teorii niezawodności, Prakseologia, nr 38, 1971.
24. J. Kłosiński, Generator elektroniczny liczb złotych, praca dyplomowa wykonana w Instytucie Maszyn Matematycznych Politechniki Warszawskiej pod kierunkiem S. Budkowskiego, 1970.
25. W. A. Kowalewskij, A. G. Semenowskij, Cifrowaja sledjaszczaja sistema czitajuszczewo awtomata, w zbiorze „Woprosy wyczislitielnoj matematiki i wyczislitielnoj techniki", Moskwa 1963.

26. S. Leptow, A. Sielicki, Wyznaczanie obszaru sprawności inwertera tranzystorowego przy użyciu elektronicznej maszyny cyfrowej UMC-1, Zeszyty Naukowe Politechniki Wrocławskiej, Automatyka I, 1964.
27. A. Sielicki, Analiza obszaru sprawności i jej wykorzystanie przy projektowaniu podstawowych układów logicznych, Rozprawa doktorska, Politechnika Wrocławska, 1963.
28. A. Sielicki, Przyrząd do rejestracji matryc sprawności inwertera tranzystorowego, Zeszyty Naukowe Politechniki Wrocławskiej, Automatyka II, 1966.
29. A. Sielicki, Pojęcie obszaru sprawności i jego wykorzystanie przy projektowaniu podstawowych układów logicznych, Prace III Krajowej Konferencji Automatyki, t. II, Politechnika Śląska, Gliwice 1966.
30. A. Sielicki, Zagadnienia optymalnej syntezy teoretycznej podstawowych układów logicznych, Zeszyty Naukowe Politechniki Wrocławskiej, Automatyka VI, 1967.
31. A. Sielicki, Optymalizacja układów logicznych ze względu na kryterium maksymalnych napięć zasilających, Zeszyty Naukowe Politechniki Wrocławskiej, Automatyka VII, 1968.
32. H. Steinhaus, Liczby złote i żelazne, Zastosowania matematyki, t. III, z. 1, 1956.
33. N. H. Taylor i inni, Designing for Reliability, Proc. of the IRE, June, 1957.
34. B. W. Wasiljew i inni, Nadiożnost i efektiwnost radioelektronnych ustrojstw, Sowietskoje Radio, 1965.
35. Ch. P. Woomack, Schmoo Plot Analysis of Coincident-Current Memory System, IEEE Trans. on Electronic Computers, February, 1965.
36. J. Zabrodzki, Eksperymentalne wyznaczanie obszarów sprawności, Rozprawa doktorska, Politechnika Warszawska, 1971.
37. J. Zabrodzki, Algorytm śledzenia brzegu obszaru sprawności, Archiwum Automatyki i Telemechaniki, T. XVII, z. 3, 1972.
38. R. Zieliński, Generator liczb losowych, WNT, 1970.

J. ZABRODZKI, S. BUDKOWSKI

METHODS FOR DETERMINING A REGION OF THE PROPER OPERATION

Summary

This paper concerns the methods for determining a region of the proper operation. The methods discussed are divided into simulating and experimental ones. In each of these groups there have been distinguished the searching methods: systematic and random ones and the methods for direct determining the boundary of the proper operation region.

J. ZABRODZKI, S. BUDKOWSKI

MÉTHODES DE DÉTERMINATION DE L'ESPACE DE BONNE FONCTIONEMENT

Résumé

Dans l'article on décrit les méthodes experimentales et celles de simulations. On distingue les méthodes qui consistent à la recherche point par point de l'espace de bonne fonctionement et les méthodes de déterminer seulement les points de borne.

J. ZABRODZKI, S. BUDKOWSKI

BESTIMMUNGSMETHODEN FÜR LEISTUNGSFÄHIGKEITSBEREICHE

Zusammenfassung

Es wurde Bestimmungsmethoden für Leistungsfähigkeitsbereiche diskutiert. Die erwogenen Methoden wurden in Simulations — und experimentelle Methoden eingeteilt. In jeder dieser Gruppen wurden wie-

derum Methoden zur Erforschung des Parameterraumes ausgesondert: systematische und zufällige sowie Methoden zur unmittelbaren Bestimmung des Randes des Leistungsfähigkeitsbereiches.

Я. ЗАБРОДЗКИ, С. БУДКОВСКИ

МЕТОДЫ ОПРЕДЕЛЕНИЯ ОБЛАСТИ РАБОТОСПОСОБНОСТИ

Р е з ю м е

В труде представлены методы определения области работоспособности. Рассматриваемые методы разделены на симуляционные и экспериментальные. В каждой группе различается систематические и случайные методы пересматривания области параметров и методы непосредственного наблюдения границы области работоспособности.

ROZPRAWY ELEKTRONICZNE 1974, 20, z. 1, ss. 41-58

621.3.019.3

Metody oceny niezawodności systemów telekomunikacyjnych

OLGA CHOREŃ (GDAŃSK), MARIAN ZIENTALSKI (GDAŃSK)
Instytut Telekomunikacji Politechniki Gdańskiej.

Otrzymano 28.8.1972

W artykule omówiono metody określania niezawodności systemów telekomunikacyjnych. Określono niezawodność systemów opisanych za pomocą dowolnego znanego operatora, podano macierzową metodę określenia niezawodności oraz metodę określenia niezawodności na podstawie równań różniczkowych Kołmogorowa. Dokonano krytycznej oceny przedstawionych metod określenia niezawodności. Podano przykład obliczeń parametrów niezawodnościowych dla stojaka generacyjnego telefonii wielokrotnej systemu TN-300.

1. WPROWADZENIE

Gwałtowny rozwój telekomunikacji, zwłaszcza w ostatnich kilkunastu latach, spowodował, że wchodzące w skład systemów telekomunikacyjnych urządzenia stają się coraz bardziej złożone i odpowiedzialne pod względem wykonywanych funkcji, a warunki ich eksploatacji są coraz trudniejsze. Również ilość przekazywanych informacji i jej znaczenie wymaga z zasady bardzo dużej niezawodności pracy urządzeń i systemów telekomunikacyjnych. W tej sytuacji znajomość parametrów niezawodnościowych systemu jest niezbędna i decyduje o możliwości jego wykorzystania. Przy określaniu parametrów niezawodnościowych systemu napotykamy jednak szereg trudności, ponieważ istniejące ogólne metody oceny niezawodności są mało przejrzyste i trudne do wykorzystania w konkretnych zastosowaniach praktycznych. W artykule przedstawiono dotychczas stosowane metody oceny niezawodności, dokonano krytycznej oceny przedstawionych metod i wskazano kierunki dalszych prac. W załączniku przytoczono przykład wykorzystania omówionych metod oceny niezawodności dla obliczeń niezawodności stojaka generacyjnego telefonii wielokrotnej systemu TN-300.

2. OKREŚLENIE NIEZAWODNOŚCI SYSTEMÓW OPISYWANYCH ZA POMOCĄ DOWOLNEGO ZNANEGO OPERATORA

Stan każdego dowolnego systemu może być scharakteryzowany na podstawie wartości parametrów wyjściowych, określających zdolność systemu do pracy. Na skutek oddziaływania różnorodnych czynników, związanych z procesem eksploatacji, parametry te w sposób losowy zmieniają swoje wartości. Proces zmian polega na tym, że wspomniane

wyżej czynniki, oddziaływując bezpośrednio na elementy systemu, zmieniają parametry tych elementów. Ponieważ wyjściowe parametry systemu związane są za pośrednictwem pewnego operatora z parametrami wejściowymi, to zmiany ostatnich powodują również zmiany parametrów wyjściowych. Jako parametry wejściowe będziemy rozpatrywać zarówno sygnały wejściowe i zakłócenia, jak i wartości parametrów elementów systemu, współczynniki wzmocnienia, napięcia zasilające, parametry charakteryzujące warunki eksploatacji itd. Do parametrów wejściowych należą również wielkości charakteryzujące wykonanie funkcji na zadanym poziomie.

W przypadku ogólnym wektorową funkcję losową $\vec{X}(t)$ charakteryzującą parametry wyjściowe systemu możemy przedstawić w sposób następujący:

$$\vec{X}(t) = \tilde{A}\vec{X}_0(t), \tag{1}$$

gdzie: $\tilde{A}$ — operator systemu,

$\vec{X}_0(t)$ — wektorowa funkcja losowa, charakteryzująca parametry wejściowe systemu.

Problem badania niezawodności systemu polega na określeniu wyjściowej funkcji losowej w postaci $\vec{T}(x)$ [7, 8]:

$$\vec{T}(x) = A_T^{-1}\tilde{A}\vec{X}_0(t), \tag{2}$$

gdzie: $\vec{T}(x)$ — wektorowa funkcja losowa odwrotna względem funkcji $\vec{X}(t)$,

A_T^{-1} — operator przekształcający funkcję $\vec{X}(t)$ w funkcję $\vec{T}(x)$.

W przypadku ogólnym określenie wektorowej funkcji losowej $\vec{T}(x)$ przedstawia sobą skomplikowany problem, składający się z dwóch zasadniczych części:

1) przekształcenie wejściowej funkcji losowej $\vec{X}_0(t)$ za pomocą operatora $\tilde{A}$,

2) przekształcenie otrzymanej wyjściowej funkcji losowej $\vec{X}(t)$ za pomocą operatora A_T^{-1}.

Zagadnienie pierwsze dostatecznie szczegółowo rozpatruje się w teorii sterowania dla przypadków liniowego i nieliniowego operatora $\tilde{A}$ [7], przy czym bardziej wszechstronnie opracowana jest teoria przekształcenia za pomocą operatora liniowego lub zlinearyzowanego. W przypadku, gdy operator $\tilde{A}$ jest nieliniowy, dąży się zwykle do linearyzacji zagadnienia, ponieważ nie ma obecnie dokładnych metod przekształcenia funkcji losowej za pomocą dowolnego nieliniowego operatora. Można wyróżnić trzy podstawowe metody linearyzacji:

1) metoda linearyzacji bezpośredniej, czyli przedstawienie nieliniowego operatora $\tilde{A}$ za pomocą operatora liniowego A,

2) linearyzacja wyjściowej funkcji losowej $\vec{X}(t)$ po zastosowaniu do funkcji $\vec{X}_0(t)$ nieliniowego operatora $\tilde{A}$,

3) metoda linearyzacji statystycznej.

Pierwsze dwie metody są metodami klasycznymi i oparte są na ogólnie znanych zasadach linearyzacji, dlatego też bardziej szczegółowo omówimy tylko metodę trzecią. Istota metody statystycznej linearyzacji polega na zamianie nieliniowych funkcji losowych takimi liniowymi funkcjami losowymi, które w sensie statystycznym są równoważne danym funkcjom nieliniowym. Warunki statystycznej równoważności można sformułować w sposób

różny. W pierwszym przybliżeniu można przyjąć jako funkcje statystycznie równoznaczne takie funkcje, które mają jednakowe momenty pierwszego i drugiego rzędu. Zgodnie z tym założeniem zagadnienie statystycznej linearyzacji wyjściowej funkcji $\vec{X}(t)$ będzie polegało na zamianie jej przez funkcję $\vec{Z}(t)$, mającą postać [7, 8];

$$\vec{Z}(t) = \vec{k}_0 \vec{E}_x(t) + \vec{k}_1 \vec{X}(t), \tag{3}$$

gdzie:

$\vec{E}_x(t)$ — wartość oczekiwana funkcji losowej $\vec{X}(t)$,

$\vec{k}_0$ i $\vec{k}_1$ — współczynniki, które należy dobrać w ten sposób, aby zostały spełnione warunki statystycznej równoważności funkcji.

Metody przekształcenia wyjściowej funkcji losowej $\vec{X}(t)$ za pomocą operatora A_T^{-1} oparte są na kanonicznym przedstawieniu funkcji [8]. W pracach poświęconych teorii funkcji losowych pokazano, że operator A_T^{-1} przedstawia sobą algebraiczne równanie stochastyczne następującej postaci [7, 8]:

$$\vec{E}_x(t) + \sum_l \vec{V}_l \vec{x}_l(t) - \vec{x} = 0, \tag{4}$$

gdzie:

$\vec{E}_x(t)$ — wartość oczekiwana funkcji losowej $\vec{X}(t)$,

$\vec{V}_l$ — nieskorelowane zmienne losowe o wartościach oczekiwanych równych zeru,

$\vec{x}_l(t)$— pewne funkcje nielosowe.

W celu określenia funkcji losowej $\vec{T}(x)$ należy znaleźć pierwiastki układu równań (4). Jednakże w przypadku ogólnym, dla dowolnego procesu losowego $\vec{X}(t)$ układ (4) nie ma dokładnego rozwiązania i wówczas należy posługiwać się metodami numerycznymi.

W wyniku kolejnego zastosowania operatorów $\tilde{A}$ i A_T^{-1} możemy przedstawić funkcję losową $\vec{T}(x)$ w postaci nielosowej funkcji losowych argumentów:

$$\vec{T}(x) = \chi(\vec{V}_1, \vec{V}_2, \ldots, \vec{V}_k; \vec{x}). \tag{5}$$

Takie właśnie rozwiązanie pozwala na określenie kompletu statystycznych charakterystyk $\vec{T}(x)$. W tym celu należy wykorzystać wzory, określające rozkład funkcji losowych argumentów [5, 6]. Należy jednak podkreślić, że w zależności od sposobu rozwiązania równania (2) funkcja $\vec{T}(x)$ może być przedstawiona analitycznie, lub też w postaci stabelaryzowanej [7]. W związku z tym należy posługiwać się odrębnymi wzorami dla określenia poszukiwanych charakterystyk.

W przypadku, gdy $\vec{T}(x)$ ma postać analityczną, celowe jest wykorzystanie wzoru na m-wymiarową gęstość rozkładu [5, 6, 8]:

$$f_x(x_1, \ldots, x_m) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_{01}, x_{02}, \ldots, x_{0n})\, \delta\big(x_1 - \varphi_1(x_{01}, \ldots, x_{0n})\big) \ldots$$

$$\ldots \delta\big(x_m - \varphi_m(x_{01}, \ldots, x_{0n})\big) \tag{6}$$

gdzie:

$\delta(x_1 - \varphi_r(x_{01}, \ldots, x_{0n}))$ — funkcja delta Diracа,

$X_r(x_0) = \varphi_r(x_{01}, \ldots, x_{0n})$ — składowa m-wymiarowego wektora $\vec{X}(x_0)$, $r = 1, 2, \ldots, m$.

Stosując wzór (5), określamy statystyczne charakterystyki wektorowej funkcji $\vec{T}(x)$.

W przypadku, gdy $\vec{T}(x)$ zadane jest w postaci stabelaryzowanej dla określenia rozkładu funkcji losowej $\vec{T}(x)$ należy wykorzystać wzór:

$$\mathscr{F}_t(t_1, \ldots, t_m) = \int_{\mathscr{B}_v} \ldots \int_{(t,x)} f(x_1, \ldots, x_n) dx_1 \ldots dx_n, \tag{7}$$

gdzie $\mathscr{B}_v(t, x)$ — obszar wartości zmiennych $x_1, \ldots, x_n$.

Załóżmy, że obszar $V_i^{(j)}$ wszystkich możliwych wartości i-tego parametru, wchodzącego w skład wejściowej funkcji losowej $X_{0j}(t)$, podzielony został na n części $v_{iv}^{(j)}$ ($v = 1, 2, \ldots, \ldots, n$). W wyniku takiej kwantyzacji wielkości losowych $V_i^{(j)}$, $k_j \times m \times n$, wymiarowa przestrzeń została rozbita na $n^{k_j m}$ części, gdzie k_j jest ogólną liczbą parametrów charakteryzujących wejściową funkcję $X_{0j}(t)$, a m jest ogólną liczbą wejściowych funkcji losowych. Macierz charakteryzująca taką przestrzeń ma postać:

$$\|v_{iv}^{(j)}\|; \quad i = 1, 2, \ldots, k_j; \quad v = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, m. \tag{8}$$

Wykonując przekształcenia zgodnie z operatorami $A_T^{-1}\tilde{A}$ dla każdego wiersza macierzy (8), znajdziemy obszar całkowania $\mathscr{B}_v(t, x)$. Następnie, posługując się wzorem (7), określimy poszukiwany rozkład funkcji losowej $\vec{T}(x)$. Rzeczywiście, dla każdego wiersza macierzy (8) można napisać równania stochastyczne:

$$\tilde{A}(v_{1v_1}^{(1)} \quad v_{2v_2}^{(1)}, \ldots, v_{k_1 v_{k_1}}^{(1)}; \quad v_{1v_1}^{(2)}, \quad v_{2v_2}^{(2)}, \ldots, v_{k_2 v_{k_2}}^{(2)}, \ldots, v_{1v_1}^{(m)}, \quad v_{2v_2}^{(m)}, \ldots, v_{k_m v_{k_m}}^{(m)}; t) - x = 0, \tag{9}$$

które mogą być rozwiązane względem t za pomocą metod numerycznych. Przy wykorzystaniu operatora A_T^{-1} należy sprawdzić warunek:

$$\tau < t_u, \tag{10}$$

gdzie:

τ — znaleziona z równania (9) wartość argumentu t,

t_u — pewna ustalona wartość.

Jeżeli warunek (10) jest spełniony, wówczas rozpatrywany s-ty wiersz macierzy (8) oznaczamy jedynką, w przeciwnym przypadku zerem. Prawdopodobieństwo spełnienia nierówności (10) określamy jako:

$$P(t_u) = P(\tau < t_u) = \sum_s dP_{ivs}^{(j)} = \sum_s f(\{v_{ivs}^{(j)}\}) d\{v_{ivs}^{(j)}\}, \tag{11}$$

gdzie:

$dP_{ivs}^{(j)}$ — prawdopodobieństwo spełnienia warunku (10) w s-tym wierszu,

$\{v_{ivs}^{(j)}\}$ — wektor charakteryzujący s-ty wiersz macierzy,

$(f\{v_{ivs}^{(j)}\})$ — gęstość rozkładu wektora $\{v_{ivs}^{(j)}\}$ w s-tym wierszu,

$d\{v_{ivs}^{(j)}\}$ — różniczka wektora $\{v_{ivs}^{(j)}\}$.

Zadając różne wartości t_u dla ustalonego x można określić poszukiwany jednowymiarowy rozkład dla każdej składowej funkcji losowej $\vec{T}(x)$:

$$\mathscr{F}(t_u, x) = P(t_u, x). \tag{12}$$

Omówione operacje są operacjami całkowania numerycznego wyrażenia (7) w obszarze $\mathscr{B}_v(t, x)$. Jeżeli obszar $\mathscr{B}_v(t, x)$ ma prosty opis analityczny, to całkę (7) można stosunkowo łatwo obliczyć, w przeciwnym zaś przypadku należy stosować metodę modelowania statystycznego.

W praktyce przy badaniu niezawodności systemów nieliniowych określa się tylko momenty wyjściowej funkcji losowej $\vec{T}(x)$. Jedna z metod [7, 8] polega na przedstawieniu każdej składowej funkcji w postaci:

$$\vec{T}(x) = A_T^{-1}\tilde{A}E_{x_0}(t) + \sum_{n=1}^{\infty}\sum_{v_1\ldots v_n} V_{v_1}\ldots V_{v_n}x_{v_1\ldots v_n}(x_0), \tag{13}$$

gdzie:

$$x_{v_1\ldots v_n}(x_0) = \frac{1}{n!}\left[\frac{\partial^n}{\partial V_{v_1}\ldots\partial V_{v_n}}A_T^{-1}\tilde{A}\left\{E_{x_0}(t) + \sum_l V_l x_l(t)\right\}\right]_{V_l=0}. \tag{14}$$

Średni czas poprawnej pracy systemu, czyli wartość oczekiwana wyrażenia (13) ma następującą postać:

$$E_t(x) = A_T^{-1}\tilde{A}E_{x_0}(t) + \sum_{n=1}^{\infty}\sum_{v_1\ldots v_n}\mu_{v_1\ldots v_n}x_{v_1\ldots v_n}(x_0), \tag{15}$$

gdzie:

$$\mu_{v_1\ldots v_n} = E[V_{v_1}V_{v_2}\ldots V_{v_n}]. \tag{16}$$

W przypadku zlinearyzowania funkcji $A_T^{-1}\tilde{A}X_0(t)$ średni czas poprawnej pracy systemu równa się:

$$E_t(x) = \mathbf{A}_T^{-1}\tilde{A}E_{x_0}(t). \tag{17}$$

Metoda druga (metoda Dostupowa) opiera się na wykorzystaniu wzoru (13) przy organiczonym szeregu [8]:

$$\vec{T}(x) = A_T^{-1}\tilde{A}E_{x_0}(t) + \sum_{n=1}^{q}\sum_{v_1,\ldots,v_n=1} V_{v_1}\ldots V_{v_n}x_{v_1\ldots v_n}(x_0). \tag{18}$$

Znając momenty wektorowej funkcji losowej $\vec{T}(x)$ oraz wykorzystując wielomiany Hermite'a i Laguerra [2] można określić poszukiwaną wielowymiarową gęstość rozkładu $\vec{T}(x)$.

3. MACIERZOWA METODA OKREŚLENIA NIEZAWODNOŚCI SYSTEMÓW

Rozpatrzmy szczególny przypadek, gdy realizacje składowych wejściowej funkcji losowej $\vec{X}_0(t)$ przedstawiają sobą funkcje skokowe, przyjmujące w różnych chwilach czasu tylko dwie wartości, które oznaczymy przez „0" i „1". Zakładamy, że system opisywany jest ta-

kim operatorem $\tilde{A}$, że w dowolnej chwili czasu t każdemu zbiorowi stanów parametrów wejściowych odpowiada jeden z dwóch stanów funkcji losowej $\tilde{A}\vec{X}_0(t)$, które oznaczymy również przez „0" i „1".

Przy poczynionych założeniach zagadnienie określenia statystycznych charakterystyk funkcji losowej $\vec{T}(x)$ znacznie się upraszcza, ponieważ zamiast funkcji $\vec{T}(x)$ można rozpatrywać ciąg zmiennych losowych $(T_1, T_2, \ldots, T_l, \ldots)$. Poszczególne człony ciągu nie zależą od parametru x, ponieważ dla $0 < x < 1$ składowe procesu $\vec{T}(x)$ są niezależne od x. W związku z powyższym celowe jest przedstawienie operatora systemu w postaci funkcji boolowskiej parametrów wejściowych:

$$X_l = \Psi_l(X_{01}, X_{02}, \ldots, X_{0k}), \quad l = 1, 2, \ldots, m, \tag{19}$$

gdzie:

k — liczba wejściowych zmiennych boolowskich X_{0i},

m — liczba wyjściowych parametrów systemu.

W celu określenia niezawodności omówionych wyżej systemów posługujemy się aparatem matematycznym algebry stanów [1], przy czym istotne znaczenie ma kolejność występowania zdarzeń w czasie; dlatego też ogólna liczba stanów M systemu wynosi:

$$M = \sum_{i=1}^{k} A_k^i, \tag{20}$$

gdzie A_k^i — liczba permutacji z k elementów po i elementów.

Proces powstawania uszkodzeń w systemie można opisać macierzą

$$\mathcal{M} = ||X_{01}, X_{02}, \ldots, X_{0k}||, \tag{21}$$

która tworzy zupełną grupę wzajemnie wyłączających się zdarzeń. Poszczególnym zbiorom stanów parametrów wejściowych odpowiada stan „1" lub „0" na wyjściu systemu. Umówmy się, że stan „1" systemu jest stanem sprzyjającym, a stan „0" — niesprzyjającym. Wówczas prawdopodobieństwo stanu sprzyjającego określimy jako:

$$P_s = \sum_{j=0}^{s-1} P(H_j), \tag{22}$$

gdzie:

H_j — j-ty sprzyjający stan systemu,

s — liczba sprzyjających stanów systemu.

Jest oczywiste, że:

$$P(H_0) = \prod_{i=1}^{k} p_i = \prod_{i=1}^{k} \left(1 - \int_0^t f_i(\tau_i)\, d\tau_i\right), \tag{23}$$

gdzie:

p_i — prawdopodobieństwo tego, że w przedziale czasu (o, t) i-ty wejściowy parametr systemu znajdzie się w stanie „1",

$f_i(\tau_i)$ — gęstość rozkładu prawdopodobieństwa czasu τ do chwili przejścia i-tego parametru ze stanu „1" do stanu „0".

Przytoczone wyżej wzory zakładają niezależność zdarzeń dla stanu H_0, a także niemożliwość pojawienia się stanu „1" lub „0" dla każdego i-tego parametru więcej niż jeden raz. Nie zmienia to w zasadzie ogólności rozważań, ponieważ każdy przypadek wielokrotnego wystąpienia „1" lub „0" można sprowadzić do przypadku poprzedniego, zwiększając początkową liczbę zmiennych.

Przy określeniu prawdopodobieństwa warunkowego wystąpienia stanu $H_{\alpha\beta\ldots\varepsilon}$ w systemie, pod warunkiem, że parametry $X_{0\alpha}, X_{0\beta}, \ldots, X_{0\varepsilon}$ przyjmują w przedziale $(0, t)$ wartość „0", istotną rolę odgrywają chwile wystąpienia tych zdarzeń $\tau_\alpha, \tau_\beta, \ldots, \tau_\varepsilon$, czyli:

$$P(H_{\alpha\beta\ldots\varepsilon} | \bar{X}_{0\alpha}, \bar{X}_{0\beta}, \ldots, \bar{X}_{0\varepsilon}) = P_{\alpha\beta\ldots\varepsilon}(\tau_\alpha, \tau_\beta, \ldots, \tau_\varepsilon), \tag{24}$$

gdzie $\bar{X}_{0i}$ — wartość i-tego parametru wejściowego, odpowiadająca stanowi „0",

Powyższy fakt należy tłumaczyć tym, że z chwilą pojawienia się zdarzeń $\bar{X}_{0i} (i = \alpha, \beta, \ldots \varepsilon)$ zmienia się prawdopodobieństwo wystąpienia stanu „0" dla pozostałych parametrów.

Stosując twierdzenie o prawdopodobieństwie całkowitym, otrzymamy:

$$P^{(j)}(H_{\alpha\beta\ldots\varepsilon}) = \overbrace{\int_{T_j} \int \ldots \int}^{\xi \text{ razy}} P^{(j)}_{\alpha\beta\ldots\varepsilon}(\tau_\alpha, \tau_\beta, \ldots, \tau_\varepsilon) f^{(j)}_{\alpha\beta\ldots\varepsilon}(\tau_\alpha, \tau_\beta, \ldots, \tau_\varepsilon) d\tau_\alpha d\tau_\beta \ldots d\tau_\varepsilon, \tag{25}$$

gdzie:

$f^{(j)}_{\alpha\beta\ldots\varepsilon}(\tau_\alpha, \tau_\beta, \ldots, \tau_\varepsilon)$ — gęstość rozkładu prawdopodobieństwa zmiennych $\tau_\alpha, \tau_\beta, \ldots, \tau_\varepsilon$,

j — numer stanu $H_{\alpha\beta\ldots\varepsilon}$,

ξ — liczba zdarzeń $\bar{X}_{0\alpha}, \bar{X}_{0\beta}, \ldots, \bar{X}_{0\varepsilon}$ w j-tym stanie systemu.

Podstawiając (25) do (22), otrzymamy:

$$P_s = \sum_{j=1}^{s-1} \int_{T_j} \int \ldots \int P^{(j)}_{\alpha\beta\ldots\varepsilon}(\tau_\alpha, \tau_\beta, \ldots, \tau_\varepsilon) f^{(j)}_{\alpha\beta\ldots\varepsilon}(\tau_\alpha, \tau_\beta, \ldots, \tau_\varepsilon) d\tau_\alpha d\tau_\beta \ldots d\tau_\varepsilon + P(H_0). \tag{26}$$

W praktyce duże znaczenie mają procesy losowe, które można opisać w sposób następujący:

$$f^{(j)}_{\alpha\beta\ldots\varepsilon}(\tau_\alpha, \tau_\beta, \ldots, \tau_\varepsilon) = f_{\alpha j}(\tau^*_\alpha) f^{(\alpha)}_{\beta j}(\tau^*_\beta) \ldots f^{(\alpha\beta\ldots\rho)}_{\varepsilon j}(\tau^*_\varepsilon), \tag{27}$$

$$\tau_\alpha = \tau^*_\alpha; \quad \tau_\beta - \tau_\alpha = \tau^*_\beta; \quad \tau_\gamma - \tau_\beta = \tau^*_\gamma; \ldots; \tau_\varepsilon - \tau_\rho = \tau^*_\varepsilon; \quad t - \tau_\varepsilon =$$

$$= t - (\tau^*_\alpha + \tau^*_\beta + \ldots + \tau^*_\varepsilon) = \tau_t.$$

Wówczas po przekształceniach mamy:

$$P_s = \sum_{j=1}^{s-1} \int_0^t \int_0^{t-\tau^*_\alpha} \int_0^{t-\tau^*_\alpha-\tau^*_\beta-\ldots-\tau^*_\rho} \prod_{i=1}^{k} P_{ij}(\tau^*_\alpha) \prod_{l=1}^{k} P^{(\alpha)}_{lj}(\tau^*_\beta) \ldots \prod_{\nu=1}^{k} P^{(\alpha\beta\ldots\rho)}_{\nu j}(\tau^*_\varepsilon) \prod_{\eta=1}^{k} P^{(\alpha\beta\ldots\varepsilon)}_{\eta j}(\tau_t) \times$$

$$\times f_{\alpha j}(\tau^*_\alpha) f^{(\alpha)}_{\beta j}(\tau^*_\beta) \ldots f^{(\alpha\beta\ldots\rho)}_{\varepsilon j}(\tau^*_\varepsilon) d\tau^*_\alpha d\tau^*_\beta \ldots d\tau^*_\varepsilon, \tag{28}$$

$i \neq \alpha; \quad l \neq \alpha, \quad l \neq \beta; \ldots; \nu \neq \alpha, \quad \nu \neq \beta, \ldots, \nu \neq \varrho; \quad \eta \neq \alpha; \quad \eta \neq \beta, \ldots, \eta \neq \varepsilon;$

gdzie $P_{ij}(\tau^*_\alpha)$, $P^{(\alpha)}_{ij}(\tau^*_\beta)$, ..., $P^{(\alpha\beta\ldots\rho)}_{\nu j}(\tau^*_\varepsilon)$, $P^{(\alpha\beta\ldots\varepsilon)}_{\eta j}(\tau_t)$ — prawdopodobieństwa warunkowe zachowania przez parametry wejściowe stanu „1" w ciągu czasu τ^*_α, τ^*_β itd., pod warunkiem, że w j-tym stanie systemu zaszły zdarzenia $\bar{X}_{0\alpha}, \bar{X}_{0\beta}, \ldots, \bar{X}_{0\rho}, \bar{X}_{0\varepsilon}$.

W przypadku szczególnym, gdy proces opisany wzorami (27) jest jednorodnym procesem Markowa, można otrzymać stosunkowo proste wzory rekurencyjne [5, 8], pozwalające określić statystyczne charakterystyki parametrów wyjściowych systemu.

Na zakończenie należy podkreślić, że omówiona w części trzeciej artykułu metoda jest szczególnym przypadkiem metody, przedstawionej w części drugiej. Mianowicie, obszar $V_i^{(j)}$ i-tego parametru, wchodzącego w skład wejściowej funkcji $X_{0j}(t)$, wyczerpują wartości „0" i „1". Wówczas indeks ν macierzy (8) może przyjmować tylko dwie wartości.

4. METODA OKREŚLENIA NIEZAWODNOŚCI NA PODSTAWIE RÓWNAŃ RÓŻNICZKOWYCH KOŁMOGOROWA

W części trzeciej artykułu zostało pokazane, że w niektórych przypadkach operator systemu może być przedstawiony w postaci funkcji boolowskiej, opisanej macierzą (21). Na podstawie tej macierzy można sporządzić grafy systemu, przedstawiające wszystkie możliwe przejścia systemu z jednego stanu do drugiego. W pracach [5, 8] podano metodę określenia niezawodności systemów za pomocą układu równań różniczkowych Kołmogorowa, opartą na wykorzystaniu grafów. Sformułowano następującą regułę ułożenia poszczególnych równań na podstawie zadanego grafu systemu: pochodna prawdopodobieństwa warunkowego $P_{il}(\tau, t)$ znajdowania się systemu w chwili t w stanie H_l, pod warunkiem że w chwili τ znajdował się on w stanie H_i, równa się algebraicznej sumie iloczynów intensywności przejść przez odpowiednie prawdopodobieństwa, przy czym składnikom, odpowiadającym wychodzącym strzałkom grafu, przypisujemy znak minus, pozostałym zaś plus. Innymi słowy:

$$\frac{\partial P_{il}(\tau, t)}{\partial t} = -\Lambda_l(t) P_{il}(\tau, t) + \sum_j P_{ij}(\tau, t) \Lambda_j(t) p_{jl}(t), \tag{29}$$

gdzie:

$P_{il}(\tau, t)$ — prawdopodobieństwo warunkowe znajdowania się systemu w chwili t w stanie H_l, pod warunkiem że w chwili τ znajdował się on w stanie H_i:

$\Lambda_l(t)$ — sumaryczna intensywność wyjścia systemu ze stanu H_l;

$p_{jl}(t)$ — prawdopodobieństwo warunkowe przejścia systemu ze stanu H_j do stanu H_l, pod warunkiem, że system wyszedł ze stanu H_j;

$\Lambda_j(t) p_{jl}(t)$ — intensywność przejścia systemu ze stanu H_j do stanu H_l.

Zaznaczono przy tym, że proces losowy, opisywany za pomocą układu równań (29), powinien spełniać szereg warunków, mianowicie:

$$\begin{aligned} &\lim_{\Delta t \to 0} \frac{1 - P_{ll}(t, t+\Delta t)}{\Delta t} = \Lambda_l(t), \\ &\lim_{\Delta t \to 0} \frac{P_{jl}(t, t+\Delta t)}{\Delta t} = \Lambda_j(t) p_{jl}(t), \quad j \neq l, \\ &\sum_l p_{jl}(t) = 1; \quad p_{jj}(t) = 0. \end{aligned} \tag{30}$$

Dla wielu systemów warunki (30) można przedstawić w postaci:

$$\Lambda_l(t) = \sum_{q=1}^{m} \lambda_q^{(l)}(t); \quad \Lambda_j(t) = \sum_{\nu=1}^{n} \lambda_\nu^{(j)}(t); \quad p_{jl}(t) = \frac{\lambda_l^{(j)}(t)}{\Lambda_j(t)}, \tag{31}$$

gdzie:

$\lambda_q^{(l)}(t)$ — intensywność przejścia systemu ze stanu H_l do stanu H_q,

m — ogólna liczba przejść systemu ze stanu H_l do innych stanów,

n — ogólna liczba przejść systemu ze stanu H_j do innych stanów.

Wówczas układ równań (29) przyjmie postać:

$$\frac{\partial P_{il}(\tau, t)}{\partial t} = -\sum_{q=1}^{m} \lambda_q^{(l)}(t) P_{il}(\tau, t) + \sum_{v=1}^{n} \lambda_v^{(j)}(t) P_{ij}(\tau, t). \tag{32}$$

W przypadku ogólnym rozwiązanie układu (32) przedstawia sobą dość skomplikowany problem, zwłaszcza gdy liczba równań jest większa niż trzy. W związku z tym wyłania się zagadnienie zamiany danego układu równań przez taki układ równań różniczkowych Kołmogorowa ze stałymi współczynnikami λ, który w sensie statystycznym jest równoważny zadanemu układowi w przedziale czasu (τ, t). Wybór kryterium równoważności statystycznej może być różny. Będziemy uważali, że realny system techniczny jest równoważny pewnemu systemowi zastępczemu, jeżeli każde lokalne przejście z jednego stanu do drugiego w systemie realnym może być zamienione przez taką grupę przejść mających stałe intensywności, że wypadkowe prawdopodobieństwo warunkowe przejścia w tej grupie zbliża się z zadaną dokładnością do prawdopodobieństwa warunkowego przejścia w systemie realnym.

Rozpatrzmy teraz kilka przykładów systemów, w których intensywności przejść są stałe w czasie. Załóżmy, że proces opisany wzorami (27) jest jednorodnym procesem Markowa. Intensywności przejść $\lambda_i, \lambda_l^{(\alpha)}, \ldots, \lambda_v^{(\alpha\beta\ldots\rho)}$ itd. są znane, przy czym górne indeksy oznaczają zajście zdarzeń $\bar{X}_{0\alpha}, \bar{X}_{0\alpha}, \bar{X}_{0\beta}, \ldots, \bar{X}_{0\alpha}, \bar{X}_{0\beta}, \ldots, \bar{X}_{0\rho}$ w chwilach τ_α^*, τ_β^* itd. Załóżmy również, że rozpatrywany system jest systemem bez odnowy, tj. nie ma powrotu z danego stanu do stanów poprzednich. Układ równań różniczkowych Kołmogorowa dla stanów $H_0, H_\alpha, H_{\alpha\beta}, H_{\alpha\beta\gamma}, \ldots$ itd. ma następującą postać:

$$\begin{aligned}
&\dot{P}_0(t) = -\sum_{i=1}^{k} \lambda_i P_0(t), \\
&\dot{P}_\alpha(t) = \lambda_\alpha P_0(t) - \sum_{l=1}^{k} \lambda_l^{(\alpha)} P_\alpha(t), \quad l \neq \alpha, \\
&\dot{P}_{\alpha\beta}(t) = \lambda_\beta^{(\alpha)} P_\alpha(t) - \sum_{j=1}^{k} \lambda_j^{(\alpha\beta)} P_{\alpha\beta}(t), \quad j \neq \alpha, \quad j \neq \beta, \\
&\dot{P}_{\beta\alpha}(t) = \lambda_\alpha^{(\beta)} P_\beta(t) - \sum_{j=1}^{k} \lambda_j^{(\beta\alpha)} P_{\beta\alpha}(t), \quad j \neq \alpha, \quad j \neq \beta, \\
&\dot{P}_{\alpha\beta\ldots\rho\varepsilon}(t) = \lambda_\varepsilon^{(\alpha\beta\ldots\rho)} P_{\alpha\beta\ldots\rho}(t) - \sum_{\Theta=1}^{k} \lambda_\Theta^{(\alpha\beta\ldots\varepsilon)} P_{\alpha\beta\ldots\varepsilon}(t), \\
&\Theta \neq \alpha, \quad \Theta \neq \beta, \ldots, \Theta \neq \varepsilon.
\end{aligned} \tag{33}$$

Układ równań różniczkowych Kołmogorowa dla systemów z odnową, a do takich należy większość systemów technicznych, ma postać:

$$\dot{P}_0(t) = -\sum_{i=1}^{k} \lambda_i P_0(t) + \sum_{i=1}^{k} \mu_i^{(i)} P_i(t),$$

$$\dot{P}_\alpha(t) = \lambda_\alpha P_0(t) + \sum_{l=1}^{k} \mu_l^{(\alpha l)} P_{\alpha l}(t) - \left(\mu_\alpha^{(\alpha)} + \sum_{l=1}^{k} \lambda_l^{(\alpha)}\right) P_\alpha(t), \quad l \neq \alpha,$$

$$\dot{P}_{\alpha\beta}(t) = \lambda_\beta^{(\alpha)} P_\alpha(t) - \sum_{j=1}^{k} \mu_j^{(\alpha\beta j)} P_{\alpha\beta j}(t) - \left(\mu_\beta^{(\alpha\beta)} + \sum_{j=1}^{k} \lambda_j^{(\alpha\beta)} P_{\alpha\beta}(t)\right), \quad j \neq \alpha, j \neq \beta, \tag{34}$$

$$\dot{P}_{\beta\alpha}(t) = \lambda_\alpha^{(\beta)} P_\beta(t) + \sum_{j=1}^{k} \mu_j^{(\beta\alpha j)} P_{\beta\alpha j}(t) - \left(\mu_\alpha^{(\beta\alpha)} + \sum_{j=1}^{k} \lambda_j^{(\beta\alpha)}\right) P_{\beta\alpha}(t)\Big), \quad j \neq \alpha, j \neq \beta,$$

$$\dot{P}_{\alpha\beta\ldots\rho\varepsilon}(t) = \lambda_\varepsilon^{(\alpha\beta\ldots\rho)} P_{\alpha\beta\ldots\rho}(t) + \sum_{\nu=1}^{k} \mu_\nu^{(\alpha\beta\ldots\varepsilon\nu)} P_{\alpha\beta\ldots\rho\varepsilon}(t),$$

$$\nu \neq \alpha, \quad \nu \neq \beta, \ldots, \nu \neq \varepsilon,$$

gdzie: $\mu_i^{(i)}, \mu_l^{(\alpha l)}, \ldots, \mu_\nu^{(\alpha\beta,\ldots\varepsilon\nu)}$ oznaczają intensywność powrotu systemu ze stanu H_i do stanu H_0, ze stanu $H_{\alpha l}$ do stanu H_α itd.

Układy równań różniczkowych (33) i (34) rozwiązujemy za pomocą znanych metod, opartych na wektorowo-macierzowym przedstawieniu układu [3, 4]:

$$\frac{d\vec{x}(t)}{dt} = \vec{B}\vec{x}(t),$$

$$x(t) = \begin{Vmatrix} P_0(t) \\ P_\alpha(t) \\ P_{\alpha\beta}(t) \\ \vdots \\ P_{\alpha\beta\ldots\rho\varepsilon}(t) \end{Vmatrix}, \quad \vec{B} = \{a_{ij}\}, \tag{35}$$

gdzie a_{ij} — intensywności przejść λ_{ij} lub μ_{ij}.

Jednakże metody te są bardzo pracochłonne i wymagają w praktyce zastosowania metod numerycznych.

Obecnie przy rozwiązywaniu układów (33) i (34) szerokie zastosowanie znalazła metoda oparta na wykorzystaniu przekształcenia Laplace'a:

$$\tilde{x}_i(p) = \int_0^\infty x_i(t) e^{-pt} dt, \tag{36}$$

gdzie $x_i(t)$ jest odpowiednikiem prawdopodobieństwa warunkowego $P_{il}(t)$ w równaniach.

W wyniku przekształcenia (36) otrzymamy układ algebraicznych równań liniowych względem $\tilde{x}_i(p)$. Stosując następnie odwrotne przekształcenie Laplace'a, znajdziemy poszukiwaną funkcję:

$$x_i(t) = \frac{1}{2\pi_j} \int_{\sigma-j\infty}^{\sigma+j\infty} \frac{\Delta_i p}{\Delta p} e^{pt} dp, \tag{37}$$

gdzie:

Δp — wyznacznik układu równań algebraicznych,

$\Delta_i p$ — wyznacznik otrzymany drogą zamiany elementów i-tej kolumny wyznacznika Δp przez człony wolne.

W tym przypadku napotkamy również duże trudności rachunkowe, ponieważ przy rozwiązywaniu całki (37) zachodzi konieczność znalezienia pierwiastków równania algebraicznego wysokiego rzędu.

Proces powstawania uszkodzeń i odnowy w systemie możemy porównać z przypadkowym ruchem pewnego punktu w przestrzeni stanów systemu. Jeżeli punkt ten napotyka na swej drodze ekran odbijający, mówimy wówczas, że istnieją przejścia systemu ze stanu niesprzyjającego do stanu sprzyjającego, jeżeli natomiast punkt napotyka ekran pochłaniający, przejście to jest niemożliwe. W ostatnim przypadku prawdopodobieństwo znajdowania się systemu w chwili t w stanie niesprzyjającym określa rozkład czasu T, w ciągu którego system znajdzie się w tym stanie:

$$\mathscr{F}(t) = P(T < t). \tag{38}$$

W terminach teorii niezawodności wzór (38) określa po prostu rozkład czasu pracy do pierwszego uszkodzenia. Na podstawie (38) można znaleźć moment n-tego rzędu zmiennej losowej T, który w przypadku ogólnym równa się:

$$\alpha_n = (-1)^n \frac{d^n}{dp^n}\left[p\frac{\Delta_i p}{\Delta p}\right]\Bigg|_{p=0}, \tag{39}$$

gdzie $x_i(p) = \dfrac{\Delta_i(p)}{\Delta p} = \mathscr{F}(p)$.

Praktyczne wykorzystanie wzoru (39), mimo jego pozornej prostoty, przysparza dużych trudności, gdyż rząd wyznaczników $\Delta_i(p)$ i Δp jest zazwyczaj bardzo wysoki.

W szeregu przypadków dla α_n można znaleźć stosunkowo proste wzory rekurencyjne, np. dla systemu bez odnowy możemy napisać [8]:

$$\alpha_{\alpha\beta\ldots\rho\varepsilon n} = \frac{\lambda_\varepsilon^{(\alpha\beta\ldots\rho)}\alpha_{\alpha\beta\ldots\rho n} - n\alpha_{\alpha\beta\ldots\rho\varepsilon(n-1)}}{\sum_{\Theta=1}^{k}\lambda_\Theta^{(\alpha\beta\ldots\rho\varepsilon)}}. \tag{40}$$

W pracach [3, 4] przedstawiono metodę przybliżonego określenia rozkładu $\mathscr{F}(t)$. Istota tej metody polega na przedstawieniu funkcji gęstości rozkładu $f(t)$ w postaci szeregu:

$$f(t) = f_0(t)[c_0 p_0(t) + c_1 p_1(t) + \ldots], \tag{41}$$

gdzie:

$f_0(t)$ — pewna „wzorcowa" gęstość rozkładu,

$p_0(t), p_1(t)$ — wielomiany ortonormalne względem funkcji $f_0(t)$,

c_0, c_1 — współczynniki określane na podstawie momentów α_n.

Dla przypadków systemów z odnową i bez odnowy jako gęstość wzorcową wykorzystuje się funkcję:

$$f_0(t) = \frac{e^{-t}t^\alpha}{\Gamma(\alpha+1)}, \tag{42}$$

gdzie $\Gamma(\alpha+1)$ — funkcja gamma.

Jako wielomiany $p_n(t)$ wykorzystuje się wielomiany Laguerra [2]:

$$p_n(t) = L_n^{(\alpha)}(t) = \sum_{k=0}^{n} \binom{n+k}{n-k} \frac{(-t)^k}{k!}. \tag{43}$$

Wówczas poszukiwana gęstość rozkładu czasu T będzie miała postać:

$$f(t) = \frac{e^{-t}t^{\alpha}}{\Gamma(\alpha+1)} \sum_{n=0}^{\infty} c_n L_n^{(\alpha)}(t), \tag{44}$$

gdzie:

$$c_n = \frac{\sum_{k=0}^{n} \frac{(-1)^k}{k!} \binom{n+\alpha}{n-k} \alpha_k}{\binom{n+\alpha}{n}}. \tag{45}$$

Momenty k-tego rzędu α_k określamy wg wzorów (39) lub (40).

Zaleta proponowanej metody polega na tym, że określenie momentów jest znacznie łatwiejsze, niż rozwiązywanie układów równań różniczkowych względem poszczególnych funkcji.

5. KRYTYCZNA OCENA KLASYCZNYCH METOD OKREŚLENIA NIEZAWODNOŚCI

Metody określenia niezawodności systemów, przedstawione w ramach klasycznej teorii niezawodności, są aktualne do dnia dzisiejszego i tworzą jeden z jej podstawowych rozdziałów. Metody te reprezentują pewne modele matematyczne realnych zjawisk fizycznych, zachodzących w systemach. Praktyczna przydatność metody oceny niezawodności danego konkretnego systemu zależy zarówno od tego, w jakim stopniu wybrany model matematyczny odzwierciedla rzeczywisty proces funkcjonowania systemu, jak i od stopnia skomplikowania operacji rachunkowych, związanych z wybraną metodą obliczeń. Należy zauważyć, że obecnie wykorzystywana jest tylko nieliczna część modeli, która nie obejmuje wszystkich, częstokroć nawet podstawowych zależności, mających miejsce w systemach.

Istotną wadą stosowanych powszechnie metod jest to, że zakładają one w przeważającej liczbie przypadków poissonowski model powstawania uszkodzeń w systemie, który jest szczególnym przypadkiem procesów Markowa. Jak wiadomo poissonowskiemu procesowi powstawania uszkodzeń w systemie odpowiada wykładniczy rozkład czasu pracy do pierwszego uszkodzenia [5, 6]. Rozkład ten posiada pewną charakterystyczną cechę wyróżniającą go spośród wszystkich innych rozkładów, mianowicie: dla rozkładu wykładniczego prawdopodobieństwo poprawnej pracy w zadanym przedziale czasu

$(t, t+\tau)$ nie zależy od długości czasu t poprzedniej jego pracy, lecz tylko od długości przedziału τ [5, 6].

Podstawową przyczyną popularności i szerokiego zastosowania rozkładu wykładniczego w dziedzinie niezawodności jest stosunkowa łatwość obliczeń i możliwość otrzymania w wielu przypadkach rozwiązania w postaci zamkniętej. Jednakże mimo tak istotnych zalet, rozkład wykładniczy ma bardzo ograniczony zasięg stosowalności. Przyczyna tkwi w tym, że model uszkodzeń odpowiadający rozkładowi wykładniczemu sprzeczny jest z wieloma pojęciami natury fizycznej, ponieważ systemy o wykładniczym rozkładzie czasu poprawnej pracy są w istocie swej systemami niestarzejącymi się.

Druga istotna wada polega na tym, że stosowane obecnie metody obliczenia niezawodności systemów przeważnie oparte są na założeniu, że system składa się z samodzielnych, niezależnych w sensie niezawodnościowym, jednostek-elementów. W związku z tym nie uwzględnia się zależności funkcjonalnej między parametrami poszczególnych elementów, czyli eliminuje się z rozpatrzenia to najważniejsze, co odróżnia dany konkretny system od przypadkowego zbioru elementów. Upraszcza to w znacznej mierze obliczenie niezawodności, ale w wielu przypadkach przedstawia zniekształcony lub wręcz błędny obraz względnej niezawodności poszczególnych wariantów systemu.

Wada ta może być zmniejszona w pewnym stopniu przez wprowadzenie zależności charakterystyk niezawodności elementów od warunków ich pracy. Z drugiej strony metody te rozwijają się w kierunku analizy wpływu poszczególnych uszkodzeń na prawdopodobieństwo pojawienia się innych uszkodzeń. Do takich metod należą między innymi: metoda macierzowa i metoda obliczeń niezawodności systemów za pomocą równań różniczkowych Kołmogorowa.

Metoda macierzowa jest metodą dość ogólną, ponieważ nie nakłada ona ograniczeń na strukturę systemu i postać rozkładu zmiennych losowych. Jednakże praktyczne zastosowanie tej metody ograniczone jest z następujących powodów.

1. Aparat matematyczny metody jest dość skomplikowany, a rozwiązanie analityczne w zamkniętej postaci udało się otrzymać tylko dla wykładniczego rozkładu czasu poprawnej pracy poszczególnych elementów systemu. W przypadku, gdy rozkłady zmiennych losowych różnią się od rozkładu wykładniczego, zachodzi konieczność stosowania metod numerycznych.

2. Matematyczny model systemu jest modelem z pełną informacją statystyczną, w związku z czym ilość informacji początkowej, niezbędnej dla określenia niezawodności systemów wielkich, musi być bardzo duża.

3. W praktyce na podstawie danych statystycznych nie ma możliwości określenia, przy dużej liczbie parametrów wejściowych, k-wymiarowej gęstości rozkładu $f_{\alpha\beta\ldots\varepsilon}(\tau_\alpha, \tau_\beta, \ldots, \tau_\varepsilon)$ oraz prawdopodobieństwa $P_{\alpha\beta\ldots\varepsilon}(\tau_\alpha, \tau_\beta, \ldots, \tau_\varepsilon)$. W związku z tym w celu określenia prawdopodobieństwa $P^{(j)}(H_{\alpha\beta\ldots\varepsilon})$ nie można, praktycznie rzecz biorąc, korzystać ani ze wzoru (25), ani też stosować modelowania statystycznego. W takich przypadkach można proponować tylko badania naturalne.

4. Nie uwzględnia się zależności funkcjonalnych w systemie, a także wpływu warunków zewnętrznych. Pomija się tym samym przyczyny i sposób formowania procesów losowych.

Metoda określenia niezawodności systemów, oparta na wykorzystaniu równań różniczkowych Kołmogorowa, jest metodą dość uniwersalną pozwalającą na badanie niezawodności zarówno systemów opisywanych za pomocą jednorodnych procesów Markowa, jak i systemów o bardziej złożonych procesach. Przejście od procesów niemarkowowskich do procesów markowowskich nie przedstawia w zasadzie dużych trudności, jednakże dokładne rozwiązanie może być otrzymane tylko dla przypadków szczególnych. Metoda Kołmogorowa posiada również wszystkie te wady, co i metoda macierzowa.

Metody określenia niezawodności przedstawione w części drugiej są metodami ogólnymi, lecz nie zdają egzaminu w praktyce. Są to metody mało przejrzyste i trudne do wykorzystania w konkretnych zastosowaniach.

Rozwiązanie w zamkniętej postaci udało się otrzymać jedynie dla pewnych przypadków szczególnych. Próby przybliżonego określenia niezawodności również nie dają zadowalających wyników, ponieważ:

1) metody analityczne dają często rezultaty obarczone niedopuszczalnie dużym błędem, zniekształcając w ten sposób rzeczywisty obraz niezawodności badanego systemu,

2) metody numeryczne, nawet przy zastosowaniu maszyn cyfrowych, wymagają zbyt dużego czasu i środków, szczególnie przy wzroście złożoności problemu.

Należy zauważyć, że stosowane w omawianych metodach sposoby opisu procesów losowych i systemów są nieadekwatne do współczesnej techniki obliczeniowej i stosowany w klasycznej teorii formalizm opisu systemów przy pomocy funkcji delta Diraca jest nieodpowiedni dla programowania na maszynach cyfrowych czy analogowych.

Reasumując należy stwierdzić, że ani metody ogólne, ani opracowanie pewnych szczególnych modeli matematycznych nie rozwiązuje na razie problemu stworzenia inżynierskiej metodyki określenia niezawodności systemów złożonych.

Na tle tych uwag wydaje się, że jedyną drogą postępowania w tym przypadku jest modelowanie procesu funkcjonowania systemu na elektronicznych maszynach cyfrowych i dlatego przy rozpracowywaniu w Instytucie Telekomunikacji Politechniki Gdańskiej metody oceny niezawodności dla powstającego w kraju systemu telefonii 960-krotnej TN 960, skoncentrowano się na metodzie modelowania statystycznego zwanej również metodą Monte-Carlo.

Metody modelowania statystycznego są szczególnie przydatne do wyznaczania wielkości, które definiowane są jako wartości oczekiwane pewnych funkcji wielkości przypadkowych. Istota metody modelowania statystycznego polega na zbudowaniu modelu probabilistycznego rozpatrywanego zagadnienia, realizacji go w sposób losowy oraz przyjęciu otrzymanych wyników jako przybliżonego rozwiązania zadania.

Metoda Monte-Carlo jest bardzo cennym narzędziem w dziedzinie rozwiązywania zagadnień niezawodnościowych, ponieważ pozwala sprowadzić ogrom tych zagadnień do stosunkowo niewielkiej liczby uniwersalnych algorytmów, charakteryzujących się prostotą i łatwością realizacji na nowoczesnych elektronicznych maszynach cyfrowych.

Zastosowanie metody modelowania statystycznego do oceny niezawodności systemów telekomunikacyjnych zostanie przez autorów przedstawione w oddzielnej pracy.

Załącznik

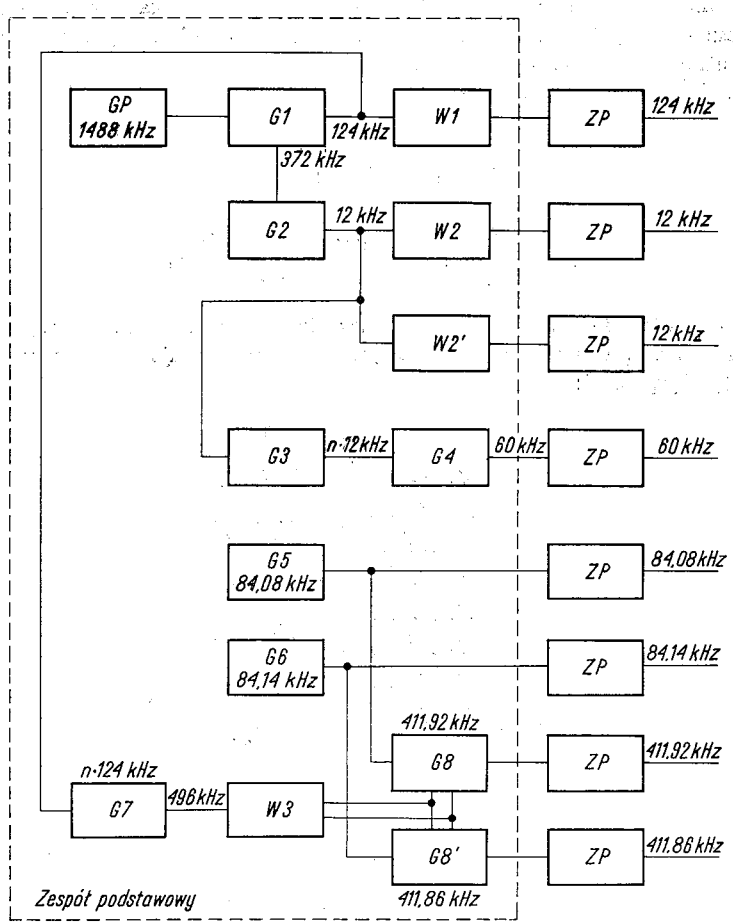
OBLICZENIE NIEZAWODNOŚCI URZĄDZEŃ STOJAKA GENERACYJNEGO SYSTEMU TN-300

Urządzenia generacyjne systemu TN-300, które były przedmiotem analizy, generują częstotliwości podstawowe 12 i 124 kHz oraz umożliwiają wytwarzanie:

prądów nośnych,
liniowych prądów pilotowych,
prądu pilotowego grupy pierwotnej,
prądu pilotowego grupy wtórnej.

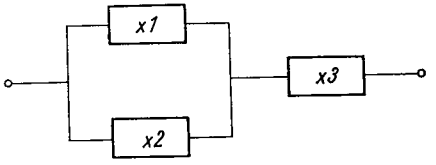
Ze względu na zasadniczą rolę jaką odgrywają urządzenia generacyjne w systemach telefonii nośnej, są one rezerwowane.

W systemie TN-300 zastosowano rezerwację gorącą, co oznacza, że podstawowy i rezerwowy zespół generacyjny znajdują się w jednakowych warunkach pracy. Schemat blokowy stojaka generacyjnego systemu TN-300 do wytwarzania częstotliwości podstawowych oraz pilotujących grup pierwotnych i wtórnych przedstawiony został na rys. 1. Przełączenie zespołów w przypadku uszkodzenia odbywa się za pomocą



Rys. 1. Schemat blokowy stojaka generacyjnego systemu TN-300 do wytwarzania częstotliwości podstawowych oraz pilotujących G_p i G_w

zespołów przełączających ZP. Zespoły generacyjne, zarówno podstawowy, jak i rezerwowy, posiadają szeregową strukturę niezawodnościową, ponieważ do niezawodnej pracy systemu generacyjnego TN-300 konieczna jest poprawna praca wszystkich zespołów. W związku z powyższym niezawodnościowy model stojaka generacyjnego można przedstawić w sposób jak na rys. 2, na którym przez $X1$ oznaczono podstawowy zespół generacyjny, przez $X2$ — rezerwowy zespół generacyjny, a przez $X3$ — zespoły przełączające.



Rys. 2. Schemat blokowy do oceny niezawodności stojaka generacyjnego systemu TN-300

Niezawodność takiego układu określimy za pomocą metody macierzowej. Zakładamy przy tym, że uszkodzenia poszczególnych zespołów są niezależne a kolejność występowania uszkodzeń nie odgrywa roli. Niezawodność zespołów $X1$, $X2$ i $X3$ została obliczona w oparciu o probabilistyczny model powstawania uszkodzeń Weibulla dla poszczególnych elementów tych zespołów.

Po przeprowadzeniu obliczeń otrzymano:

$$p_1 = p_2 = \exp\left[-\left(\frac{t}{27\,500}\right)^{3,8}\right],$$

$$p_3 = \exp\left[-2\left(\frac{t}{40\,400}\right)^{3,8} + \left(\frac{t}{34\,300}\right)^{3,8}\right],$$

gdzie p_1, p_2, p_3 — prawdopodobieństwo poprawnej pracy odpowiednio zespołów $X1$, $X2$, $X3$.

Macierz, przedstawiająca zupełną grupę zdarzeń, w których może znajdować się rozpatrywany układ, będzie miała następującą postać:

$$M = \begin{array}{l} \text{Stan} \\ \left[\begin{array}{lll} X1 & X2 & X3 \end{array}\right] \rightarrow \mathrm{H}_0(\text{„1”}) \\ \left.\begin{array}{lll} \overline{X1} & X2 & X3 \\ X1 & \overline{X2} & X3 \end{array}\right\} \rightarrow H_1\ (\text{„1)”}) \\ \left[\begin{array}{lll} X1 & X2 & \overline{X3} \end{array}\right] \rightarrow H_2\ (\text{„0”}) \\ \left.\begin{array}{lll} \overline{X1} & \overline{X2} & X3 \\ \overline{X2} & \overline{X1} & X3 \end{array}\right\} \rightarrow H_3\ (\text{„0”}) \\ \left.\begin{array}{lll} \overline{X1} & X2 & \overline{X3} \\ \overline{X3} & X2 & \overline{X1} \\ X1 & \overline{X2} & \overline{X3} \\ X1 & \overline{X3} & \overline{X2} \end{array}\right\} \rightarrow H_4\ (\text{„0”}) \\ \left.\begin{array}{lll} \overline{X1} & \overline{X2} & \overline{X3} \\ \overline{X2} & \overline{X1} & \overline{X3} \\ \overline{X1} & \overline{X3} & \overline{X2} \\ \overline{X3} & \overline{X1} & \overline{X2} \\ \overline{X2} & \overline{X3} & \overline{X1} \\ \overline{X3} & \overline{X2} & \overline{X1} \end{array}\right\} \rightarrow H_5\ (\text{„0”}) \end{array}$$

Stany H_0 i H_1 są stanami sprzyjającymi, tj. odpowiadają poprawnej pracy urządzeń generacyjnych systemu TN-300. Prawdopodobieństwo znajdowania się systemu w tych stanach równa się:

$$P(H_0) = \prod_{i=1}^{3} p_i,$$

$$P(H_1) = (1-p_1)p_2p_3+p_1(1-p_2)p_3.$$

Na podstawie wzoru (22) prawdopodobieństwo poprawnej pracy rozpatrywanych urządzeń jest równe:

$$P_s = P(H_0)+P(H_1).$$

Obliczenie funkcji $P_s(t)$ przeprowadzono na maszynie cyfrowej. Wyniki zostały przedstawione w tablicy 1.

Tablica 1

Wyniki obliczeń funkcji $P_s(t)$ określającej prawdopodobieństwo poprawnej pracy stojaka generacyjnego TN-300

t[h]	$P_s(t)$
0	1
4 000	0,9993
8 000	0,9906
12 000	0,9553
16 000	0,8640
20 000	0,6853
24 000	0,4288
28 000	0,1849
32 000	0,0477
36 000	0,0064

Po wytestowaniu otrzymanych wyników za pomocą siatki Weibulla mamy:

$$P_s(t) = \exp\left[-\left(\frac{t}{24\,800}\right)^{4,023}\right].$$

Znając prawdopodobieństwo poprawnej pracy urządzeń generacyjnych systemu TN-300 można określić średni czas poprawnej pracy:

$$T_s = 24\,800\Gamma\left(1+\frac{1}{4,023}\right) = 22\,471 \text{ [h]},$$

gdzie $\Gamma(x)$ — jest funkcją gamma.

Uzyskany średni czas poprawnej pracy, który dla półprzewodnikowej wersji stojaka generacyjnego systemu TN-300 wynosi około dwa i pół roku, można będzie zwiększyć w wyniku zastosowania obwodów scalonych.

LITERATURA CYTOWANA W TEKŚCIE

1. E. Berki, Simwoliczeskaja logika i razumnyje maszyny, Izd. Inostr. Lit., Moskwa 1961.
2. W. Ł. Daniłow i współautorzy, Funkcje, granice, szeregi, ułamki łańcuchowe, PWN, Warszawa 1970.
3. B. P. Demidowicz, J. A. Maron, Osnowy wycislitelnoj matematyki, Nauka, Moskwa 1966.
4. B. P. Demidowicz, J A. Maron, E. Z. Szuwałowa, Cislennyje metody analiza, Fizmatgiz, Moskwa 1963.
5. B. W. Gniedenko, J. K. Bielajew, A. D. Sołowiew, Matematyczne metody w teorii niezawodności, PWN, Warszawa 1968.
6. M. Fisz, Rachunek prawdopodobieństwa i statystyka matematyczna, PWN, Warszawa 1958.
7. B. S. Pugaczew, Teoria słuczajnych funkcji, Fizmatgiz, Moskwa 1960.
8. B. W. Wasiliew, B. A. Kozłow, L. G. Tkaczenko, Nadiożnost i efektiwnost radioelektronnych ustrojstw, Sow. Radio, Moskwa 1964.

O. CHOREŃ, M. ZIENTALSKI

METHODS OF RELIABILITY MEASURE OF COMMUNICATION NETWORK

Summary

The article discusses methods for analysing and solving a reliability of communication networks. The reliability of systems, which are described by means of any known operator, along with the matrix method of reliability calculation and method based on Komogorov's differential equations have been established. Finally, a critical opinion of presented methods of reliability measure is performed. As an example a numerical calculation of reliability parameters for the multiple telephone network roch, type TN-300 is presented.

O. CHOREŃ, M. ZIENTALSKI

BEWERTUNGSMETHODEN DER ZUVERLÄSSIGKEIT VON SYSTEMEN DER FERNMELDETECHNIK

Zusammenfassung

Im Artikel wurden Bestimmungsmethoden der Zuverlässigkeit von Systemen der Fernmeldetechnik erwogen. Es wurde die Zuverlässigkeit der beschriebenen Systeme mittels eines beliebigen Operators bestimmt, das Matrixverfahren für die Sicherheitsbestimmung sowie die Methoden auf Basis von Kołmogorowschen Differenzialgleichungen angeführt. Die vorgetragenen Bestimmungsmethoden der Sicherheit wurden kritisch analysiert und bewertet. Am Beispiel eines Generationsständers vervielfachter Fernsprechtechnik Typ TN-300 wurden Parameterberechnungen der Zuverlässigkeit vollzogen.

О. ХОРЕНЬ, М. ЗЕНТАЛЬСКИ

МЕТОДЫ ОПРЕДЕЛЕНИЯ НАДЕЖНОСТИ СИСТЕМ ДАЛЬНЕЙ СВЯЗИ

Резюме

В статье рассмотрены методы определения надежности систем дальней связи. Определена надежность систем описываемых с помощью произвольного заданного оператора. Представлен матричный метод определения надежности, а также метод определения надежности с помощью дифференциальных уравнений Колмогорова. Проведена критическая оценка представленных методов. В заключение дан пример определения характеристик надежности генерирующих устройств для многоканальной системы связи TN-300.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 1, ss. 59-68

621.317.33

Metoda czasowego równoważenia mostka w pomiarach zmian rezystancji

LECH JAN WEISS (POZNAŃ)
Instytut Elektrotechniki Przemysłowej Politechniki Poznańskiej

Otrzymano 31.1.1973

W artykule opisano mostkową metodę pomiaru zmian rezystancji. Metoda polega na tym, że rezystancja jednego z ramion, będąca elementem regulacyjnym w mostku, zmienia swoją wartość liniowo w czasie. Dzięki temu mierzona zmiana rezystancji przetworzona zostaje na analogowy odcinek czasu. Czas ten mierzony numerycznie może wyrażać bezpośrednio wartość wielkości mierzonej. Jako rezystor regulacyjny w mostku zastosowano tranzystor polowy.

1. WSTĘP

Pomiary zmian rezystancji zyskały w ostatnich czasach szczególnie duże znaczenie. W dziedzinie pomiarów elektrycznych stanowią one podstawę określania tolerancji wykonania wszelkich elementów rezystancyjnych, a mianowicie oporników, elektrycznych elementów grzejnych i rezystancyjnych przetworników pomiarowych. Przetworniki te pozwalają sprowadzić do pomiarów zmian rezystancji pomiary wielu wielkości fizycznych [1, 2, 3] jak wydłużenia, siły, ciśnienia, niektórych wielkości cieplnych oraz magnetycznych. Na zasadzie pomiaru zmian rezystancji dokonuje się również analizy składu mieszaniny gazów, mierzy się szybkość przepływu gazów, dokonuje się pomiarów wilgotności i wielu innych.

Do pomiaru zmian rezystancji opracowano mostki zwane procentowymi [4, 5, 6] lub mostkami tolerancji [7]. Postęp techniczny, objawiający się między innymi skracaniem czasu produkcji oraz zwężaniem tolerancji wykonania produktów, narzuca coraz ostrzejsze wymagania również w stosunku do metod pomiarowych. Zostały więc opracowane automatyczne metody pomiarowe o krótkim czasie działania [8, 9]. Dalsze skrócenie czasu pomiaru oraz znaczne zwiększenie dokładności umożliwiły metody numeryczne [10, 11]. Automatyczne cyfrowe mostki pomiarowe [12, 13, 14] stanowią obecnie najdokładniejsze automatyczne przyrządy do pomiaru rezystancji i jej zmian.

W pomiarach przemysłowych wystarcza jednak umiarkowana dokładność, którą można szacunkowo przyjąć jako 0,1%. Pożądany jest natomiast możliwie krótki czas pomiaru [15]. Wymaganie to dotyczy przede wszystkim pomiarów wielopunktowych w układach automatyki z centralną rejestracją i przetwarzaniem danych [16, 17].

W dotychczasowej praktyce wielkość mierzoną przetwarza się najczęściej na analogowe napięcie, a to z kolei mierzy się za pomocą woltomierza numerycznego [18, 19]. Czas pomiaru wynosi tu zazwyczaj 20 milisekund [20].

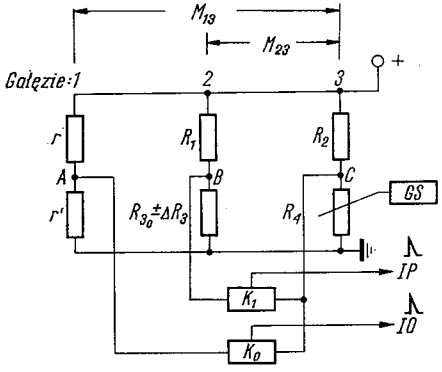
W niniejszej pracy opisano opracowaną przez autora [21, 22, 23] metodę pomiaru zmian rezystancji. Metoda ta, zwana metodą czasowego równoważenia mostka (*CRM*), polega na sterowaniu rezystora regulacyjnego w mostku pomiarowym w taki sposób, że czas potrzebny do zrównoważenia tego mostka wyraża bezpośrednio wartość mierzonej zmiany rezystancji. Czas ten mierzy się metodą numeryczną. Omawiana metoda polega więc na zastosowaniu układu rozwijającego [24] do równoważenia mostka. W odróżnieniu jednak od metod znanych, gdzie funkcją rozwijającą jest napięcie, w omawianej metodzie rolę funkcji rozwijającej spełnia rezystancja. Ta okoliczność znacznie upraszcza realizację przetwarzania zmian rezystancji na czas.

Jako rezystor regulacyjny w mostku zastosowany został tranzystor polowy, którego rezystancja ściek—źródło R_{DS} zmieniana jest liniowo w czasie.

2. ZASADA DZIAŁANIA METODY CRM

2.1. Przetwarzanie zmiany rezystancji na odcinek czasu

Istota metody [21, 22, 23] polega na przetworzeniu mierzonej zmiany rezystancji na analogowy odcinek czasu. Przetworzenie następuje w mostku rezystancyjnym przedstawionym na rys. 1.



Rys. 1. Mostek rezystancyjny równoważony metodą czasową

Układ pomiarowy składa się z gałęzi odniesienia *1*, gałęzi pomiarowej *2* oraz gałęzi regulacyjnej *3* połączonych równolegle i tworzących dwa mostki pomiarowe M_{13} oraz M_{23}, dla których regulacyjna gałąź *3* jest wspólna. Na przekątne pomiarowe tych mostków załączone są komparatory napięcia K_0 i K_1. Wartość rezystora regulacyjnego $R_4(t)$ sterowana jest za pomocą generatora sterującego (*GS*) tak, że zmienia się ona liniowo w czasie. Rezystory R_1, R_2 oraz r i r' mają wartości stałe, R_3 jest rezystorem, którego zmiany lub tolerancja w stosunku do wartości nominalnej ma być mierzona. Wartości nominalne rezystorów R_3 i R_4 oznaczone są jako R_{30} i R_{40}. Mamy więc:

$$R_3 = R_{30} + \Delta R_3,$$
$$R_4(t) = R_{40} + \Delta R_4.$$

Celem uproszczenia opisu rozpatrzymy na razie pomiar dodatnich zmian rezystancji.

Gdy $\Delta R_3 = 0$ mostek M_{23} znajduje się w równowadze, ponieważ rezystory R_1, R_2, R_{30}, R_{40} są tak dobrane, że spełnione jest równanie

$$R_1 R_{40} = R_2 R_{30}. \tag{2.1}$$

Zmiana rezystancji $\Delta R_3 \neq 0$ wyprowadza mostek z równowagi. Celem powtórnego zrównoważenia mostka należy zmienić wartość rezystora regulacyjnego R_4 do wartości $R_{40}+\Delta R_4$, przy czym ΔR_4 określona jest równaniem

$$R_1 R_{40}\left(1+\frac{\Delta R_4}{R_{40}}\right) = R_2 R_{30}\left(1+\frac{\Delta R_3}{R_{30}}\right). \tag{2.2}$$

Z równania (2.2) widać, że równowaga mostka następuje powtórnie, gdy względne zmiany rezystancji mierzonej i regulacyjnej są równe sobie, czyli gdy

$$\frac{\Delta R_3}{R_{30}} = \frac{\Delta R_4}{R_{40}}. \tag{2.3}$$

Warunek ten może być wyrażony również jako

$$\Delta R_3 = \frac{R_{30}}{R_{40}} \Delta R_4, \tag{2.4}$$

lub

$$\Delta R_3 = \frac{R_1}{R_2} \Delta R_4. \tag{2.4a}$$

Wartość rezystancji regulacyjnej R_4 zmienia się liniowo w czasie według równania

$$R_4(t) = R_{40} + S\Delta t, \tag{2.5}$$

gdzie stała $S = \frac{dR_4}{dt}$ jest stromością zmiany rezystancji $R_4(t)$ w czasie.

Zmiana ta następuje od wartości początkowej (np. minimalnej) do końcowej (maksymalnej), po czym następuje szybki powrót do wartości początkowej. Zmiana ta może odbywać się jednorazowo lub powtarzać cyklicznie.

Podstawiając $\Delta R_4 = S\Delta t$ do równania (2.4a), mierzoną zmianę rezystancji można wyrazić jako funkcję czasu

$$\Delta R_3 = \frac{R_1}{R_2} S\Delta t. \tag{2.6}$$

Na zasadzie równania (2.6) zmiana rezystancji zostaje przetworzona na proporcjonalny do niej odcinek czasowy; jest to istota metody czasowego równoważenia mostka.

Stałą

$$\frac{R_1}{R_2} \cdot S = \frac{d(\Delta R_3)}{dt} = S_p \tag{2.7}$$

nazwiemy stromością przetwarzania bezwzględnej zmiany rezystancji na czas. Wartość jej powinna być równa całkowitej potędze dziesięciu; wówczas mierzony numerycznie czas pomiaru wyraża bezpośrednio mierzoną wartość ΔR_3.

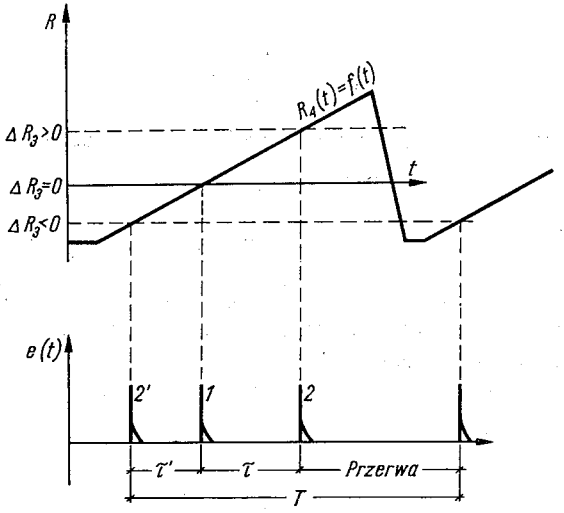
Stromość przetwarzania względnej zmiany rezystancji na czas wyraża się wzorem

$$S_p' = \frac{S_p}{R_{30}}. \tag{2.8}$$

2.2. Pomiar zmian rezystancji

W stanie początkowym wartość rezystancji $R_4(t)$ jest nieco mniejsza niż R_{40}. Dla $\Delta R_3 > 0$ cykl pomiarowy przebiega jak następuje. Rezystancja regulacyjna $R_4(t)$ zwiększa swoją wartość, osiągając po pewnym czasie wartość R_{40}. Odpowiada to równowadze mostka M_{23} dla $\Delta R_3 = 0$. Mostek M_{13} jest uregulowany tak, że przechodzi on w tym momencie przez położenie równowagi i w komparatorze K_0 powstaje impuls, który zaznacza początek czasu pomiaru t_1. Ponieważ rezystancje gałęzi 1 mostka M_{13} są stałe, impuls ten pojawia się zawsze w tym samym punkcie na osi czasu, to znaczy na początku cyklu pomiarowego niezależnie od wartości R_3, dlatego nazywać go będziemy impulsem *odniesienia* (*IO*).

Ponieważ wartość $R_4(t)$ wzrasta liniowo, mostek M_{23} przechodzi przez położenie równowagi po czasie określonym równaniem (2.6). Chwila t_2 przejścia mostka M_{23} przez położenie równowagi oznaczona jest drugim impulsem powstałym w komparatorze K_1, przy czym $t_2 - t_1 = \Delta t$. Położenie tego impulsu na osi czasu zależy od wartości ΔR_3. Impuls ten nazywamy *pomiarowym* (*IP*).



Rys. 2. Przebiegi czasowe w mostku w przypadku pomiaru zmian rezystancji
1 — impuls odniesienia (*IO*), 2 — impuls pomiarowy (*IP*), 2' — impuls pomiarowy gdy $\Delta R_3 < 0$, τ — czas pomiaru, T — okres pomiaru

W przypadku gdy $\Delta R_3 > 0$, impuls odniesienia pojawia się pierwszy, stając się impulsem START, drugi impuls pomiarowy powstaje jako impuls STOP. Dla $\Delta R_3 < 0$ role impulsów są odwrotne. Kolejność powstawania impulsów w komparatorach wykorzystana została do oznaczania znaku mierzonej zmiany oporu. Czas między impulsami jest czasem pomiaru.

Należy zauważyć, że mostki M_{13} i M_{23} nie pozostają w stanie równowagi. Mostki zbliżają się do stanu równowagi osiągają ją, po czym zostają rozrównoważone w przeciwnym kierunku.

Po osiągnięciu przez $R_4(t)$ maksymalnej wartości, która ogranicza zakres pomiaru, wartość $R_4(t)$ wraca do swojej wartości początkowej. Po przerwie cykl pomiarowy może być powtórzony.

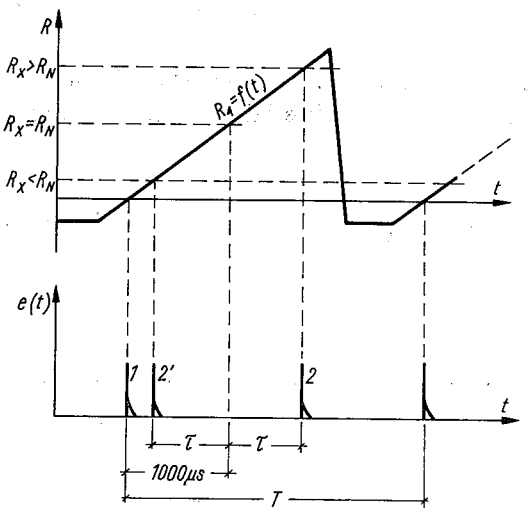
Przebiegi czasowe podczas pomiaru pokazane są na rys. 2.

2.3. Pomiar stosunku rezystancji

Wynik pomiaru może być wyrażony również jako stosunek wartości zmierzonej (aktualnej) do wartości nominalnej rezystancji badanej, tj. jako

$$\frac{R_{30}+\Delta R_3}{R_{30}} \quad \text{lub} \quad \frac{R_x}{R_N}.$$

Gdy $R_x = R_N$ (co oznacza, że $\Delta R_3 = 0$), stosunek ten równa się jedności. W polu odczytowym urządzenia pomiarowego uwidoczniona powinna zostać jedynka (zamiast zera). Można tego dokonać przesuwając impuls (IO) o 1000 μs do tyłu na osi czasu, czyli generując go o 1000 μs wcześniej przez odpowiednią regulację stosunku oporów r/r' w mostku M_{13}.



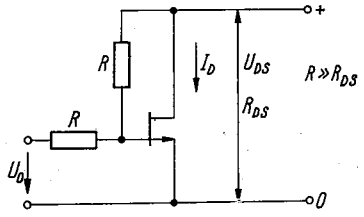
Rys. 3. Przebiegi czasowe w mostku w przypadku pomiaru stosunku rezystancji
1 — impuls odniesienia (IO), 2 — impuls pomiarowy gdy $R_x > R_N$, 2' — impuls pomiarowy gdy $R_x < R_N$, τ — czas pomiaru T — okres pomiaru

Dla $\Delta R_3 = 0$ czas jaki upłynie między pojawieniem się impulsów IO i IP wynosi 1000 μs. W polu odczytowym miernika czasu zostanie uwidoczniona liczba 1000. Jeżeli za jedynką zostanie umieszczony przecinek, to wskazanie przyrządu równe będzie 1,000, co oznacza, że stosunek aktualnej wartości rezystancji mierzonej do nominalnej równa się jedności. Przebiegi czasowe w przypadku pomiaru stosunku rezystancji przedstawione są na rys. 3.

3. REALIZACJA METODY

Istotnym dla realizacji metody podzespołem aparatury jest rezystor o liniowej zmienności w czasie. Pożądane parametry tego rezystora to: zakres zmian rezystancji około 30%, okres zmienności około 1 ms, wykonanie bez kontaktów ruchomych. Własności te zapewnić może tylko rozwiązanie czysto elektroniczne.

Celem skonstruowania takiego rezystora autor przeprowadził szczegółową analizę znanego z literatury układu [25], w którym rezystorem o zmiennej wartości jest tranzystor polowy. Wyniki tej analizy przedstawionej w pracy [26] pozwoliły na skonstruowanie rezystora spełniającego warunki metody *CRM*. Myśl przewodnia tej analizy jest następująca.



Rys. 4. Tranzystor polowy z dzielnikiem linearyzującym charakterystykę $I'_D = f(U_{DS})$

Zależność rezystancji ściek—źródło R_{DS} od napięcia U_0 sterującego bramkę dla tranzystora polowego, pracującego w układzie jak na rys. 4, można według [26, 27, 28] wyrazić wzorem:

$$R_{DS} = \frac{1}{k(2U_{p1} + U_0)} \quad \text{dla} \quad |U_0| < |2U_{p1}|, \tag{3.1}$$

gdzie:

k — stała,

U_{p1} — napięcie *pinch-off*.

Jeżeli do sterowania tranzystora polowego użyte zostanie napięcie U_0 uzależnione od U_1 według wzoru

$$U_0 = \frac{1}{CU_1} - D, \tag{3.2}$$

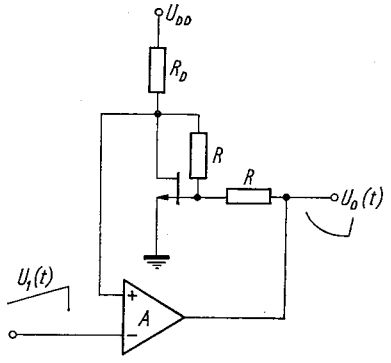
gdzie C i D są stałymi, to zależność R_{DS} od U_1 ma postać

$$R_{DS} = \frac{CU_1}{k + kCU_1(2U_{p1} - D)}. \tag{3.3}$$

Dla $D = 2U_p$ zależność ta jest liniowa

$$R_{DS} = \frac{C}{k} U_1. \tag{3.4}$$

Zależność (3.2) realizowana jest w układzie formującym, którego schemat przedstawiony jest na rys. 5.



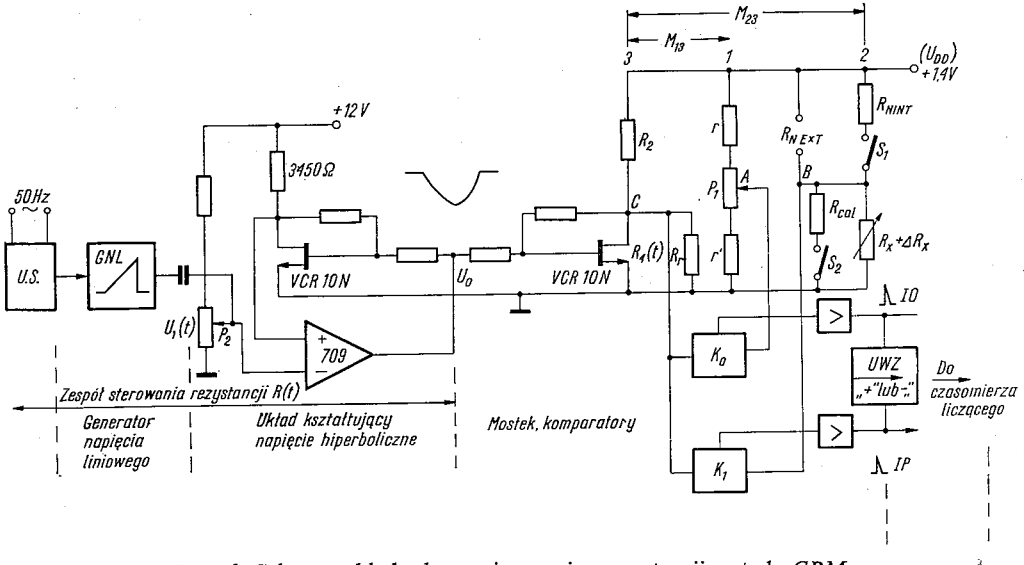
Rys. 5. Schemat układu kształtującego napięcie

Przy przyłożonym na wejściu napięciu liniowym, napięcie na wyjściu układu ma kształt zbliżony do hiperbolicznego. Zależność między U_1 i U_0 wyrażona jest funkcją, której postać uwikłaną przedstawia wzór

$$U_0 U_1 + \left(2U_{p2} + \frac{1}{k_2 R D_2}\right) U_1 - \frac{U_{DD}}{k_2 R D_2} = \frac{-U_0}{A}\left(U_0 + 2U_{p2} + \frac{1}{k_2 R D_2}\right) \qquad (3.5)$$

$$\text{dla } -2U_{p2} < U_0 < 0,$$

gdzie A — wzmocnienie wzmacniacza operacyjnego.



Rys. 6. Schemat układu do pomiaru zmian rezystancji metodą *CRM*

Kompletny schemat układu formującego oraz mostka zawierającego tranzystor polowy T_1 jako rezystor sterowany przedstawiony jest na rys. 6.

Zmiany rezystancji na zaciskach ściek—źródło tranzystora polowego T_1 wyraża wzór

$$R'_{DS} = \frac{1}{k_2\left[2(U_{p2}-U_{p1})+\frac{1}{k_2 R_r}-\frac{1}{k_1 R_{D1}}+\frac{U_{DD}}{k_1 R_{D1} U_1}\right]}. \tag{3.6}$$

Indeksy 1 odnoszą się do tranzystora polowego T_1 w układzie rezystora sterowanego, indeksy 2 do tranzystora polowego w układzie formującym.

Przy spełnieniu warunku

$$2(U_{p2}-U_{p1})+\frac{1}{k_2 R_r}-\frac{1}{k_1 R_{D1}}=0, \tag{3.7}$$

zależność $R'_{DS}=f(U_1)$ jest liniowa

$$R'_{DS}=\frac{R_D}{U_{DD}}\cdot U_1. \tag{3.8}$$

Celem spełnienia warunku (3.7) tranzystory polowe T_1 i T_2 powinny być tego samego typu o zbliżonych danych charakterystycznych.

Działanie aparatu, którego schemat przedstawiony został na rys. 6, jest następujące. Układ startowy *US* dostarcza co 20 ms impulsu wyzwalającego generator napięć liniowych *GNL*. Napięcie liniowe z tego generatora dochodzi do wejścia wzmacniacza operacyjnego 709. Wzmacniacz ten wraz z układem tranzystora polowego znajdującym się w pętli sprzężenia zwrotnego tego wzmacniacza stanowi układ formujący. Układ ten przetwarza liniowe napięcie U_1 w napięcie U_0 hiperbolicznie uzależnione od czasu. Napięcie to steruje tranzystor polowy T_1, który spełnia rolę rezystora regulacyjnego w mostku. Na przekątną mostka odniesienia M_{13} połączony jest komparator K_0, generujący w chwili równowagi tego mostka impuls odniesienia. Impuls ten po wzmocnieniu pojawia się na wyjściu aparatu jako *IO*. Na przekątnej mostka pomiarowego M_{12} połączony jest komparator K_1, który w chwili równowagi tego mostka generuje impuls pomiarowy. Impuls ten po wzmocnieniu pojawia się na wyjściu aparatu jako *IP*. Kolejność pojawienia się tych impulsów, a tym samym znak mierzonej zmiany rezystancji wykazuje układ wykazania znaku *UWZ*. Impulsy te doprowadzane są następnie do czasomierza numerycznego.

Na schemacie mostka pomiarowego (rys. 6) zaciski oznaczone $R_{N\,ext}$ oraz wyłącznik S_1 służą do załączenia zewnętrznego rezystora wzorcowego. Rezystor oznaczony R_{cal} oraz wyłącznik S_2 służą do kalibracji aparatu.

4. WYNIKI BADAŃ URZĄDZENIA PROTOTYPOWEGO

W zbudowanej aparaturze uzyskano zakres pomiaru względnych zmian rezystancji w granicach od $-5{,}0\%$ do $+20{,}0\%$ przy zdolności rozdzielczej równej 0,1%, przy czym czas pomiaru dla maksymalnej wartości wielkości mierzonej wynosi 200 mikrosekund. Przy zmianie zakresu pomiarowego, uzyskanej przez zmianę stromości wzrostu napięcia liniowego z *GNL* dla wielkości mierzonej zawartej w granicach od $-1{,}00\%$ do $+10{,}00\%$ zdolność rozdzielcza wynosiła 0,01%, a maksymalny czas pomiaru był równy jednej milisekundzie. W powyższych zakresach niestałość wyniku pomiaru nie przekraczała wartości ± 1 jednostki ostatniego miejsca dziesiętnego.

5. WNIOSKI

Na podstawie wyników badania urządzenia prototypowego wyciągnąć można następujące wnioski co do możliwości zastosowania metody.

1. W oparciu o metodę *CRM* można budować proste i wygodne w obsłudze przyrządy do pomiaru tolerancji rezystorów w zastosowaniu do kontroli technicznej lub sortowania.

2. Duża zdolność rozdzielcza aparatury przy ograniczeniu zakresu pomiarowego pozwala na zastosowanie metody w kontroli technicznej i pomiarach tolerancji rezystancyjnych przetworników energii i przetworników pomiarowych.

3. Metoda może być wykorzystana w pomiarach wielkości nieelektrycznych dokonywanych w oparciu o rezystancyjne przetworniki pomiarowe.

4. Krótki czas pomiaru oraz możliwości przedstawienia wyniku pomiaru w postaci numerycznej predestynują metodę do stosowania jej w automatycznych układach pomiarów i kontroli.

LITERATURA CYTOWANA W TEKŚCIE

1. H. Wiedmer, Technische Informationen — Messen — Steuern — Regeln, 4. Aufl. VEB Verlag Technik, Berlin 1966.
2. M. Łapiński, Czujniki pomiarowe, PWT, 1957.
3. A. M. Turiczin, Elektriczeskie Izmierienia Nielektriczeskich Wieliczin, Gosenergoizdat, 1954
4. G. Sink, Prozent-Messbrücken, ATM, Juni 1937, J-912-2.
5. W. Hunsinger, Prozent-Messbrücken mit unmittelbarer Anzeige, ATM, Juni 1939, J-912-3.
6. K. W. Karandiejew, Pomiary elektryczne metodami mostkowymi i kompensacyjnymi, WNT, W-wa 1969.
7. J. Hajek, Automatische Prozentmessbrücke zum Gebrauch in der Industrie, ATM, Okt. 1970, s. 229, J-912-5.
8. J. F. Kinnard, Applied Electrical Measurements, John Wiley a. Sons Inc., 1956.
9. K. Homilius, Selbstabgleichende Kompensations-Schaltungen in der elektrischen Messtechnik, Teil I-V, ATM, 8, 10, 11, 12 — 1964, 1 — 1965.
10. E. Weber, Digitale Messmethoden, Interkama 1957, Sektion C, S. 85, Oldenbourg, München.
11. A. Sowiński, Cyfrowa technika pomiarowa, Wyd. 2, WKŁ, W-wa 1970.
12. E. Lorenz, Ein Präzisions-Digital-Ohmmeter, Siemens Zeitschrift, 34 Ihg., 1960, H. 10, S. 733.
13. L. Borucki, E. Lorenz, Neue Messgeräte für Präzisions- und Betriebsmessungen, Zeitschr. für Instrumentenkunde, Apr. 1962, S. 81—89.
14. Solartron, Instruction Manual for the Resistor Digital Test Set Type EM 1 006.
15. K. Kleinteich, Informationsgewinnung für digitale Anlagen, msr — 3, 1968, s. 67.
16. J. Nowaczyk, Cyfrowy woltomierz całkujący jako typ woltomierza przemysłowego, Z.N.P. Wrocł. Miernictwo X, nr 200, 1968, s. 93.
17. A. Ege, Einführung in die Realzeit-Datenverarbeitung, Elektronik, 6/71, S. 205.
18. R. Staritz, Die Umwandlung elektrischer Messgrössen in Digitalwerte, Elektronik, 9/1959, S. 266.
19. B. J. Szweckij, Elektronnyje izmieritielnyje pribory s cifrowym otsczetom, Technika, Kiew 1964.
20. B. G. Kay, Selecting the Right Digital Voltmeter, Electronics, April, 4, 1966, p. 84.
21. L. J. Weiss, Sposób pomiaru zmian impedancji przy pomocy mostka elektrycznego oraz urządzenia do stosowania tego sposobu, Pat. PRL nr 55517-1964.
22. L. J. Weiss, Nowa metoda numerycznego pomiaru zmian oporu, Publikacja Ośrodka Postępu Technicznego, W-wa, kwiecień 1969.
23. L. J. Weiss, Digital Measurement of Resistance Deviations, Electr. Eng., Aug. 1972, p. 41.
24. F. E. Tiemnikow, Teoria układów rozwijających, WNT, W-wa 1965.

25. P. Seyfried, Elektronische Leistungsmessung und ihre Anwendungen, Interkama 1968, Oldenbourg, München 1968.
26. L. J. Weiss, Tranzystor polowy jako opór o wartości zmiennej liniowo w czasie, Elektronika, nr 5,1973.
27. R. D. Middlebrook, A Simple Derivation of Field-Effect Transistor Characteristics, Proc. IEEE, Aug. 1963, pp. 1146—1147.
28. Th. Molinga, The FET as a Voltage-Controlled Resistor, EEE, Jan. 1970, pp. 50—61.

L. J. WEISS

THE METHOD OF TIME-BALANCING OF A BRIDGE IN RESISTANCE DEVIATIONS MEASUREMENTS

Summary

In this article a method for measuring of resistance deviations by means of a bridge is described. Due to linearly versus time variation of the regulating resistance in the measuring bridge the measured resistance deviation is performed into an analogous time interval, which measured digitaly, may directly express the value of the said resistance deviation. As the regulating resistor in the bridge a *FET* is used.

L. J. WEISS

METHODE ZUM ZEITAUSGLEICH EINER BRÜCKE BEI MESSUNGEN VON WIDERSTANDSÄNDERUNGEN

Zusammenfassung

In diesem Aufsatz ist eine Brückenmethode zur Messung von Widerstandsänderungen beschrieben worden. Dank linearer Zeitabhängigkeit des Regulationswiderstandswertes in der Brücke wird die zu messende Widerstandsänderung in ein analoges Zeitinterval umgeformt, welches digital gemessen, die Messgrösse unmittelbar in Zahlenform darstellen kann. Als Regulationswiderstand in der Brücke wird ein Feldeffekttransistor angewandt.

Л. Я. ВАЙС

МЕТОД ВРЕМЕННОЙ БАЛАНСИРОВКИ МОСТОВОЙ СХЕМЫ В ИЗМЕРЕНИЯХ ИЗМЕНЕНИЙ СОПРОТИВЛЕНИЯ

Резюме

В статье представлен мостовой метод измерения изменений сопротивления. Этот метод заключается в том, что сопротивление одной из ветвей, являющейся регуляционным элементом в мосте, изменяет свое значение линейно во времени. Благодаря этому измеряемое изменение сопротивления будет преобразовано на аналоговый промежуток времени. Это время, измеряемое нумерически, может выражать непосредственно значение измеряемой величины.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 1, ss. 69-81

621.372.4

Analiza dynamiki układów drugiego rzędu w których tłumienie jest funkcją uchybu

KRZYSZTOF KUŹMIŃSKI (ŁÓDŹ)
Instytut Automatyki i Elektroniki Politechniki Łódzkiej

Otrzymano 28.12.1972

W pracy dokonano analizy właściwości dynamicznych układów opisanych równaniami różniczkowymi:

$$\ddot{\varepsilon}+2(a-b\varepsilon)\dot{\varepsilon}+\varepsilon = 0$$

oraz

$$\ddot{\varepsilon}+2(a-b|\varepsilon|)\dot{\varepsilon}+\varepsilon = 0.$$

Podano portrety fazowe tych układów oraz odpowiednie przebiegi czasowe przy $x_0(t) = A1(t)$.

Przedyskutowano zalety i wady rozważanej korekcji nieliniowej układów automatycznej regulacji.

Niech będzie dany liniowy układ regulacji drugiego rzędu opisany równaniem

$$\ddot{\varepsilon}(t)+2\zeta\omega\dot{\varepsilon}(t)+\omega^2\varepsilon(t) = 0, \tag{1}$$

gdzie:

ε — uchyb regulacji,

ω — częstotliwość drgań własnych,

ζ — współczynnik tłumienia.

Dla układów opisanych równaniem (1), w konwencjonalnej teorii regulacji przyjmuje się, że najlepszym ze względu na czas trwania procesu przejściowego oraz przeregulowanie jest współczynnik tłumienia $\zeta = \frac{1}{\sqrt{2}}$ [3]. Wprowadzenie jednak bardziej precyzyjnych kryteriów jakości regulacji powoduje zmiany wymagań w odniesieniu do tego współczynnika. Określenie jego wartości dla różnych kryteriów można znaleźć między innymi w pracach [1], [4]. Warto tutaj zwrócić uwagę na to, że wybór kryterium jakości, według którego oceniany jest układ, nie jest zagadnieniem jednoznacznie określonym, co wyraźnie jest podkreślane w literaturze, np. [8], [9]. Istotnym w odniesieniu do układu liniowego jest jednak to, że współczynnik tłumienia w trakcie procesu przejściowego pozostaje stały, a dobór jego może być różny w zależności od przyjętego kryterium. Zmniejszenie współczynnika tłumienia w porównaniu z wartością optymalną powoduje przyspieszenie przebiegów przejściowych przy jednoczesnym zwiększeniu przeregulowania. Natomiast zwiększenie jego wartości zwalnia przebiegi przejściowe przy jednoczesnym zmniejszeniu lub całkowitym wy-

eliminowaniu przeregulowania. To spostrzeżenie stało się podstawą syntezy układów, w których współczynnik tłumienia jest funkcją błędu regulacji $\zeta = f(\varepsilon)$. Realizację takiego układu można uzyskać przez wprowadzenie nieliniowego sprzężenia korekcyjnego. Równanie różniczkowe opisujące taki układ przyjmuje wtedy postać

$$\ddot{\varepsilon} + 2f(\varepsilon)\omega\dot{\varepsilon} + \omega^2\varepsilon = 0. \tag{2}$$

Omówienie oraz obszerną bibliografię tego rodzaju układów można znaleźć w pracach [10], [13], [14]. Badania nad tymi układami rozpoczęto stosunkowo dawno, mianowicie już w 1950 r. D. C. Mac Donald [6] rozważał układy, w których

$$f(\varepsilon) = \frac{c}{1 + a|\varepsilon|} \tag{3}$$

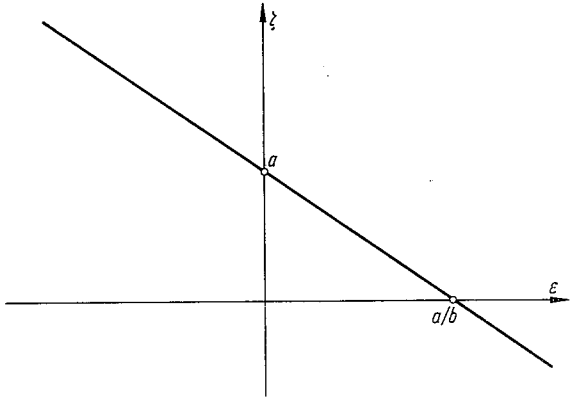
oraz

$$f(\varepsilon) = \frac{c}{1 + a|\varepsilon + b\dot{\varepsilon}|}. \tag{4}$$

Niedługo potem J. B. Lewis [5] opisał układ, w którym

$$\zeta = f(\varepsilon) = a - b\varepsilon. \tag{5}$$

Układ ten jest często przytaczany w różnych pracach, np. [7], [10], [11], [13], [14] jako klasyczny przykład korekcji nieliniowej. Jednak odczuwa się brak wyczerpującej analizy warunków pracy układu tego typu. Ponadto niektórzy autorzy (np. [7]) nie zwracają uwagi na to,

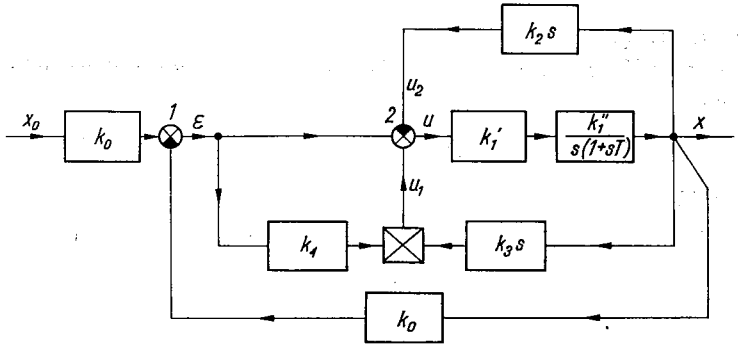


Rys. 1. Wykres zmian współczynnika tłumienia $\zeta = a - b\varepsilon$

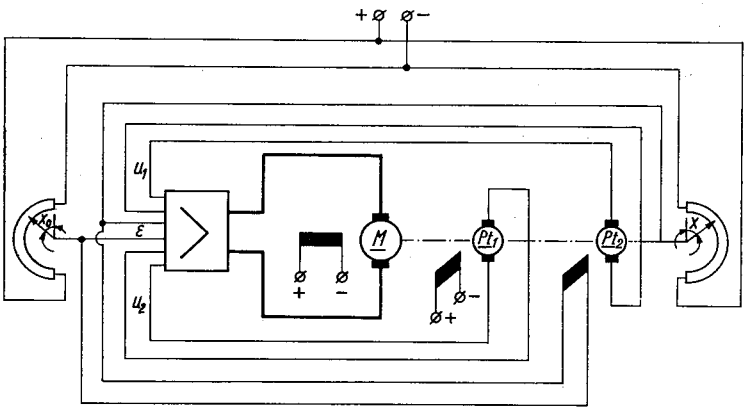
że poprawienie właściwości dynamicznych w układzie według (2) i (5) w porównaniu z układem liniowym według (1) uzyskuje się tylko dla $\varepsilon > 0$. Natomiast dla $\varepsilon < 0$ następuje pogorszenie przebiegów, co wynika ze wzrostu współczynnika tłumienia, zgodnie z zależnością (5) uwidocznioną graficznie na rys. 1.

Jest to wyraźnie podkreślone w pracach [10] i [14]. Warto jednak zająć się bliżej tym zagadnieniem ze względu na to, że układ Lewisa stanowi punkt wyjścia dla stosunkowo licznej grupy układów z nieliniowym tłumieniem.

Schemat strukturalny takiegu układu przedstawiony jest na rys. 2. Strukturze przedstawionej na tym rysunku odpowiada np. układ o schemacie ideowym jak na rys. 3. W układzie tym prądniczka tachometryczna Pt_2 poza różniczkowaniem sygnału wyjściowego x, którym jest kąt obrotu wału silnika M, dokonuje także operacji przemnożenia pochodnej



Rys. 2. Schemat strukturalny układu opisanego równaniem
$\ddot{\varepsilon}+2(a-b\varepsilon)\dot{\varepsilon}+\varepsilon = 0$, przy $x_0(t) = A1(t)$



Rys. 3. Schemat ideowy układu opisanego równaniem
$\ddot{\varepsilon}+2(a-b\varepsilon)\dot{\varepsilon}+\varepsilon = 0$, przy $x_0(t) = A1(t)$

tego sygnału przez sygnał odpowiadający uchybowi regulacji. Zatem Pt_2 pełni również rolę elementu mnożącego. Jej sygnał wyjściowy u_1, którym jest napięcie na zaciskach twornika, określa następująca zależność

$$u_1 = k_3 k_4 \varepsilon \frac{dx}{dt}. \tag{6}$$

W sumatorze 2 sygnał ten jest dodawany do sygnału błędu, od którego jednocześnie odejmuje się sygnał u_2 stanowiący napięcie prądniczki techometrycznej Pt_1

$$u_2 = k_2 \frac{dx}{dt}. \tag{7}$$

Wobec powyższego na wyjściu sumatora otrzymuje się sygnał

$$u = \varepsilon + u_1 - u_2. \tag{8}$$

Jednocześnie

$$\varepsilon = k_0(x_0 - x), \tag{9}$$

gdzie x_0 — sygnał zadający.

Przyjęto, że do wejścia układu doprowadzone jest wymuszenie standardowe w postaci funkcji skokowej

$$x_0(t) = A1(t), \tag{10}$$

gdzie A — wartość skoku danej funkcji.

Wtedy po uwzględnieniu wzorów (6), (7), (8) otrzymuje się

$$u = \varepsilon - \frac{k_3 k_4}{k_0}\varepsilon\frac{d\varepsilon}{dt} + \frac{k_2}{k_0}\frac{d\varepsilon}{dt}. \tag{11}$$

Ze wzoru tego wynika, że zależnie od położenia punktu opisującego na płaszczyźnie fazowej o współrzędnych ε i $\dot{\varepsilon}$, sygnał określony przez drugi składnik wzoru (11) współdziała z sygnałem uchybu lub mu przeciwdziała. Sygnały te współdziałają ze sobą, tzn. mają ten sam znak w III i IV kwadrancie płaszczyzny fazowej, natomiast w I i II przeciwdziałają sobie, tzn. mają różne znaki.

Zatem w rozważanym układzie występuje niepożądane zjawisko polegające na tym, że korekcja nieliniowa dla określonych warunków początkowych zamiast poprawiać będzie pogarszać dynamiczne właściwości tego układu; tak jest np. gdy $\varepsilon(0) < 0$ i $\dot{\varepsilon}(0) \geqslant 0$.

Naturalnie zjawisko to wyraźnie wpływa na przebiegi przejściowe, które aby wyznaczyć należałoby rozwiązać równanie (2) przy uwzględnieniu zależności (5). Wygodnym jest wtedy wprowadzić do równania (2) bezwymiarowy czas τ, gdzie

$$\tau = \omega t. \tag{12}$$

Otrzymuje się wówczas równanie

$$\frac{d^2\varepsilon}{d\tau^2} + 2(a - b\varepsilon)\frac{d\varepsilon}{d\tau} + \varepsilon = 0. \tag{13}$$

Dokonując następnie podstawienia

$$\frac{d\varepsilon}{d\tau} = y(\varepsilon), \tag{14}$$

skąd

$$\frac{d^2\varepsilon}{d\tau^2} = \frac{dy}{d\varepsilon}\frac{d\varepsilon}{d\tau} = y\frac{dy}{d\varepsilon}, \tag{15}$$

można więc równanie (13) przekształcić do postaci

$$y'(\varepsilon)y(\varepsilon) = 2(b\varepsilon - a)y(\varepsilon) - \varepsilon. \tag{16}$$

Równanie (16) jest równaniem Abela drugiego rodzaju [2]. Równanie to można przez kolejne podstawienie

$$y(\varepsilon) = \frac{1}{z(\varepsilon)}, \tag{17}$$

skąd

$$y'(\varepsilon) = -\frac{z'(\varepsilon)}{z^2(\varepsilon)}, \tag{18}$$

doprowadzić do postaci

$$z'(\varepsilon) = \varepsilon z^3(\varepsilon) + 2(a - b\varepsilon)z^2(\varepsilon). \tag{19}$$

Postać ta stanowi równanie Abela pierwszego rodzaju [2]. Równania Abela pierwszego i drugiego rodzaju w pewnych szczególnych przypadkach dają się rozwiązać efektywnie lub przez kwadraturę. Niestety nie dotyczy to postaci określonych zależnościami (16) i (19). W związku z tym najwygodniej w rozważanym przypadku wyznaczyć portret fazowy równania (13). Z równania tego otrzymuje się

$$\frac{dv}{d\varepsilon} = \frac{2(b\varepsilon - a)v - \varepsilon}{v}, \tag{20}$$

gdzie

$$v = \frac{d\varepsilon}{d\tau}. \tag{21}$$

Z zależności (20) wynika, że punkt osobliwy umiejscowiony jest w środku układu współrzędnych. Punkt ten zgodnie z obowiązującą klasyfikacją jest węzłem stabilnym dla $a \geqslant 1$ oraz ogniskiem stabilnym dla $0 < a < 1$.

Z kolei izokliny są rodziną hiperbol

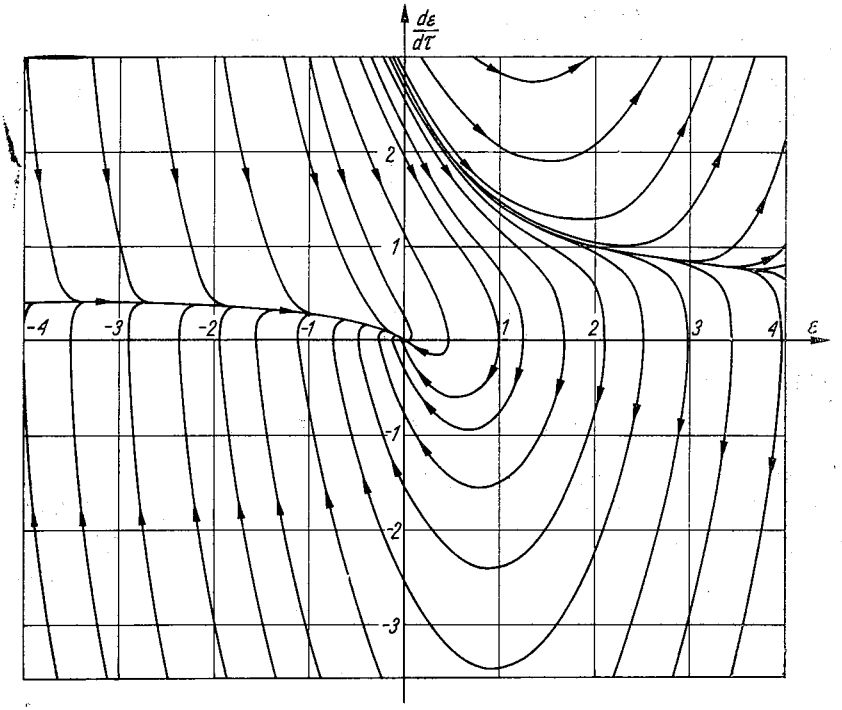
$$v = \frac{\varepsilon}{2b\varepsilon - (2a + m)}, \tag{22}$$

gdzie m — tangens kąta nachylenia trajektorii fazowych. Wyjątek stanowią: oś rzędnych, która jest izokliną odpowiadającą nachyleniu $m = -2a$ oraz asymptota pozioma o równaniu

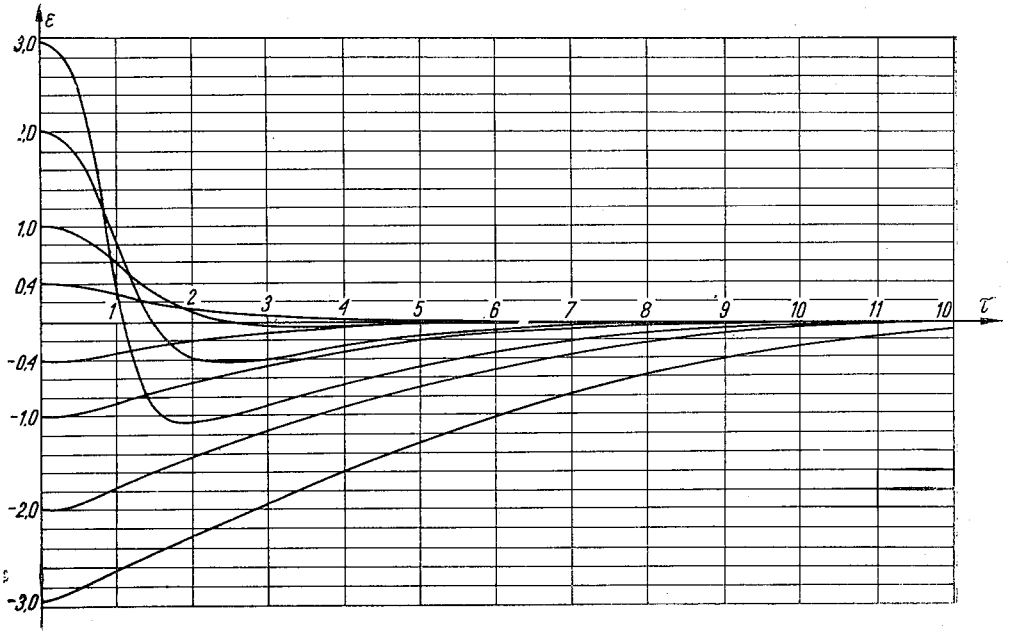
$$v = \frac{1}{2b}, \tag{23}$$

stanowiąca również izoklinę odpowiadającą temu samemu nachyleniu, tzn. $m = -2a$.

Portret fazowy rozpatrywanego układu przedstawiono na rys. 4, po przyjęciu zgodnie z wytycznymi Lewisa [5] $a = 1$ i $b = 0{,}9$ jako wartości najdogodniejszych ze względu na wymagany przebieg odpowiedzi jednostkowej, tzn. gdy we wzorze (10) $A = 1$. Zatem w tym przypadku punkt osobliwy jest węzłem stabilnym. Jednak przebieg trajektorii fazowych z prawej strony tego punktu (dla $\varepsilon > 0$) wskazuje, że posiada on również cechy stabilnego ogniska. Jednocześnie z portretu fazowego widać, że układ posiada obszar niestabilności odpowiadający prawej górnej części płaszczyzny fazowej.



Rys. 4. Portret fazowy układu opisanego równaniem
$\ddot{\varepsilon}+2(a-b\varepsilon)\dot{\varepsilon}+\varepsilon = 0$ przy $x_0(t) = A1(t)$



Rys. 5. Przebiegi czasowe $\varepsilon(t)$ w układzie opisanym równaniem
$\ddot{\varepsilon}+2(a-b\varepsilon)\dot{\varepsilon}+\varepsilon = 0$ przy $x_0(t) = A1(t)$ i $\dot{\varepsilon}(0) = 0$

Ponadto w oparciu o przedstawiony portret fazowy można stwierdzić, że empirycznie ustalone przez Lewisa [5] dane określające współczynniki a i b we wzorze (5) są w zasadzie słuszne. Poleca on przyjmować takie wartości a i b, aby spełnione były warunki: $f(0) = 0$ i $f(A) = 1$. Jednak przy takim doborze a i b dla $A = 1$, występuje już pewne bardzo małe przeregulowanie przy skoku jednostkowym, gdy $\varepsilon(0) > 0$. Widać to na portrecie fazowym oraz na rys. 5, gdzie przedstawiono przebiegi czasowe uzyskane w oparciu o ten portret. Pokazano tam przebiegi przy $\dot{\varepsilon}(0) = 0$, $\varepsilon(0) > 0$ oraz $\varepsilon(0) < 0$. Na rys. 4 i 5 widać, że przy $\varepsilon(0) > 0$, gdy $A \geqslant 1$ (a i b dobrano dla $A = 1$) występuje pojedyncze przeregulowanie. Natomiast gdy $A < 1$ przebieg jest monotoniczny. W przypadku, gdy $\varepsilon(0) < 0$ wszystkie przebiegi są monotoniczne, niezależnie od wartości $|A|$, przy czym czas ustalenia jest wtedy bardzo długi.

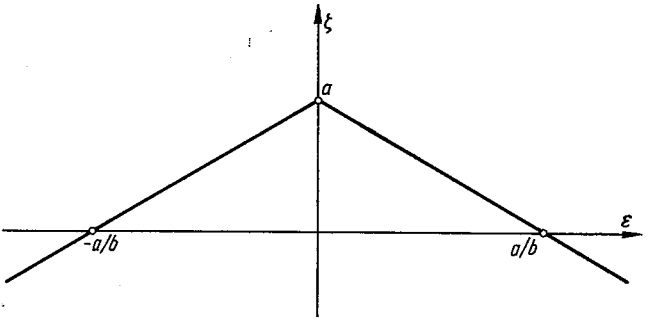
Zasadniczą wadą rozpatrywanego dotychczas układu, polegającą na różnych właściwościach dynamicznych w zależności od tego czy $\varepsilon(0) > 0$, czy też $\varepsilon(0) < 0$, można wyeliminować przechodząc do układu opisanego następującym równaniem różniczkowym

$$\frac{d^2\varepsilon}{d\tau^2} + 2(a - b|\varepsilon|)\frac{d\varepsilon}{d\tau} + \varepsilon = 0. \tag{24}$$

W tym przypadku

$$\zeta = f(\varepsilon) = a - b|\varepsilon|, \tag{25}$$

co pokazano na wykresie rys. 6.



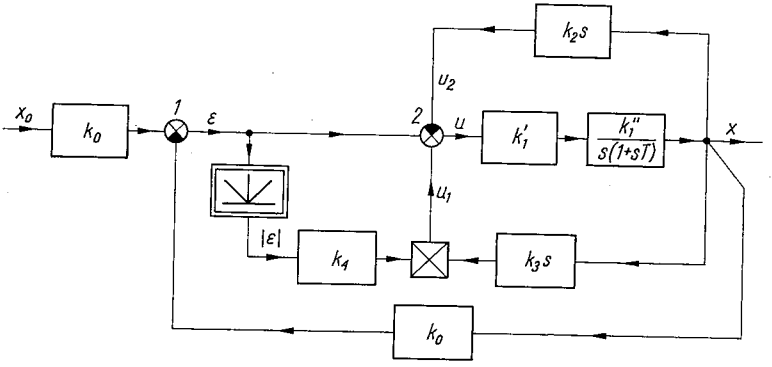
Rys. 6. Wykres zmian współczynnika tłumienia $\zeta = a - b|\varepsilon|$

Propozycję takiego układu można znaleźć między innymi w pracach [10], [11], [13], [14]. W celu uzyskania układu opisanego równaniem (24) trzeba schemat strukturalny z rys. 2 zmodyfikować wprowadzając element dający na wyjściu sygnał równy bezwzględnej wartości uchybu regulacji, jak pokazano na rys. 7. Z kolei realizacja takiego układu będzie się różniła od układu przedstawionego na rys. 3 tym, że uzwojenie wzbudzające prądniczki Pt_2 musi być zasilane napięciem odpowiadającym uchybowi regulacji poprzez mostek Graetza. Pokazano to na rys. 8.

W przypadku takiegu układu napięcie u spełnia zależność

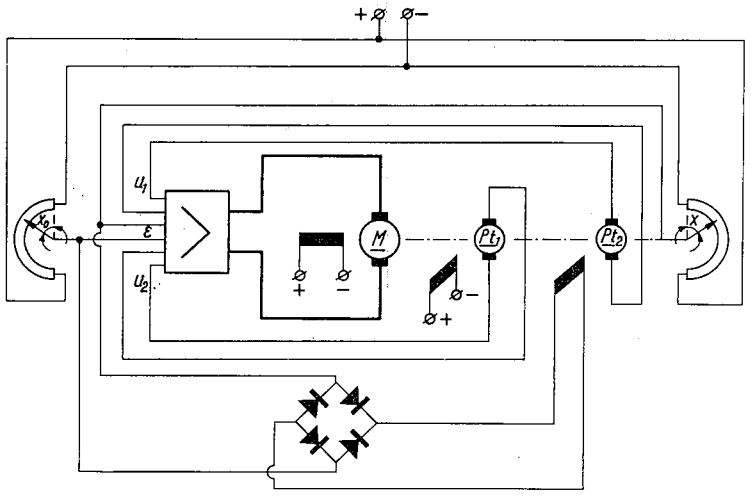
$$u = \varepsilon - \frac{k_3 k_4}{k_0}|\varepsilon|\frac{d\varepsilon}{dt} + \frac{k_2}{k_0}\frac{d\varepsilon}{dt}. \tag{26}$$

Wtedy sygnał korekcji nieliniowej współdziała z sygnałem uchybu regulacji w II i IV kwadrancie płaszczyzny fazowej. Natomiast w I i III kwadrancie przeciwdziała sygnałowi uchybu.



Rys. 7. Schemat strukturalny układu opisanego równaniem
$\ddot{\varepsilon}+2(a-b|\varepsilon|)\dot{\varepsilon}+\varepsilon=0$, przy $x_0(t)=A1(t)$

A zatem w tym układzie występują również obszary, w których korekcja nieliniowa pogarsza właściwości dynamiczne układu. Jednak w tym przypadku występuje symetria tych obszarów względem środka układu współrzędnych płaszczyzny fazowej, co zapewnia analogiczne działanie układu dla $\varepsilon > 0$ i $\varepsilon < 0$.



Rys. 8. Schemat ideowy układu opisanego równaniem
$\ddot{\varepsilon}+2(a-b|\varepsilon|)\dot{\varepsilon}+\varepsilon=0$, przy $x_0(t)=A1(t)$

Dla tego układu izokliny stanowią rodzinę krzywych opisanych równaniem

$$v=\frac{\varepsilon}{2b|\varepsilon|-(2a+m)}. \tag{27}$$

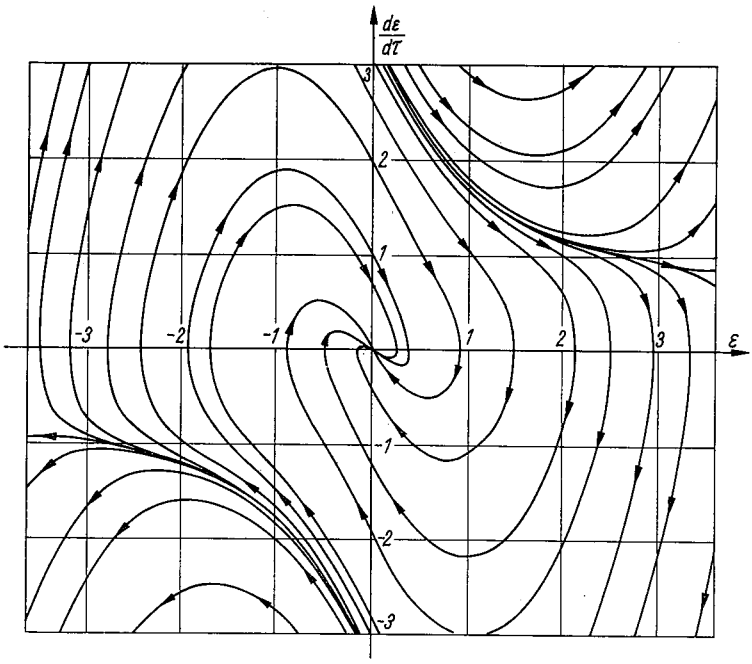
Wyjątkiem jest tutaj oś rzędnych, będąca izokliną odpowiadającą nachyleniu $m = -2a$ oraz asymptoty poziome o równaniach:

$$v = \frac{1}{2b} \quad \text{dla} \quad \varepsilon > 0, \tag{28}$$

$$v = -\frac{1}{2b} \quad \text{dla} \quad \varepsilon < 0, \tag{29}$$

będące również izoklinami o nachyleniu $m = -2a$.

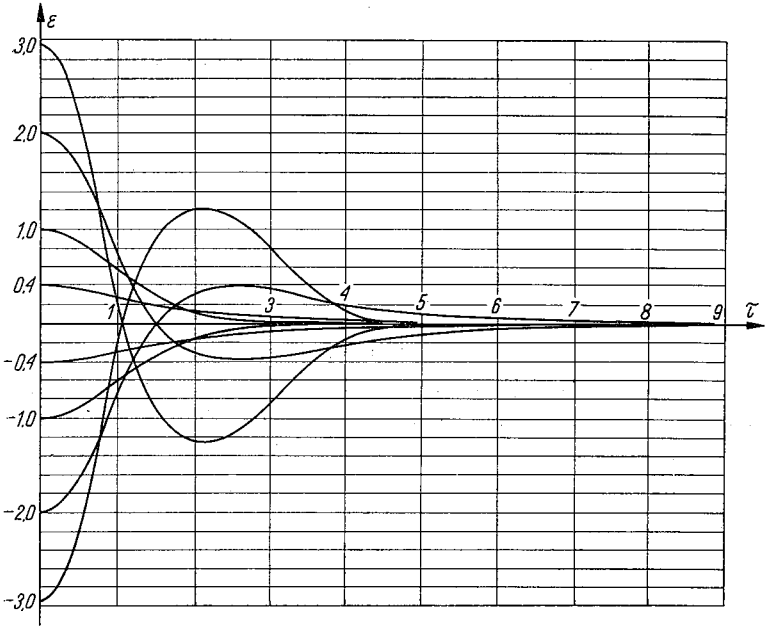
Portret fazowy tego układu przedstawiono na rys. 9, przyjmując takie same wartości a i b jak poprzednio. Z portretu tego widać, że środek układu współrzędnych ma charakter węzła stabilnego, a przebiegi są oscylacyjne przy $|A| > 1$ (a i b dobrano dla $A = 1$). Jest



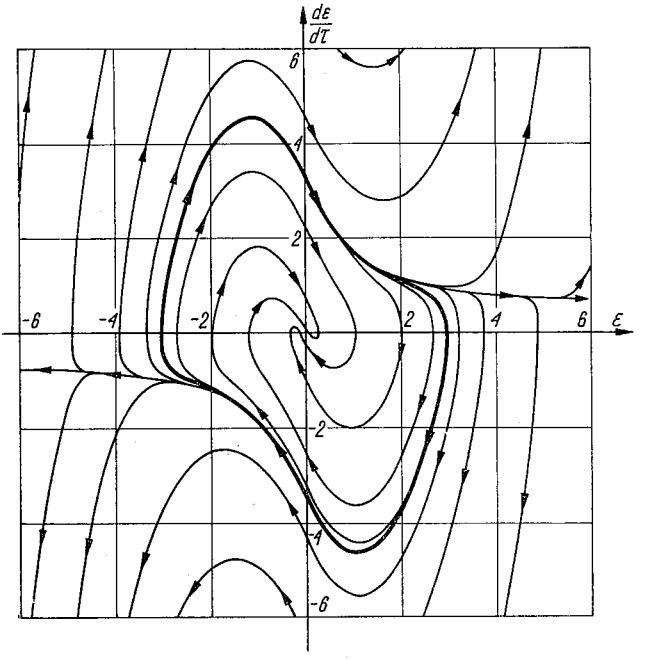
Rys. 9. Portret fazowy układu opisanego równaniem
$\ddot{\varepsilon} + 2(a - b|\varepsilon|)\dot{\varepsilon} + \varepsilon = 0$ przy $x_0(t) = A1(t)$

to ewidentne, jeżeli weźmie się pod uwagę zmniejszenie współczynnika tłumienia określonego wzorem (25) z chwilą wystąpienia przeregulowania, tzn. z chwilą zmiany znaku ε. Natomiast gdy $|A| < 1$ przebiegi są monotoniczne. W tym przypadku nie występuje przeregulowanie i badany układ zarówno dla $\varepsilon > 0$, jak i dla $\varepsilon < 0$ działa jak układ opisany przez równanie (13), gdy $\varepsilon > 0$ i $A < 1$. Odpowiednie przebiegi czasowe dla $\dot{\varepsilon}(0) = 0$ i $\varepsilon(0) > 0$ oraz $\varepsilon(0) < 0$ przedstawiono na rys. 10.

Analizując równanie (24) można przypuszczać, że w rozpatrywanym układzie występuje cykl graniczny. Jeżeli bowiem w równaniu (24) zastąpi się τ przez $-\tau$, to trajektorie pozo-



Rys. 10. Przebiegi czasowe $\varepsilon(t)$ w układzie opisanym równaniem $\ddot{\varepsilon}(t)+2(a-b|\varepsilon|)\dot{\varepsilon}+\varepsilon = 0$ przy $x_0(t) = A1(t)$ i $\dot{\varepsilon}(0) = 0$



Rys. 11. Portret fazowy układu opisanego równaniem $\ddot{\varepsilon}+2(a-b|\varepsilon|)\dot{\varepsilon}+\varepsilon = 0$ przy $x_0(t) = A1(t)$, uwidaczniający istnienie cyklu granicznego

staną takie same, lecz będą przebiegać w odwrotnym kierunku, a równanie to przyjmie wtedy postać

$$\frac{d^2\varepsilon}{d\tau^2} - 2(a - b|\varepsilon|)\frac{d\varepsilon}{d\tau} + \varepsilon = 0. \tag{30}$$

Równanie (30), przy założeniu, że $|\varepsilon|$ stanowi zgrubną aproksymację ε^2, można w przybliżeniu traktować jako równanie Van der Pola, o którym wiadomo, że posiada cykl graniczny. Cykl graniczny w przypadku równania Van der Pola jest stabilny. Natomiast dla równania (30) powinien on być niestabilny, ponieważ równanie (30) otrzymano z (24) przez zamianę τ na $-\tau$, a więc zmianę kierunku przebiegu trajektorii. Przedstawioną hipotezę potwierdziły badania rozpatrywanego układu, przeprowadzone przy użyciu maszyny analogowej. Wyznaczony w ten sposób portret fazowy układu przedstawiono na rys. 11.

Jak widać z tego portretu obszar stabilności układu jest ograniczony niestabilnym cyklem granicznym. Ta właściwość układu nie jest podawana w pozycjach literatury dotyczących rozważanego układu.

Jak wynika z przeprowadzonych rozważań, układ opisany równaniem (13) jest w zasadzie bezużyteczny z punktu widzenia zastosowań praktycznych. Natomiast układ określony przez równanie (24) stwarza możliwości zastosowań praktycznych przy ograniczeniu warunków początkowych ze względu na stabilność, przy czym pożądane przebiegi wystąpią w nim tylko wtedy, gdy A będzie mniejsze od wartości, dla której dobrano a i b. Niemniej omówione układy, ze względu na ich właściwości, niektórzy autorzy [12] zaliczają do zamkniętych układów samonastrajających się, a więc układów adaptacyjnych.

Dalsze poprawienie właściwości dynamicznych rozpatrywanych układów można uzyskać uzależniając współczynnik tłumienia nie tylko od błędu regulacji, ale również i od pochodnej błędu regulacji, tzn. realizując $\zeta = f(\varepsilon, \dot{\varepsilon})$.

LITERATURA CYTOWANA W TEKŚCIE

1. R. E. Kalman, When is a Linear Control System Optimal, Trans. ASME, ser. D, Vol. 8, 1964, p. 51.
2. E. Kamke, Differentialgleichungen, Leipzig, 1959.
3. A. Krasowski, G. Pospiełow, Podstawy automatyki i cybernetyki technicznej, WNT, Warszawa 1965.
4. K. Kuźmiński, Optymalny współczynnik tłumienia w sensie pewnego kwadratowego kryterium jakości, Zeszyty Naukowe Politechniki Łódzkiej „Elektryka", z. 35, 1971, s. 183—192.
5. J. B. Lewis, The Use of Nonlinear Feedback to Improve the Transient Response of a Servomechanism, Trans. AJEE, Vol. 71, part II, 1952, p. 449—453.
6. D. C. Mac Donald, Nonlinear Techniques for Improving Servo Performance, Proc. Nat. Electron Conf., Chicago, Vol. 6, 1950, p. 400—421.
7. G. J. Thaler, M. P. Pastel, Nieliniowe układy automatycznego sterowania, WNT, Warszawa 1965.
8. W. A. Besekerskij, E. P. Popow, Teorija sistem awtomatičeskogo regulirowanija, „Nauka", Moskwa 1972.
9. A. A. Feldbaum, A. G. Butkowskij, Metody teorii awtomatičeskogo uprawlenija, „Nauka", Moskwa 1971.

10. Nelinejnye korrektirujuščie ustrojstwa w sistemach awtomaticeskogo uprawlenija, pod red. Ju. I. Topčewa, „Mašinostroenie", Moskwa 1971.
11. G. M. Ostrowskij, Primenenie nelinejnych korrektirujuščich ustrojstw w sistemach awtomatičeskogo regulirowanija wtorogo porjadka, Awtomatika i Telemechanika, t. XVII, no. 11, 1956, s. 979—984.
12. W. W. Solodownikow, L. S. Šramko, Rasčet i proektirowanie analitičeskich samonastraiwajuščichsja sistem s étalonnymi modeljami, „Mašinostroenie", Moskwa 1972.
13. W. A. Taran, Primenenie nelinejnoj korrekcii i peremennoj struktury dlja ulučšenija dinamičeskich swojstw sistem awtomatičeskogo regulirowanija, Awtomatika i Telemechanika, t. XXV, no. 1, 1964 s. 140—149.
14. E. I. Chlypalo, Nelinejnye sistemy awtomatičeskogo regulirowanija, „Énergija", Leningrad 1967.

K. KUŹMIŃSKI

ANALYSIS OF DYNAMIC PROPERTIES OF SECOND ORDER SYSTEMS, IN WHICH ATTENUATION IS A FUNCTION OF ERROR

Summary

The work contains the analysis of dynamic properties of systems described by the following differential equations:

$$\ddot{\varepsilon}+2(a-b\varepsilon)\dot{\varepsilon}+\varepsilon = 0$$

and

$$\ddot{\varepsilon}+2(a-b|\varepsilon|)\dot{\varepsilon}+\varepsilon = 0.$$

The phase pictures of these systems and corresponding time runs have been presented at $x_0(t) = A1(t)$.

The advantages and defects of the considered nonlinear correction of automatic regulation systems have been discussed.

K. KUŹMIŃSKI

ANALYSE DER DYNAMIK VON SYSTEMEN 2. ORDNUNG, WORIN DÄMPFUNG FUNKTION DER ABWEICHUNG IST

Zusammenfassung

Im Aufsatz wurden dynamischen Eigenschaften der Systeme analysiert, die mit folgenden Differentialgleichungen

$$\ddot{\varepsilon}+2(a-b\varepsilon)\dot{\varepsilon}+\varepsilon = 0$$

und

$$\ddot{\varepsilon}+2(a-b|\varepsilon|)\dot{\varepsilon}+\varepsilon = 0$$

beschrieben werden.

Es wurden Phasenporträten dieser Systeme und entsprechende Zeitverläufe bei $x_0(t) = A1(t)$ angegeben.

Vorteile und Nachteile der betrachteten, nichtlinearen Korrektionen automatischer Regelungssysteme wurden diskutiert.

К. КУЗЬМИНЬСКИ

АНАЛИЗ ДИНАМИЧЕСКИХ СВОЙСТВ СИСТЕМ ВТОРОГО ПОРЯДКА С ЗАТУХАНИЕМ, ЯВЛЯЮЩИМСЯ ФУНКЦИЕЙ ПОГРЕШНОСТИ

Р е з ю м е

В статье приведен анализ динамических свойств систем описанных дифференциальными уравнениями:

$$\ddot{\varepsilon}+2(a-b\varepsilon)\dot{\varepsilon}+\varepsilon=0$$

и

$$\ddot{\varepsilon}+2(a-b|\varepsilon|)\dot{\varepsilon}+\varepsilon=0.$$

Даны фазовые портреты этих систем и соответствующие переходные процессы для $x_0(t) = A1(t)$.

Проанализированы недостатки и преимущества описываемой нелинейной коррекции систем автоматического регулирования.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 1, ss. 83-93

621.372.4

Analiza dynamiki układów drugiego rzędu w których tłumienie jest funkcją uchybu i jego pochodnej

KRZYSZTOF KUŹMIŃSKI (ŁÓDŹ)

Instytut Automatyki i Elektroniki Politechniki Łódzkiej

Otrzymano 28.12.1972

W pracy dokonano analizy właściwości dynamicznych układów opisanych równaniami różniczkowymi:

$$\ddot{\varepsilon}+2(a-b|\varepsilon|)\dot{\varepsilon}+\varepsilon=0 \quad \text{dla} \quad \varepsilon\dot{\varepsilon}\leqslant 0,$$
$$\ddot{\varepsilon}+2a\dot{\varepsilon}+\varepsilon=0 \quad \text{dla} \quad \varepsilon\dot{\varepsilon}\geqslant 0$$

oraz

$$\ddot{\varepsilon}+2(a+b\varepsilon \operatorname{sign}\dot{\varepsilon})\dot{\varepsilon}+\varepsilon=0.$$

Podano wykresy zmian współczynnika tłumienia, portrety fazowe tych układów oraz odpowiednie przebiegi czasowe przy $x_0(t)=A1(t)$.

Przedyskutowano wpływ rozważanej korekcji nieliniowej na pracę układu automatycznej regulacji oraz wskazano możliwości wykorzystania w syntezie układów optymalnych.

W pracy [3] autor przeanalizował właściwości dynamiczne układów, w których współczynnik tłumienia ζ zależał od błędu regulacji, czyli $\zeta=f(\varepsilon)$. Jak zostało tam wykazane, tego rodzaju uzależnienie, uzyskane przez korekcję nieliniową, poprawia — w określonym obszarze warunków początkowych — właściwości dynamiczne układu w porównaniu ze zwykłym liniowym układem drugiego rzędu. Jednocześnie jednak na płaszczyźnie fazowej $F(\varepsilon, \dot{\varepsilon})$ występują wtedy obszary, w których korekcja nieliniowa pogarsza te właściwości, a nawet może spowodować niestabilność układu.

W związku z powyższym celowym staje się uzależnienie współczynnika tłumienia nie tylko od błędu regulacji, ale i od pochodnej tego błędu, tzn. $\zeta=f(\varepsilon, \dot{\varepsilon})$. Ogólne równanie różniczkowe opisujące wtedy układ, można przedstawić w postaci

$$\ddot{\varepsilon}+2f(\varepsilon, \dot{\varepsilon})\omega\dot{\varepsilon}+\omega^2\varepsilon=0, \tag{1}$$

gdzie ω — częstotliwość drgań własnych, lub po wprowadzeniu czasu bezwymiarowego $\tau=\omega t$

$$\frac{d^2\varepsilon}{d\tau^2}+2f\left(\varepsilon, \frac{d\varepsilon}{d\tau}\right)\frac{d\varepsilon}{d\tau}+\varepsilon=0. \tag{2}$$

Propozycje takiego układu podano w pracach [4] i [7]. Autorzy tych prac opisują tam układ ze zmienną strukturą, który można opisać następującymi równaniami:

$$\frac{d^2\varepsilon}{d\tau^2}+2(a-b|\varepsilon|)\frac{d\varepsilon}{d\tau}+\varepsilon=0 \quad \text{dla} \quad \varepsilon\dot{\varepsilon}\leqslant 0, \tag{3}$$

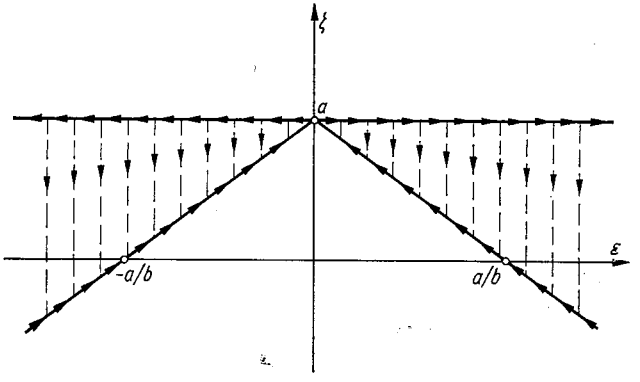
$$\frac{d^2\varepsilon}{d\tau^2}+2a\frac{d\varepsilon}{d\tau}+\varepsilon=0 \quad \text{dla} \quad \varepsilon\dot{\varepsilon}\geqslant 0. \tag{4}$$

Jak stąd wynika, w tym przypadku współczynnik tłumienia określony jest zależnościami

$$\zeta=a-b|\varepsilon| \quad \text{dla} \quad \varepsilon\dot{\varepsilon}\leqslant 0, \tag{5}$$

$$\zeta=a \quad \text{dla} \quad \varepsilon\dot{\varepsilon}\geqslant 0. \tag{6}$$

Przedstawiono to graficznie na rys. 1. Tutaj korzystniejsze właściwości dynamiczne w porównaniu z układem, dla którego obowiązuje równanie (5) w całej płaszczyźnie $F(\varepsilon, \dot{\varepsilon})$, wynikają stąd, że w układzie według (4), (5) i (6) przy oddalaniu się punktu opisującego na



Rys. 1. Wykres zmian współczynnika tłumienia określonych zależnościami $\zeta = a-b|\varepsilon|$ dla $\varepsilon\dot{\varepsilon}\leqslant 0$ i $\zeta = a$ dla $\varepsilon\dot{\varepsilon}\geqslant 0$

płaszczyźnie fazowej od położenia równowagi tłumienie jest większe, mianowicie $\zeta = a$. Następnie zmniejsza się ono skokowo do wartości określonej przez zależność (5), przy wartości ε odpowiadającej przeregulowaniu, które wystąpiło przy $\dot{\varepsilon} = 0$. Wtedy jednak punkt opisujący podąża w kierunku położenia równowagi. Powyższe zmiany współczynnika tłumienia (nie podawane w literaturze) tłumaczą właściwości rozpatrywanego układu na tle układów przeanalizowanych w pracy [3]. Dla pełności rozważań wydaje się celowym podanie portretu fazowego tego układu, który wygodnie jest wyznaczyć metodą izoklin. Zgodnie z równaniem (4) w I i III kwadrancie płaszczyzny fazowej $F(\varepsilon, \dot{\varepsilon})$ izokliny stanowią proste o równaniu:

$$v=-\frac{\varepsilon}{m+2a}, \tag{7}$$

gdzie:

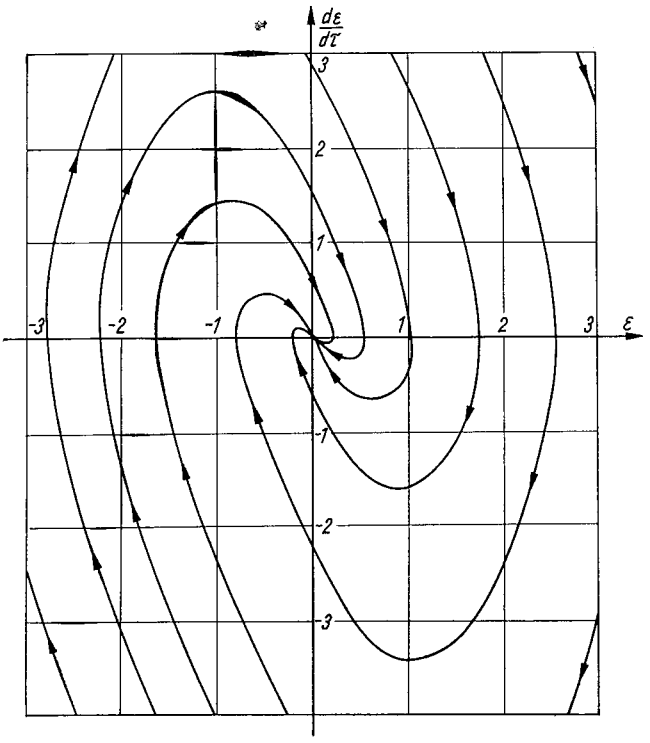
$$v=\frac{d\varepsilon}{d\tau}, \tag{8}$$

m — tangens kąta nachylenia trajektorii fazowych.

Natomiast w II i IV kwadrancie izokliny, zgodnie z równaniem (3), stanowią rodzinę krzywych o równaniu

$$v = \frac{\varepsilon}{2b|\varepsilon| - (2a+m)}. \tag{9}$$

Ponadto oś rzędnych jest izokliną odpowiadającą nachyleniu $m = -2a$.

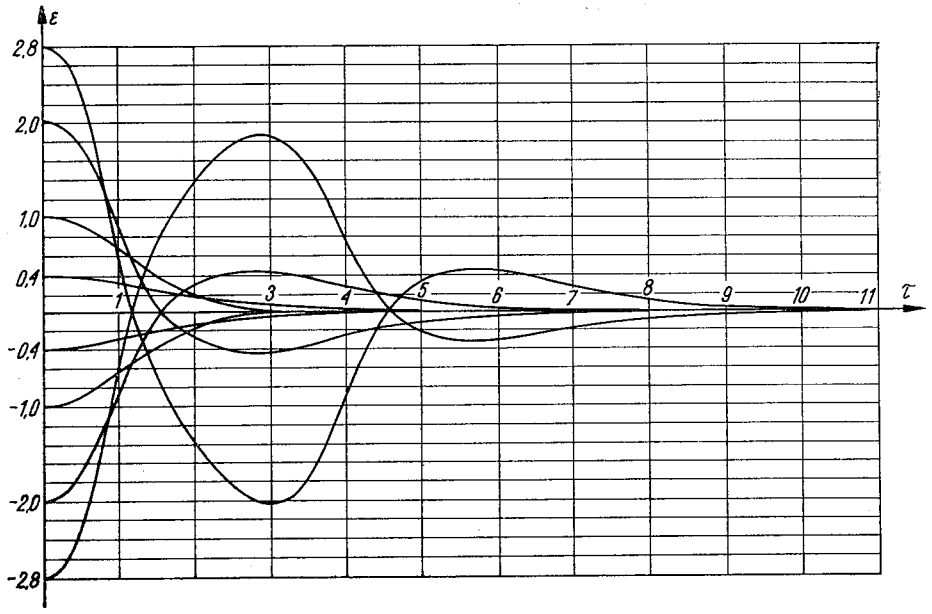


Rys. 2. Portret fazowy układu opisanego równaniami: $\ddot{\varepsilon}+2(a-b|\varepsilon|)\dot{\varepsilon}+\varepsilon = 0$ dla $\varepsilon\dot{\varepsilon} \leqslant 0$ i $\ddot{\varepsilon}+2a\dot{\varepsilon}+\varepsilon = 0$ dla $\varepsilon\dot{\varepsilon} \geqslant 0$, przy $x_0(t) = A1(t)$

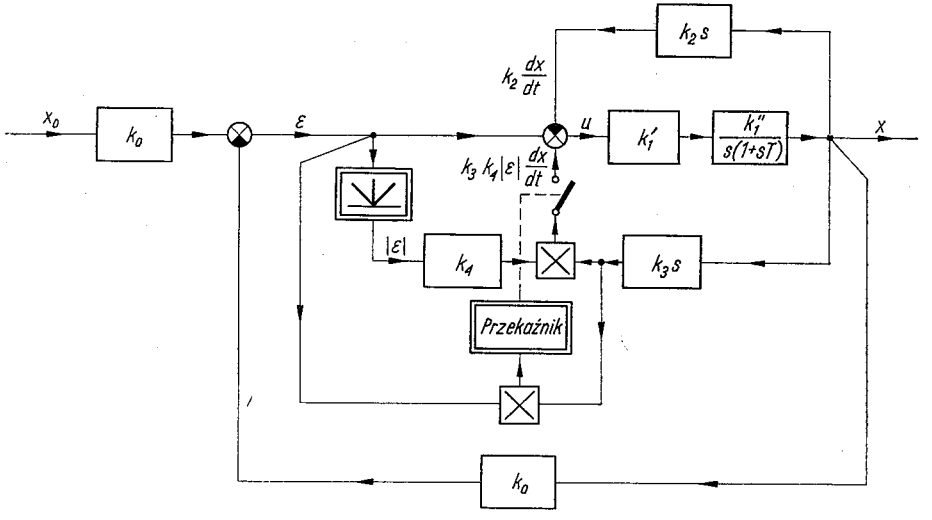
Portret fazowy rozpatrywanego układu przedstawiono na rys. 2. Przyjęto przy tym $a = 1$ i $b = 0{,}9$ jako wartości najdogodniejsze ze względu na wymagany przebieg odpowiedzi jednostkowej [3], tzn. gdy $A = 1$ przy sygnale zadającym:

$$x_0(t) = A1(t). \tag{10}$$

Jak widać na rys. 2 punkt osobliwy, umiejscowiony w środku układu współrzędnych, ma w tym przypadku charakter węzła stabilnego. Jednocześnie na podstawie tego portretu można przypuszczać, że układ jest globalnie asymptotycznie stabilny. Ponadto z portretu fazowego widać, że przy $|A| \leqslant 1$ przebiegi czasowe są monotoniczne, podobnie jak w układzie opisanym równaniem (3) w całej płaszczyźnie fazowej. Natomiast przy $|A| > 1$ przebiegi te stają się oscylacyjne, przy czym liczba przeregulowań jest tym większa im większa jest wartość $|A|$. Odpowiednie przebiegi czasowe przedstawiono na rys. 3 dla $\dot{\varepsilon}(0) = 0$ i $\varepsilon(0) > 0$ oraz $\varepsilon(0) < 0$. Przebiegi te uwidaczniają jednakowe własności układu dla $\varepsilon > 0$



Rys. 3. Przebiegi czasowe $\varepsilon(t)$ w układzie opisanym równaniami: $\ddot{\varepsilon}+2(a-b|\varepsilon|)\dot{\varepsilon}+\varepsilon=0$ dla $\varepsilon\dot{\varepsilon}\leqslant 0$ i $\ddot{\varepsilon}+2a\dot{\varepsilon}+\varepsilon=0$ dla $\varepsilon\dot{\varepsilon}\geqslant 0$, przy $x_0(t)=A1(t)$ i $\dot{\varepsilon}(0)=0$



Rys. 4. Schemat strukturalny układu opisanego równaniem $\ddot{\varepsilon}+2(a-b|\varepsilon|)\dot{\varepsilon}+\varepsilon=0$ dla $\varepsilon\dot{\varepsilon}\leqslant 0$ i $\ddot{\varepsilon}+2a\dot{\varepsilon}+\varepsilon=0$ dla $\varepsilon\varepsilon\geqslant 0$, przy $x_0(t)=A1(t)$

i $\varepsilon < 0$, o czym można już było sądzić na podstawie symetrii wykresu przedstawionego na rys. 1. Porównanie tych przebiegów z przebiegami układu opisanego równaniem (3) w całej płaszczyźnie fazowej [3], pozwala stwierdzić zmniejszenie wartości przeregulowania przy $|A| > 1$, gdy a i b były dobrane dla $A = 1$, przy czym czasy ustalenia pozostają w zasadzie takie same. Układ o opisanych własnościach, tzn. określony przez równania (3) i (4) można zrealizować według schematu strukturalnego przedstawionego na rys. 4. Ze schematu tego oraz równań (3) i (4) wynika, że sygnał sterujący u

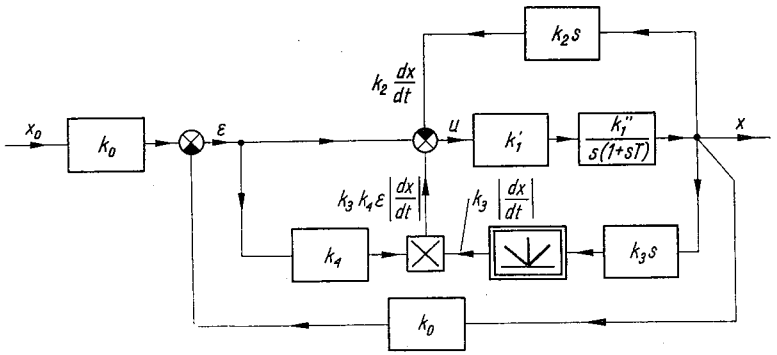
$$u = \varepsilon + k_3 k_4 |\varepsilon| \frac{dx}{dt} - k_2 \frac{dx}{dt} \quad \text{dla } \varepsilon\dot{\varepsilon} \leqslant 0 \tag{11}$$

i

$$u = \varepsilon - k_2 \frac{dx}{dt} \quad \text{dla } \varepsilon\dot{\varepsilon} \geqslant 0. \tag{12}$$

Jednocześnie

$$\varepsilon = k_0(x_0 - x), \tag{13}$$



Rys. 5. Schemat strukturalny układu opisanego równaniem
$\ddot{\varepsilon} + 2a\dot{\varepsilon} + \varepsilon + 2b|\dot{\varepsilon}| = 0$, przy $x_0(t) = A1(t)$

co przy rozważaniach dotyczących reakcji układu na wymuszenie standardowe, w postaci funkcji skokowej określonej przez wzór (10), prowadzi do następujących zależności:

$$u = \varepsilon - \frac{k_3 k_4}{k_0} |\varepsilon| \frac{d\varepsilon}{dt} + \frac{k_2}{k_0} \frac{d\varepsilon}{dt} \quad \text{dla} \quad \varepsilon\dot{\varepsilon} \leqslant 0 \tag{14}$$

i

$$u = \varepsilon + \frac{k_2}{k_0} \frac{d\varepsilon}{dt} \quad \text{dla} \quad \varepsilon\dot{\varepsilon} \geqslant 0. \tag{15}$$

Jak wynika ze wzorów (14) i (15) w II i IV kwadrancie płaszczyzny fazowej $F(\varepsilon, \dot{\varepsilon})$, sygnał korekcji nieliniowej odpowiadający drugiemu składnikowi wyrażenia (14) współdziała z sygnałem błędu regulacji, tzn. mają one jednakowe znaki. Natomiast w I i III kwadrancie sygnał korekcji nieliniowej jest równy zeru, podczas gdy w układzie opisanym równaniem (3) w całej płaszczyźnie fazowej sygnał ten przeciwdziałał sygnałowi uchybu. W wyni-

ku tego rozpatrywany układ ma lepsze właściwości dynamiczne, jak to było stwierdzone już poprzednio.

Analizowany układ posiada stosunkowo rozbudowany schemat strukturalny. Można jednak uzyskać podobne właściwości dynamiczne w układzie o bardziej prostym schemacie i mniejszej liczbie elementów w części realizującej korekcję nieliniową. Schemat takiego układu, proponowanego w pracy [7], przedstawiono na rys. 5. Schemat ten, przy założeniu, że do wejścia układu doprowadzony jest sygnał określony wzorem (10), wynika z równania

$$\ddot{\varepsilon}+2a\dot{\varepsilon}+\varepsilon+2b\varepsilon|\dot{\varepsilon}| = 0. \tag{16}$$

Okazuje się, że przekształcając równanie (16) można określić funkcję zmian współczynnika tłumienia, tzn. $\zeta = f(\varepsilon, \dot{\varepsilon})$. W tym celu należy skorzystać z zależności:

$$y \operatorname{sign} y = |y|, \tag{17}$$

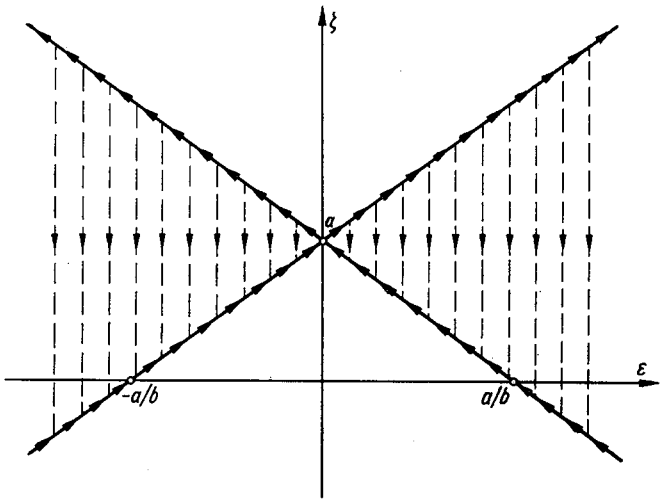
po uwzględnieniu której w równaniu (16) otrzymuje się

$$\ddot{\varepsilon}+2(a+b\varepsilon \operatorname{sign} \dot{\varepsilon})\dot{\varepsilon}+\varepsilon = 0. \tag{18}$$

Zatem wynika stąd

$$\zeta = f(\varepsilon, \dot{\varepsilon}) = a+b\varepsilon \operatorname{sign} \dot{\varepsilon}. \tag{19}$$

Zilustrowano to graficznie na rys. 6. Widać tu, że przy oddalaniu się punktu opisującego na płaszczyźnie fazowej od położenia równowagi w badanym układzie tłumienie jest większe



Rys. 6. Wykres zmian współczynnika tłumienia $\zeta = a+b\varepsilon \operatorname{sign} \dot{\varepsilon}$

od tłumienia występującego w analogicznej sytuacji w poprzednio rozpatrywanym układzie. Podobnie jak poprzednio, zmniejsza się ono następnie skokowo z chwilą, gdy $\dot{\varepsilon} = 0$ i punkt opisujący zacznie podążać w kierunku położenia równowagi.

Przechodząc z kolei do portretu fazowego, otrzymuje się z równania (16) zależności określające izokliny, a mianowicie:

$$v = -\frac{\varepsilon}{2b\varepsilon+(m+2a)} \quad \text{dla} \quad v > 0 \tag{20}$$

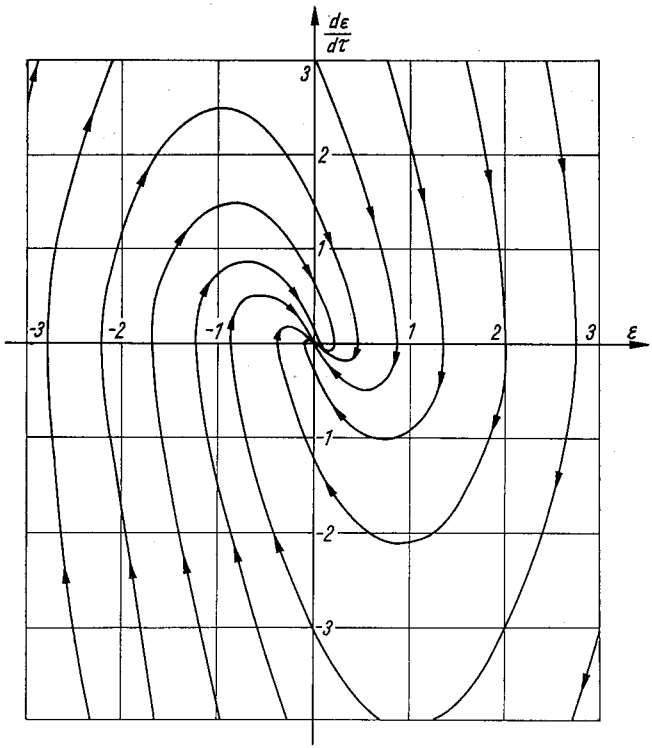
i

$$v = \frac{\varepsilon}{2b\varepsilon-(m+2a)} \quad \text{dla} \quad v < 0, \tag{21}$$

co można także zapisać w nieco inny sposób

$$v = \pm\frac{\varepsilon}{2b|\varepsilon|-(m+2a)}. \tag{22}$$

Ponadto oś rzędnych stanowi izoklinę odpowiadającą nachyleniu $m = -2a$. Uzyskany w ten sposób portret fazowy przedstawiono na rys. 7, przy czym przyjęto jak poprzednio $a = 1$ i $b = 0{,}9$. Z portretu tego widać, że w przypadku rozpatrywanego układu punkt osobliwy umiejscowiony w środku układu współrzędnych ma również charakter węzła stabilnego. Można także stwierdzić, że układ jest globalnie asymptotycznie stabilny.



Rys. 7. Portret fazowy układu opisanego równaniem
$\ddot{\varepsilon}+2(a+b\varepsilon \operatorname{sign}\dot{\varepsilon})\dot{\varepsilon}+\varepsilon = 0$ przy $x_0(t) = A1(t)$

Aby to udowodnić, wygodnie jest przedstawić równanie (16) w postaci układu równań pierwszego rzędu, po wprowadzeniu oznaczenia $\varepsilon = x_1$. Wtedy otrzymuje się:

$$\dot{x}_1 = x_2, \tag{23}$$

$$\dot{x}_2 = -2ax_2-(1+2b|x_2|)x_1. \tag{24}$$

Następnie szuka się funkcji Lapunowa w postaci

$$V(x_1, x_2) = F_1(x_1)+F_2(x_2). \tag{25}$$

Pochodna tej funkcji jest wobec tego określona zależnością

$$\frac{dV}{d\tau} = \frac{\partial F_1}{\partial x_1} x_2 - \frac{\partial F_2}{\partial x_2} [2ax_2 + (1+2b|x_2|)x_1]. \tag{26}$$

Dalej można postawić żądanie, żeby funkcja $\frac{dV}{d\tau}$ miała taką samą budowę jak V, tzn. żeby był tożsamościowo spełniony związek

$$\frac{\partial F_1}{\partial x_1} x_2 \equiv \frac{\partial F_2}{\partial x_2} (1+2b|x_2|)x_1, \tag{27}$$

czyli

$$\frac{F_1'(x_1)}{x_1} = \frac{F_2'(x_2)(1+2b|x_2|)}{x_2}. \tag{28}$$

Równość ta może być spełniona tylko wtedy, gdy każde z wyrażeń po obu stronach tej równości jest stałe. Założono, że wyrażenia te są równe 1. Wtedy

$$F_1(x_1) = \int_0^{x_1} x_1 \, dx_1 = \frac{x_1^2}{2}, \tag{29}$$

$$F_2(x_2) = \int_0^{x_2} \frac{x_2 dx_2}{1+2b|x_2|}. \tag{30}$$

Zatem po podstawieniu wzorów (29) i (30) do (25) otrzymuje się

$$V(x_1, x_2) = \frac{x_1^2}{2} + \int_0^{x_2} \frac{x_2 \, dx_2}{1+2b|x_2|}. \tag{31}$$

Przy $b \geqslant 0$ słusznym jest

$$1+2b|x_2| > 0, \tag{32}$$

skąd

$$F_2(x_2) > 0 \quad \text{dla} \quad x_2 \neq 0. \tag{33}$$

Z warunku (33) i wzoru (31) wynika, że $V(x_1, x_2)$ jest funkcją dodatnio określoną. Z kolei pochodna tej funkcji

$$\frac{dV}{d\tau} = -\frac{2ax_2^2}{1+2b|x_2|}, \tag{34}$$

przy $b \geqslant 0$ i $a > 0$, jest funkcją stale ujemną z wyjątkiem punktów położonych na osi x_1, gdzie przyjmuje wartość zero. Jednak ten zbiór punktów, w których $\frac{dV}{d\tau}$ jest równa zeru,

nie zawiera trajektorii zamkniętych z wyjątkiem środka układu współrzędnych. Wobec powyższego oraz ponieważ

$$F_1(x_1) \to \infty \quad \text{przy} \quad |x_1| \to \infty \tag{35}$$

i

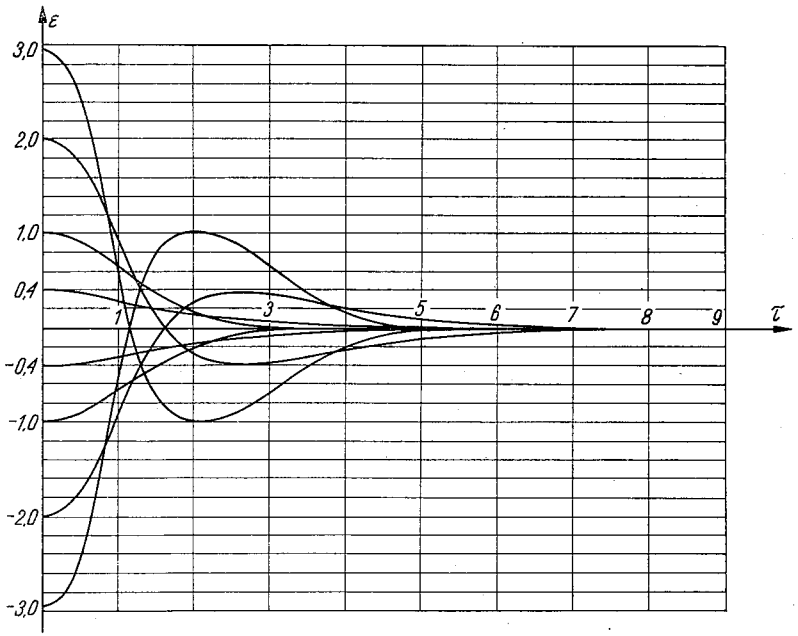
$$F_2(x_2) \to \infty \quad \text{przy} \quad |x_2| \to \infty, \tag{36}$$

czyli

$$V(x_1, x_2) \to \infty \quad \text{przy} \quad |x_1| \to \infty \quad \text{i} \quad |x_2| \to \infty, \tag{37}$$

rozważany układ jest rzeczywiście globalnie asymptotycznie stabilny.

Wracając do portretu fazowego analizowanego układu widać, że przebieg trajektorii jest zbliżony do przebiegu trajektorii poprzednio badanego układu. Wynikające stąd podobieństwo właściwości dynamicznych uwidaczniają także przebiegi czasowe przedstawione



Rys. 8. Przebiegi czasowe $\varepsilon(t)$ w układzie opisanym równaniem $\ddot{\varepsilon}+2(a+b\varepsilon \operatorname{sign}\dot{\varepsilon})\dot{\varepsilon}+\varepsilon = 0$ przy $x_0(t) = A1(t)$ i $\dot{\varepsilon}(0) = 0$

na rys. 8. Z porównania tych przebiegów z przebiegami przedstawionymi na rys. 3 widać, że w rozważanym układzie uzyskuje się nieco mniejsze przeregulowania w przypadku, gdy $|A| > 1$, przy jednoczesnym niewielkim wydłużeniu czasu ustalenia wszystkich przebiegów.

Badany układ ma tę istotną właściwość, że we wszystkich kwadrantach płaszczyzny fazowej $F(\varepsilon, \dot{\varepsilon})$ występuje współdziałanie sygnału korekcji nieliniowej z sygnałem uchybu. Wynika to z zależności określającej sygnał sterujący, mianowicie

$$u = \varepsilon + k_3 k_4 \varepsilon \left|\frac{dx}{dt}\right| - k_2 \frac{dx}{dt}, \tag{38}$$

skąd po uwzględnieniu wzorów (10) i (13) otrzymuje się

$$u = \varepsilon + \frac{k_3 k_4}{k_0}\varepsilon\left|\frac{d\varepsilon}{dt}\right| + \frac{k_2}{k_0}\frac{d\varepsilon}{dt}. \tag{39}$$

W układach, w których $\zeta = f(\varepsilon, \dot{\varepsilon})$ można uzyskać przebiegi czasowo optymalne przy skokowych sygnałach wejściowych, jeżeli narzuci się skokową zmianę ζ od wartości najmniejszej do wartości największej w określonych punktach płaszczyzny fazowej. Opis takiego układu podano w pracy [5]. Dowód optymalności przebiegów w układzie, w którym $\zeta = f(\varepsilon, \dot{\varepsilon})$ jest zmienną skokowo funkcją dwuwartościową, przytoczono w pracy [6], stosując zasadę maksimum Pontriagina.

Układy, w których $\zeta = f(\varepsilon, \dot{\varepsilon})$, stwarzają także możliwości syntezy quasi optymalnych układów nieprzełączających, zgodnie z sugestiami Gibsona [1] opartymi na pracy [2].

LITERATURA CYTOWANA W TEKŚCIE

1. J. E. Gibson, Nieliniowe układy sterowania automatycznego, WNT, Warszawa 1968.
2. R. E. Kuba, L. F. Kazda: A Phasespace Method for the Synthesis of Nonlinear Servomechanisms, Trans. AIEE, Vol. 75, part II, 1956, pp. 282—290.
3. K. Kuźmiński: Analiza dynamiki układów drugiego rzędu, w których tłumienie jest funkcją uchybu, Rozprawy Elektrotechniczne, z. 1, 1974.
4. Nelinejnye korrektirujuščie ustrojstwa w sistemach awtomatičeskogo uprawlenija, pod red. Ju. I. Topčewa, „Mašinostroenie", Moskwa 1971.
5. G. M. Ostrowskij, Primenenie nelinejnych korrektirujuščich ustrojstw w sistemach awtomatičeskogo regulirowanija wtorogo porjadka, Awtomatika i Telemechanika, t. XVII, no. 11, 1956, s. 979—984.
6. W. A. Taran, Primenenie nelinejnoj korrekcii i peremennoj struktury dlja ulučšenija dinamičeskich swojstw sistem awtomatičeskogo regulirowanija, Awtomatika i Telemechanika, t. XXV, no. 1, 1964, s. 140—149.
7. E. I. Chlypalo, Nelinejnye sistemy awtomatičeskogo regulirowanija, „Énergija", Leningrad 1967.

K. KUŹMIŃSKI

ANALYSIS OF DYNAMIC PROPERTIES OF SECOND ORDER SYSTEMS, IN WHICH ATTENUATION IS A FUNCTION OF ERROR AND ITS DERIVATIVE

Summary

The work contains the analysis of dynamic properties of systems described by the following differential equations:

$$\ddot{\varepsilon} + 2(a - b|\varepsilon|)\dot{\varepsilon} + \varepsilon = 0 \quad \text{for} \quad \varepsilon\dot{\varepsilon} \leqslant 0,$$

$$\ddot{\varepsilon} + 2a\dot{\varepsilon} + \varepsilon = 0 \quad \text{for} \quad \varepsilon\dot{\varepsilon} \geqslant 0$$

and

$$\ddot{\varepsilon} + 2(a + b\varepsilon \operatorname{sign}\dot{\varepsilon})\dot{\varepsilon} + \varepsilon = 0.$$

The diagrams of damping coefficient changes and phase pictures of these systems and corresponding time runs at $x_0(t) = A1(t)$ have been presented.

The influence of considered nonlinear correction on operation of automatic regulation system has been discussed.

K. KUŹMIŃSKI

ANALYSE DER DYNAMIK VON SYSTEMEN 2. ORDNUNG, WORIN DÄMPFUNG FUNKTION DER ABWEICHUNG UND DEREN ABLEITUNG IST

Zusammenfassung

Im Aufsatz wurden dynamischen Eigenschaften der Systeme analysiert, die mit folgenden Differentialgleichungen

$$\ddot{\varepsilon}+2(a-b|\varepsilon|)\dot{\varepsilon}+\varepsilon=0 \quad \text{für} \quad \varepsilon\dot{\varepsilon}\leqslant 0,$$

$$\ddot{\varepsilon}+2a\dot{\varepsilon}+\varepsilon=0 \quad \text{für} \quad \varepsilon\dot{\varepsilon}\geqslant 0$$

und

$$\ddot{\varepsilon}+2(a+b\varepsilon \operatorname{sign}\dot{\varepsilon})\dot{\varepsilon}+\varepsilon=0$$

beschrieben werden.

Es wurden Diagramme der Dämpfungskoeffizient-Veränderungen, Phasenporträte dieser Systeme und entsprechende Zeitverläufe bei $x_0(t) = A1(t)$ angegeben.

Der Einfluss der betrachteten, nichtlinearen Korrektion auf die Arbeit des automatischen Regelungs systems wurde diskutiert. Es wurde auch auf die Möglichkeiten einer Auswertung in der Synthese der Optimalsysteme hingewiesen.

К. КУЗЬМИНЬСКИ

АНАЛИЗ ДИНАМИЧЕСКИХ СВОЙСТВ СИСТЕМ ВТОРОГО ПОРЯДКА С ЗАТУХАНИЕМ, ЯВЛЯЮЩИМСЯ ФУНКЦИЕЙ ПОГРЕШНОСТИ И ЕЕ ПРОИЗВОДНОЙ

Резюме

В статье приведен анализ динамических свойств систем описанных дифференциальными уравнениями:

$$\ddot{\varepsilon}+2(a-b|\varepsilon|)\dot{\varepsilon}+\varepsilon=0 \quad \text{для} \quad \varepsilon\dot{\varepsilon}\leqslant 0$$

$$\ddot{\varepsilon}+2a\dot{\varepsilon}+\varepsilon=0 \quad \text{для} \quad \varepsilon\dot{\varepsilon}\geqslant 0$$

и

$$\ddot{\varepsilon}+2(a+b\varepsilon \operatorname{sign}\dot{\varepsilon})\dot{\varepsilon}+\varepsilon=0.$$

Даны диаграммы изменений коэффициента затухания, фазовые портреты этих систем и соответствующие переходные процессы для $x_0(t) = \mathrm{A}1(t)$.

Проанализировано влияние описываемой нелинейной коррекции на работу системы автоматического регулирования и указана возможность ее использования в синтезе оптимальных систем.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 1, ss. 95-103

621.372.4

Analiza dynamiki układu, w którym tłumienie i częstotliwość drgań własnych są funkcją uchybu i jego pochodnej

KRZYSZTOF KUŹMIŃSKI (ŁÓDŹ)
Instytut Automatyki i Elektroniki Politechniki Łódzkiej

Otrzymano 15.1.1973

W pracy zaproponowano układ z korekcją nieliniową, opisany równaniem różniczkowym:

$$\ddot{\varepsilon}+2a\dot{\varepsilon}+2b\varepsilon|\dot{\varepsilon}|+\varepsilon+\varepsilon|\varepsilon| = 0.$$

Przedyskutowano właściwości dynamiczne tego układu na tle układów analizowanych w pracach [1] i [2]. W tym celu wyznaczono portret fazowy tego układu oraz podano odpowiednie przebiegi czasowe przy $x_0(t) = A1(t)$. Ponadto dla rozważanego układu podano wykresy zmian częstotliwości drgań własnych oraz współczynnika tłumienia w zależności od uchybu i jego pochodnej. W oparciu o II metodę Lapunowa udowodniono globalną stabilność asymptotyczną badanego układu.

W pracy zamieszczono także schemat strukturalny układu według proponowanego sposobu korekcji oraz odpowiadający tej strukturze schemat ideowy serwomechanizmu.

W pracach [1] i [2] autor przeanalizował właściwości dynamiczne wybranych układów 2-go rzędu, w których współczynnik tłumienia ζ zależał od błędu regulacji ε [1], tzn. $\zeta = f_1(\varepsilon)$, lub od błędu regulacji i jego pochodnej [2], tzn. $\zeta = f_2(\varepsilon, \dot{\varepsilon})$. Przedyskutowano tam zalety i wady tego rodzaju korekcji nieliniowej. Na tle tych rozważań okazało się, że najbardziej korzystnym jest układ, opisany następującym równaniem różniczkowym:

$$\frac{d^2\varepsilon}{d\tau^2}+2a\frac{d\varepsilon}{d\tau}+\varepsilon+2b\varepsilon\left|\frac{d\varepsilon}{d\tau}\right| = 0, \tag{1}$$

gdzie

a i b — stałe współczynniki,

τ — czas bezwymiarowy określony zależnością

$$\tau = \omega t \tag{2}$$

ω — częstotliwość drgań własnych.

Równanie (1) można również przedstawić w postaci:

$$\frac{d^2\varepsilon}{d\tau^2}+2\left(a+b\varepsilon\,\mathrm{sign}\frac{d\varepsilon}{d\tau}\right)\frac{d\varepsilon}{d\tau}+\varepsilon = 0, \tag{3}$$

lub

$$\frac{d^2\varepsilon}{dt^2}+2f_2(\varepsilon,\dot{\varepsilon})\omega\frac{d\varepsilon}{dt}+\omega^2\varepsilon=0, \tag{4}$$

gdzie

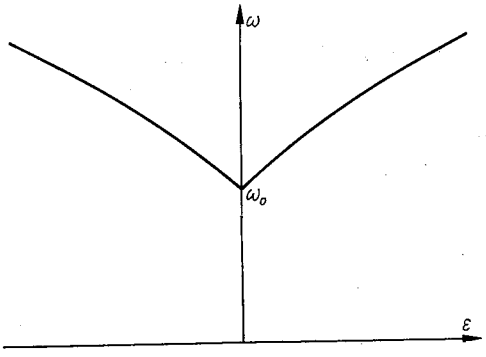
$$f_2(\varepsilon,\dot{\varepsilon})=\zeta=a+b\varepsilon\,\mathrm{sign}\frac{d\varepsilon}{dt}. \tag{5}$$

W celu dalszej poprawy właściwości dynamicznych układu opisanego równaniami (4) i (5), przy wymuszeniach w postaci funkcji skokowych

$$x_0(t)=A1(t) \tag{6}$$

słusznym jest odpowiednie uzależnienie od błędu regulacji nie tylko współczynnika tłumienia, ale również częstotliwości drgań własnych. Realizując ten warunek, otrzymuje się układ, który ogólnie opisuje następujące równanie różniczkowe:

$$\frac{d^2\varepsilon}{dt^2}+2f_2(\varepsilon,\dot{\varepsilon})\omega_0\frac{d\varepsilon}{dt}+\omega_0^2\varphi_1(\varepsilon)\varepsilon=0. \tag{7}$$



Rys. 1. Wykres zmian częstotliwości drgań własnych $\omega=\psi_1(\varepsilon)$

W równaniu (7), można przyjąć

$$\varphi_1(\varepsilon)=1+|\varepsilon|. \tag{8}$$

Wtedy częstotliwość drgań własnych określona jest zależnością

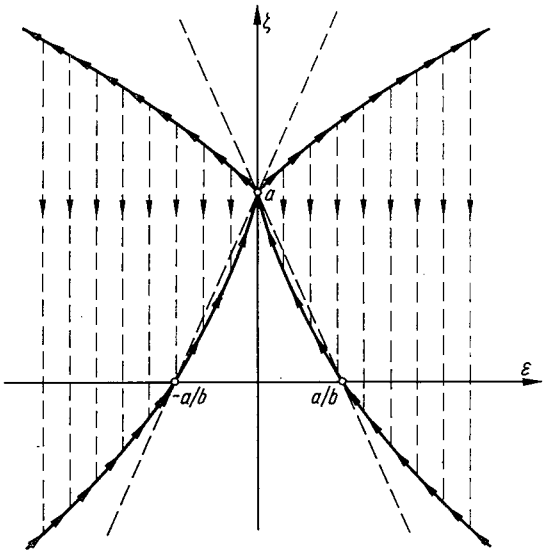
$$\omega=\psi_1(\varepsilon)=\omega_0\sqrt{1+|\varepsilon|}, \tag{9}$$

co uwidoczniono graficznie na rys. 1. Z kolei współczynnik tłumienia zmienia się zgodnie z zależnością

$$\zeta=\varphi_2(\varepsilon,\dot{\varepsilon})=\frac{a+b\varepsilon\,\mathrm{sign}\dot{\varepsilon}}{\sqrt{1+|\varepsilon|}}, \tag{10}$$

przedstawioną graficznie na rys. 2.

Jak wynika ze wzoru (9), wraz ze wzrostem bezwzględnej wartości błędu regulacji rośnie częstotliwość drgań własnych. W przypadku odwrotnym, gdy $|\varepsilon| \to 0$, częstotliwość drgań własnych dąży do wartości minimalnej ω_0. A zatem jest to zjawisko korzystne, ze względu na szybkość przebiegów przejściowych.



Rys. 2. Wykres zmian współczynnika tłumienia $\zeta = \varphi_2(\varepsilon, \dot{\varepsilon})$

Współczynnik tłumienia w rozważanym układzie, podobnie jak w układzie opisanym równaniem (3) [2], zmienia się w ten sposób, że wzrasta przy oddalaniu się punktu opisującego na płaszczyźnie fazowej $F(\varepsilon, \dot{\varepsilon})$ od położenia równowagi. Następnie zmniejsza się on skokowo do wartości określonej przez wzór (10), przy wartości ε odpowiadającej przeregulowaniu, które wystąpiło gdy $\dot{\varepsilon} = 0$. Wtedy jednak punkt opisujący podąża w kierunku położenia równowagi, przy czym im większa jest ta wartość ε, z tym mniejszym tłumieniem rozpoczyna się ruch w tym kierunku. Następnie tłumienie narasta przy zmniejszaniu się wartości uchybu. Jednak w proponowanym układzie wykres $\varphi_2(\varepsilon, \dot{\varepsilon})$ jest bardziej korzystny, niż wykres $f_2(\varepsilon, \dot{\varepsilon})$ odpowiadający układowi opisanemu równaniem (3), który zaznaczono linią przerywaną na rys. 2. Wynika to stąd, że przy danej wartości $|\varepsilon|$ w przedziale $-\frac{a}{b} \leqslant \varepsilon \leqslant \frac{a}{b}$, przy zbliżaniu się do punktu równowagi na płaszczyźnie fazowej $F(\varepsilon, \dot{\varepsilon})$, zależność (10) narzuca mniejsze wartości tłumienia w porównaniu z zależnością (5).

Tym samym w zasadniczym przedziale zmian ε, dla którego dobiera się parametry układu, zwiększa to szybkość przebiegów przejściowych. Natomiast dla $|\varepsilon| > \frac{a}{b}$, gdy punkt opisujący podąża w kierunku punktu równowagi, ujemne tłumienie narzucone przez zależność (10) jest większe w porównaniu z wartością określoną przez wzór (5). Jest to również korzystne, ponieważ uniemożliwia osiągnięcie przez układ zbyt dużej prędkości, przy $|\varepsilon(0)| \gg \frac{a}{b}$, takiej jaka wystąpiłaby przy obowiązującej zależności (5), która spowodowa-

łaby wystąpienie odpowiednio większego przeregulowania. Jedynie w przypadku oddalania się od położenia równowagi, układ dla którego obowiązuje wzór (10) ma nieco gorsze właściwości w porównaniu z układem opisanym równaniem (3), gdyż dla danej wartości $|\varepsilon|$ tłumienie osiąga wtedy mniejsze wartości.

Równanie (7) przy uwzględnieniu zależności (5), (8) i wprowadzeniu czasu bezwymiarowego odniesionego do ω_0, tzn.

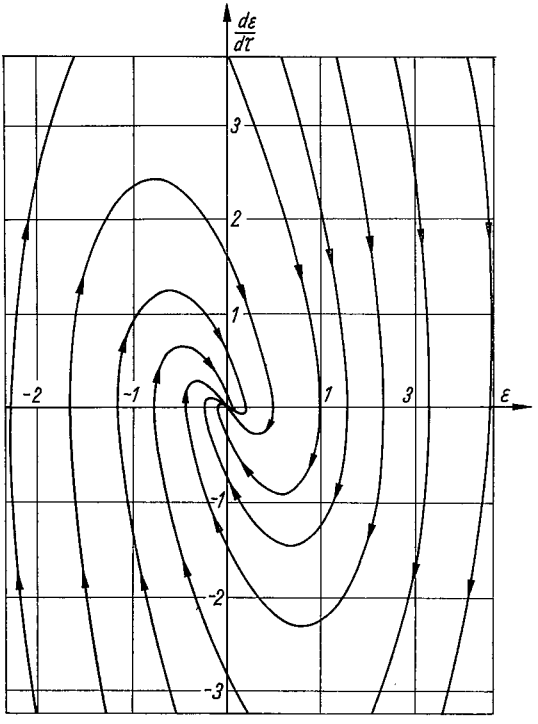
$$\tau = \omega_0 t, \tag{11}$$

można przedstawić w postaci:

$$\frac{d^2\varepsilon}{d\tau^2} + 2\left(a + b\varepsilon \operatorname{sign}\frac{d\varepsilon}{d\tau}\right)\frac{d\varepsilon}{d\tau} + (1 + |\varepsilon|)\varepsilon = 0, \tag{12}$$

lub

$$\ddot{\varepsilon} + 2a\dot{\varepsilon} + \varepsilon + 2b\varepsilon|\dot{\varepsilon}| + \varepsilon|\varepsilon| = 0. \tag{13}$$



Rys. 3. Portret fazowy układu opisanego równaniem $\ddot{\varepsilon} + 2a\dot{\varepsilon} + \varepsilon + 2b\varepsilon|\dot{\varepsilon}| + \varepsilon|\varepsilon| = 0$ przy $x_0(t) = A1(t)$

Posługując się równaniem (13) wygodnie jest określić portret fazowy rozpatrywanego układu. Dokonując podstawienia

$$v = \frac{d\varepsilon}{d\tau} \tag{14}$$

i przyjmując oznaczenie tangensa kąta nachylenia trajektorii fazowych jako m, otrzymuje się równanie izoklin w postaci:

$$v = -\frac{\varepsilon(1+|\varepsilon|)}{(m+2a)+2b\varepsilon} \quad \text{dla} \quad v > 0, \tag{15}$$

$$v = -\frac{\varepsilon(1+|\varepsilon|)}{(m+2a)-2b\varepsilon} \quad \text{dla} \quad v < 0. \tag{16}$$

Izoklinę stanowi także oś rzędnych odpowiadająca nachyleniu $m = -2a$. Portret fazowy układu przedstawiono na rys. 3, przyjmując podobnie jak w pracach [1], [2] $a = 1$ i $b = 0{,}9$. Z portretu tego widać, że punkt osobliwy umiejscowiony w środku układu współrzędnych ma charakter węzła stabilnego. Można także udowodnić, że układ jest globalnie asymptotycznie stabilny. W tym celu należy przyjąć funkcję Lapunowa w następującej postaci:

$$V(\varepsilon, \dot{\varepsilon}) = \frac{\dot{\varepsilon}^2}{2} + 2b\int_0^{\varepsilon} |\dot{\varepsilon}|\varepsilon d\varepsilon + \int_0^{\varepsilon} |\varepsilon|\varepsilon d\varepsilon + \frac{\varepsilon^2}{2}. \tag{17}$$

Z wyrażenia (17) wynika, że ponieważ

$$\begin{matrix} |\dot{\varepsilon}|\varepsilon^2 > 0 \\ |\varepsilon|\varepsilon^2 > 0 \end{matrix} \quad \text{dla} \quad \varepsilon \neq 0 \quad \text{i} \quad \dot{\varepsilon} \neq 0, \tag{18}$$

to przy $b \geqslant 0$, $V(\varepsilon, \dot{\varepsilon})$ jest funkcją dodatnio określoną.

Z kolei aby wyznaczyć pochodną funkcji Lapunowa, tzn. $\frac{dv}{d\tau}$, należy przekształcić równanie (13), które po scałkowaniu można zapisać w postaci:

$$\int_0^{\varepsilon} \dot{\varepsilon}\frac{d\dot{\varepsilon}}{d\varepsilon}d\varepsilon + 2a\int_0^{\varepsilon} \dot{\varepsilon}d\varepsilon + \int_0^{\varepsilon} \varepsilon d\varepsilon + 2b\int_0^{\varepsilon} |\dot{\varepsilon}|\varepsilon d\varepsilon + \int_0^{\varepsilon} |\varepsilon|\varepsilon d\varepsilon = 0. \tag{19}$$

Stąd dalej wynika związek

$$\frac{\varepsilon^2}{2} + \frac{\dot{\varepsilon}^2}{2} + 2b\int_0^{\varepsilon} |\dot{\varepsilon}|\varepsilon d\varepsilon + \int_0^{\varepsilon} |\varepsilon|\varepsilon d\varepsilon = -2a\int_0^{\tau} \dot{\varepsilon}^2 d\tau. \tag{20}$$

Zatem porównując wzory (17) i (20), otrzymuje się

$$V(\varepsilon, \dot{\varepsilon}) = -2a\int_0^{\tau} \dot{\varepsilon}^2 d\tau. \tag{21}$$

Wobec powyższego

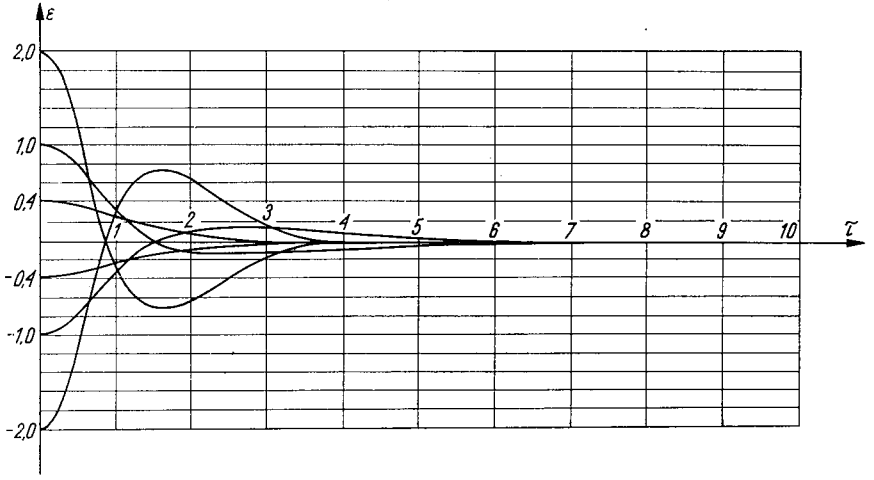
$$\frac{dV}{d\tau} = -2a\dot{\varepsilon}^2 \tag{22}$$

Jak wynika z zależności (22), przy $a > 0$, $\frac{dV}{d\tau} < 0$ w całej płaszczyźnie $F(\varepsilon, \dot{\varepsilon})$ z wyjątkiem punktów położonych na osi odciętych, gdzie przyjmuje wartość zero. Jednak ten

zbiór punktów, w których $\frac{dV}{d\tau}$ jest równa zeru, nie zawiera trajektorii zamkniętych oprócz środka układu współrzędnych. Ponadto ze wzoru (17) wynika, że

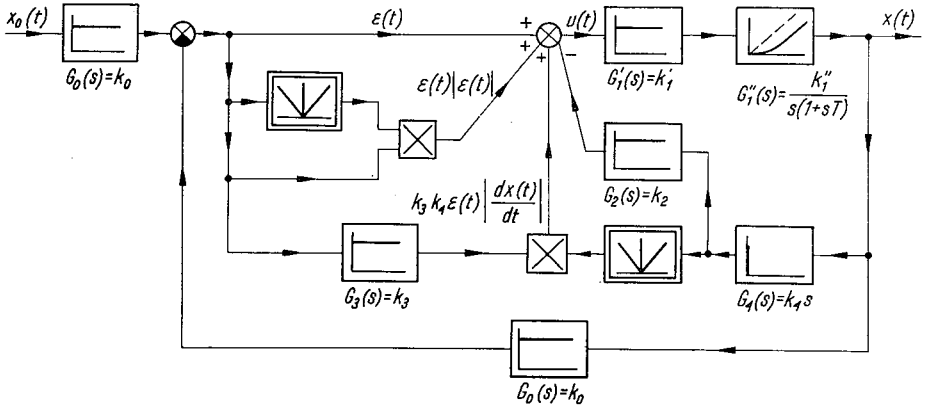
$$V(\varepsilon, \dot{\varepsilon}) \to \infty \text{ przy } |\varepsilon| \to \infty \text{ i } |\dot{\varepsilon}| \to \infty. \tag{23}$$

Zatem zostało wykazane, że układ opisany równaniem (13) jest rzeczywiście globalnie asymptotycznie stabilny [3].



Rys. 4. Przebiegi czasowe $\varepsilon(t)$ w układzie opisanym równaniem
$\ddot{\varepsilon}+2a\dot{\varepsilon}+\varepsilon+2b\varepsilon|\dot{\varepsilon}|+\varepsilon|\varepsilon| = 0$ przy $x_0(t) = A1(t)$ i $\dot{\varepsilon}(0) = 0$

Porównanie portretu fazowego z rys. *3* oraz przebiegów czasowych, przedstawionych na rys. 4 dla $\dot{\varepsilon}(0) = 0$ i $\varepsilon(0) > 0$ oraz $\varepsilon(0) < 0$ z portretami oraz przebiegami występującymi w układach opisanych w pracach [1] i [2] wskazuje na poprawę dynamiki układu. W rozpatrywanym bowiem układzie czas trwania procesów przejściowych jest krótszy przy pewnym jednak zwiększeniu przeregulowań, występujących gdy $|A| > 1$ (6). Podobnie

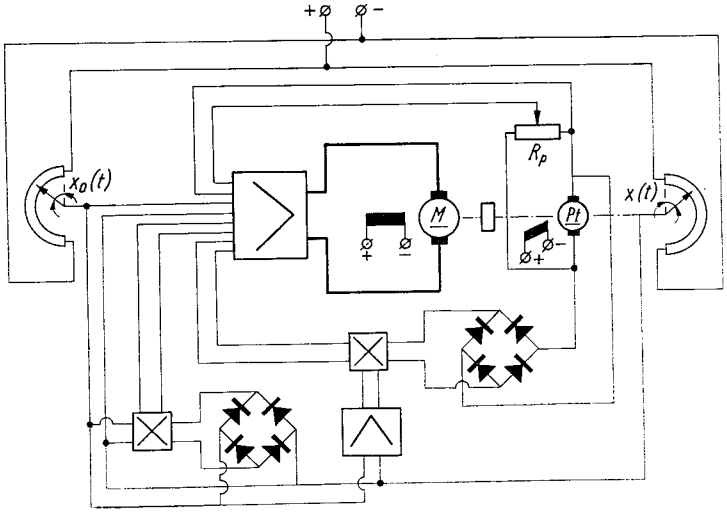


Rys. 5. Schemat strukturalny układu opisanego równaniem
$\ddot{\varepsilon}+2a\dot{\varepsilon}+\varepsilon+2b\varepsilon|\dot{\varepsilon}|+\varepsilon|\varepsilon| = 0$ przy $x_0(t) = A1(t)$

jak w układach opisanych w [1] i [2], w przypadku gdy $|A| < 1$ przebiegi są monotoniczne, z tym że czas ustalenia jest również krótszy. Rozpatrywany układ spełnia także warunek konieczny ze względu na zastosowanie praktyczne, mianowicie wykazuje symetrię działania w odniesieniu do znaku ε.

Schemat strukturalny układu, który opisuje równanie (13), przy spełnionym warunku (6) przedstawiono na rys. 5. Jak wynika z tego schematu, sygnał sterujący u określony jest zależnością

$$u = \varepsilon + k_3 k_4 \varepsilon \left| \frac{dx}{dt} \right| + \varepsilon|\varepsilon| - k_2 k_4 \frac{dx}{dt}. \tag{24}$$



Rys. 6. Schemat ideowy układu opisanego równaniem
$\ddot{\varepsilon} + 2a\dot{\varepsilon} + \varepsilon + 2b\varepsilon|\dot{\varepsilon}| + \varepsilon|\varepsilon| = 0$ przy $x_0(t) = A1(t)$

Uwzględniając jeszcze równanie sumatora

$$\varepsilon = k_0(x_0 - x) \tag{25}$$

oraz wzór (6) otrzymuje się ostatecznie:

$$u = \varepsilon + \frac{k_3 k_4}{k_0} \varepsilon \left| \frac{d\varepsilon}{dt} \right| + \varepsilon|\varepsilon| + \frac{k_2 k_4}{k_0} \frac{d\varepsilon}{dt}. \tag{26}$$

Jak wynika ze wzoru (26), we wszystkich kwadrantach płaszczyzny fazowej $F(\varepsilon, \dot{\varepsilon})$ występuje współdziałanie sygnału korekcji nieliniowej, określonego przez drugi i trzeci składnik tego wzoru z sygnałem uchybu ε, tzn. występuje zgodność ich znaków.

Na rys. 6 przedstawiono schemat ideowy serwomechanizmu zrealizowanego według struktury podanej na rys. 5. Przyjęto, że poszczególne elementy tego układu określone są przez odpowiednie transmitancje, zaznaczone na schemacie strukturalnym, mianowicie:

$G_0(s)$ — potencjometr zadający i wyjściowy,

$G_1(s)$ — wzmacniacz tranzystorowy lub tyrystorowy,

$G_1(s)$ — silnik obcowzbudny prądu stałego wraz z przekładnią,

$G_2(s)$ — potencjometr w obwodzie prądniczki tachometrycznej,
$G_3(s)$ — wzmacniacz wstępny tranzystorowy w obwodzie korekcji nieliniowej,
$G_4(s)$ — prądniczka tachometryczna.

Jako bloki wartości bezwzględnej przyjęto mostki Graetza. Zaś węzły mnożące, na schemacie ideowym zaznaczone umownie, można zrealizować stosując hallotrony.

Autor badał także właściwości dynamiczne układu ze zmienną strukturą, opisanego równaniami: (13) dla $\varepsilon\dot{\varepsilon} < 0$ i (1) dla $\varepsilon\dot{\varepsilon} > 0$. W tym przypadku przy oddalaniu się punktu opisującego na płaszczyźnie fazowej $F(\varepsilon, \dot{\varepsilon})$ od położenia równowagi obowiązuje wzór (5). Zatem wtedy współczynnik tłumienia zmienia się jak w układzie opisanym równaniem (1), czyli przyjmuje wartości większe w porównaniu z tymi, które przyjmowałby przy obowiązującej zależności (10). Natomiast przy zbliżaniu się punktu opisującego do położenia równowagi słuszny jest wzór (10).

W ten sposób można wyeliminować poprzednio wymienioną wadę układu opisanego równaniem (13) w porównaniu z właściwościami układu opisanego równaniem (1). Układ ze zmienną strukturą nie wprowadza jednak istotnych zmian właściwości w stosunku do układu opisanego równaniem (13). Można wprawdzie zauważyć minimalne zmniejszenie przeregulowań w przypadku, gdy $|A| > 1$, jednak czas trwania procesów przejściowych ulega jednocześnie wydłużeniu. Ponadto struktura takiego układu, wymagająca zastosowania elementu przełączającego, wypada bardziej skomplikowana w porównaniu z przedstawioną na rys. 5.

LITERATURA CYTOWANA W TEKŚCIE

1. K. Kuźmiński, Analiza dynamiki układów drugiego rzędu, w których tłumienie jest funkcją uchybu, Rozprawy Elektrotechniczne, z. 1, 1974.
2. K. Kuźmiński, Analiza dynamiki układów drugiego rzędu, w których tłumienie jest funkcją uchybu i jego pochodnej, Rozprawy Elektrotechniczne, z. 1, 1974.
3. J. La Salle, S. Lefschetz, Zarys teorii stabilności Lapunowa i jego metody bezpośredniej, PWN, 1966.

K. KUŹMIŃSKI

ANALYSIS OF DYNAMIC PROPERTIES OF A SYSTEM, IN WHICH ATTENUATION AND FREQUENCY OF CHARACTERISTIC OSCILLATIONS ARE FUNCTIONS OF ERROR AND ITS DERIVATIVE

Summary

The paper presents a control system with nonlinear correction described by differential equation:

$$\ddot{\varepsilon} + 2a\dot{\varepsilon} + 2b\varepsilon|\dot{\varepsilon}| + \varepsilon + \varepsilon|\varepsilon| = 0.$$

Dynamic proprieties of the system are discussed in comparison to the systems analysed in papers [1] and [2]. For this purpose the phase portrait and respective time characteristics at $x_0 = A1(t)$ are determined. Moreover the curves of natural frequency and damping ratio versus error and its derivative are given for the system in consideration. The overall asympthotic stability of the system is proved by apllication of the 2-nd method of Lapunow.

The paper also gives the block diagram of the system according to the suggested correction and a respective scheme of the servomechanism.

K. KUŹMIŃSKI

ANALYSE DER DYNAMIK VON SYSTEMEN 2. ORDNUNG, WORIN DÄMPFUNG UND FREQUENZ FUNKTIONEN DER ABWEICHUNG UND DEREN ABLEITUNG SIND

Zusammenfassung

In der vorliegenden Arbeit wurde ein System mit nichtlinearer Korrektion vorgeschlagen, das durch die Differentialgleichung

$$\ddot{\varepsilon}+2a\dot{\varepsilon}+2b\varepsilon|\dot{\varepsilon}|+\varepsilon+\varepsilon|\varepsilon| = 0$$

beschrieben wird.

Es wurden die Eigenschaften dieses Systems in Vergleich mit den in den Arbeiten [1] und [2] betrachteten Systemen diskutiert. Dann wurde das Phasenporträt angegeben und entsprechende Übergangsvorgänge bei $x_0 = A1(t)$ ermittelt. Asserdem wurden Kurven angeführt, welche Änderungen der Frequenzen von Einschwingungen und des Dämpfungsfaktors in Abhängigkeit von der Abweichung und deren Ableitung darstellen. Bei Anwendung der zweiten Methode von Lapunow wurde bewiesen, dass das betrachtete System im globalen Sinne asymptotisch stabil ist.

In der Arbeit wurde auch das Strukturschema gezeigt, welches der vorgeschlagenen Korrektion entspricht.

К. КУЗЬМИНЬСКИ

АНАЛИЗ ДИНАМИЧЕСКИХ СВОЙСТВ СИСТЕМ С ЗАТУХАНИЕМ И ЧАСТОТОЙ СОБСТВЕННЫХ КОЛЕБАНИЙ, ЯВЛЯЮЩИМИСЯ ФУНКЦИЯМИ ПОГРЕШНОСТИ И ЕЕ ПРОИЗВОДНОЙ

Резюме

В статье предложена система с нелинейной коррекцией, описанной дифференциальным уравнением:

$$\ddot{\varepsilon}+2a\dot{\varepsilon}+2b\varepsilon|\dot{\varepsilon}|+\varepsilon|\varepsilon|+\varepsilon = 0.$$

Проанализированы динамические свойства этой системы на основании анализа систем описанных в статьях [1] и [2]. С этой целью определен фазовый портрет этой системы и приведены соответствующие переходные процессы для $x_0(t) = A1(t)$. Кроме того, для описываемой системы даны диаграммы изменений собственной частоты и коэффициента затухания в зависимости от погрешности и ее производной. На базе второго метода Ляпунова доказано, что исследуемая система в большинстве асимптотически устойчива.

В статье приведена также структурная схема системы с предлагаемым способом коррекции и соответствующая этой структуре идейная схема сервомеханизма.

ROZPRAWY ELEKTRONICZNE 1974, 20, z. 1, ss. 105-122

621.391.81

Równanie transmisji między antenami w obszarze Fresnela

SEWERYN POGORZELSKI (WARSZAWA)

Przemysłowy Instytut Telekomunikacji

Otrzymano 19.12.1972

Rozpatruje się dwie bezstratne anteny (nadawczą i odbiorczą) o dowolnej strukturze i polaryzacji rozmieszczone dowolnie względem siebie w pustej przestrzeni. Wyprowadza się zależności dla współczynnika transmisji oraz dla energii odebranej w tych warunkach. Rozważa się przypadki, gdy każda z anten leży w strefie dalekiej drugiej anteny oraz gdy anteny są położone bliżej względem siebie. Zależności teoretyczne są ilustrowane prostym przykładem rachunku numerycznego dla dwu jednakowych anten parabolicznych.

1. WSTĘP

Przedmiotem rozważań jest problem transmisji energii między dwiema bezstratnymi antenami dowolnie względem siebie usytuowanymi w pustej przestrzeni. Ze względu na późniejsze uproszczenia rachunkowe zakłada się jednak, że odległość między antenami nie jest zbyt mała. Przyjmuje się też, że jedna z anten (oznaczona numerem I) jest nadawczą, a druga (oznaczona numerem II) — odbiorczą.

Pola elektromagnetyczne, jakie będą w związku z problemem transmisji rozpatrywane, mają harmoniczną zależność od czasu (wg funkcji $e^{-i\omega t}$). Są one rozwiązaniami układu równań Maxwella o postaci

$$\nabla \times \mathbf{H} = -i\omega\varepsilon\mathbf{E}, \tag{1}$$

$$\nabla \times \mathbf{E} = i\omega\mu\mathbf{H}. \tag{2}$$

Ważną rolę w rozważaniach odgrywać będzie uśredniony w czasie wektor Poyntinga spełniający zależność (gwiazdka oznacza wielkość sprzężoną)

$$\mathbf{S} = \frac{1}{4}(\mathbf{E} \times \mathbf{H}^* + \mathbf{E}^* \times \mathbf{H}), \tag{3}$$

lub równoważnie

$$\mathbf{S} = \frac{1}{2}\mathrm{Re}(\mathbf{E} \times \mathbf{H}^*). \tag{4}$$

Wektory $\mathbf{E}$, $\mathbf{H}$ są na ogół zespolone, natomiast wektor $\mathbf{S}$ jest rzeczywisty. Strumień wektora $\mathbf{S}$ przez dowolną gładką powierzchnię zamkniętą jest równy mocy czynnej jaka

wpływa do obszaru (lub wypływa z obszaru) ograniczonego tą powierzchnią. W dalszych rozważaniach moc oznaczać będzie zawsze moc czynną.

Wprowadza się symbole:

P_t — moc promieniowana przez antenę nadawczą,

P_r — moc odbierana przez antenę odbiorczą.

Moce te można mierzyć w liniach transmisyjnych zasilających obie anteny zakładając, że linie są dla danej częstotliwości obustronnie dopasowane.

Równanie o postaci

$$\frac{P_r}{P_t} = U(x_1, x_2, \ldots, x_n) \tag{5}$$

nazywamy równaniem transmisji. Zmienne x_i reprezentują parametry obu anten, ich orientację wzajemną, odległość, długość fali itd. Ważnym problemem w teorii anten jest znalezienie możliwie ogólnego wyrażenia dla funkcji U oraz dyskusja przypadków szczególnych.

Równanie transmisji jest znane od dawna dla przypadku, gdy każda z anten leży w strefie dalekiej drugiej anteny. Obowiązuje wtedy tzw. formuła Friisa

$$\frac{P_r}{P_t} = \left(\frac{\lambda}{4\pi R}\right)^2 G_1 G_2 \cos^2 \alpha, \tag{6}$$

gdzie:

λ — długość fali,

R — odległość anten,

G_1, G_2 — zyski anten,

2α — kąt między polaryzacjami obu anten na sferze Poincaré [7].

Jeśli anteny są położone bliżej względem siebie, równanie (6) staje się niedokładne i wymaga zastąpienia przez bardziej ogólną zależność. Przyjmuje się zwykle

$$\frac{P_r}{P_t} = TT^*, \tag{7}$$

gdzie T jest tzw. współczynnikiem transmisji (jest to wielkość zespolona). Koncepcja współczynnika transmisji została wprowadzona przez Ming-Kuei-Hu [1], a niezależnie od niego przez Robieux [2]. W pracach tych podaje się po raz pierwszy wyrażenia dla tego współczynnika. Dalsze rozwinięcie tej koncepcji ze wskazaniem zastosowań w miernictwie antenowym nastąpiło w pracach Kinbera i Cejtlina [5] oraz Pace'a [6].

Celem niniejszej pracy jest możliwe proste wyprowadzenie zależności dla współczynnika transmisji oraz równania transmisji w postaci (7). Wskazuje się również na możliwości przeprowadzania rachunków numerycznych przy użyciu maszyny cyfrowej, co ma ogromne znaczenie w technice antenowej. Wyniki podstawowe, tj. wyrażenia dla współczynnika transmisji oraz równanie transmisji w postaci (7), pokrywają się z wynikami zawartymi w cytowanych pracach, natomiast wyprowadzenia wzorów, dyskusja przypadków szczególnych oraz przykład rachunkowy są oryginalne.

2. POLA WŁASNE ANTEN NADAWCZEJ I ODBIORCZEJ

Przy wyprowadzeniu równania transmisji rozpatruje się pola własne anten nadawczej i odbiorczej.

Polem własnym anteny nadawczej nazywa się pole wytworzone przez źródła elementarne rozmieszczone w tej antenie, przy uwzględnieniu biernego oddziaływania na nie anteny odbiorczej. Pole własne anteny odbiorczej określa się w ten sam sposób jak pole anteny nadawczej zakładając, że role obu anten można zamienić (antena odbiorcza staje się nadawczą, zaś nadawcza — odbiorczą). Gdy każda z anten jest umieszczona oddzielnie w pustej przestrzeni, jej pole własne pokrywa się całkowicie z polem, jakie ona wytwarza. Sprawa komplikuje się przy obecności drugiej anteny. Niech $\mathbf{E}$, $\mathbf{H}$ oznacza pole pierwszej anteny wytworzone przy obecności drugiej anteny. Pole to jest sumą pola własnego pierwszej anteny i pola wytworzonego przez wtórne źródła elementarne rozmieszczone w drugiej antenie:

$$\mathbf{E} = \mathbf{E}^{\text{własne}} + \mathbf{E}^{\text{wtórne}},$$
$$\mathbf{H} = \mathbf{H}^{\text{własne}} + \mathbf{H}^{\text{wtórne}}.$$

Bierny wpływ drugiej anteny manifestuje się w dwojaki sposób: pojawia się pole wtórne, a pole własne pierwszej anteny wskutek oddziaływania między antenami różni się od pola własnego w pustej przestrzeni.

Ścisłe wyznaczenie pola własnego jest zwykle trudnym problemem dyfrakcyjnym, praktycznie niemożliwym do rozwiązania. Gdy odległość między antenami nie jest zbyt mała, zakłada się, że modyfikacja pola własnego jest niewielka i można wtedy przyjmować pole anteny w pustej przestrzeni w charakterze zadawalającego przybliżenia pola własnego. Takie uproszczenie wprowadza się zwykle w obliczeniach numerycznych współczynnika transmisji i równania transmisji. Nie rozwiązuje to sprawy oceny dopuszczalnej minimalnej odległości między antenami. Ze względu na trudności teorii dokładna ocena wydaje się możliwa głównie na drodze eksperymentalnej. Problem ten wykracza poza ramy obecnej pracy.

3. WSPÓŁCZYNNIK TRANSMISJI

Współczynnik transmisji jest wielkością odgrywającą zasadniczą rolę w równaniu transmisji (7). Koncepcja tego współczynnika opiera się na założeniu, że pola własne anteny nadawczej można rozdzielić na dwa odpowiednio dobrane pola składowe. Jedno z tych pól ma wszędzie topografię zgodną (por. dodatek) z polem własnym anteny odbiorczej, natomiast posiada przeciwny kierunek przepływu energii czynnej, tj. przeciwny kierunek uśrednionego w czasie wektora Poyntinga. Energia tego pola jest w całości odbierana przez antenę odbiorczą. Pozostałe pole składowe jest w pewnym sensie ortogonalne względem pola własnego anteny odbiorczej i nie wydziela energii czynnej w odbiorniku.

Wprowadzi się następujące oznaczenia:

$\mathbf{E}_1$, $\mathbf{H}_1$ — pole własne anteny I,

P_1 — odpowiednia moc wypromieniowana przez antenę I (w ogólności $P_1 \neq P_t$), $\mathbf{E}_2, \mathbf{H}_2$ — pole własne anteny II,

P_2 — odpowiednia moc wypromieniowana przez antenę II.

Pole własne anteny odbiorczej o przeciwnym kierunku uśrednionego wektora Poyntinga ma postać (por. dodatek)

$$-\mathbf{E}_2^*, \mathbf{H}_2^*.$$

Stosownie do uczynionego na wstępie założenia przedstawi się pole własne anteny nadawczej w postaci

$$\mathbf{E}_1 = -T\sqrt{\frac{P_1}{P_2}}\mathbf{E}_2^* + \mathbf{E}_d, \tag{8}$$

$$\mathbf{H}_1 = T\sqrt{\frac{P_1}{P_2}}\mathbf{H}_2^* + \mathbf{H}_d; \tag{9}$$

$T = \text{const}$ jest zespolonym współczynnikiem, który nazywa się współczynnikiem transmisji. Czynnik $\sqrt{\frac{P_1}{P_2}}$ uniezależnia wartość T od poziomu mocy P_1 i P_2.

W układzie równań (8), (9) wyodrębniono z pola własnego anteny nadawczej składnik o zgodnej topografii z polem własnym anteny odbiorczej, ale o przeciwnym kierunku przepływu energii czynnej. Reprezentuje on energię odebraną. Drugi składnik, tj. pole $\mathbf{E}_d, \mathbf{H}_d$, jest ortogonalny względem pola własnego anteny odbiorczej. Energia tego pola nie dociera do odbiornika. Warunek ortogonalności zostanie sformułowany w toku rachunku. Należy podkreślić, że bez sformułowania takiego dodatkowego warunku układ (8), (9) pozostaje nieokreślony — dla dowolnej wartości T istnieje zawsze pole $\mathbf{E}_d, \mathbf{H}_d$ spełniające ten układ.

Obliczenie współczynnika transmisji może być przeprowadzone w następujący sposób. Mnożymy wektorowo równanie (8) przez $\mathbf{H}_2$, zaś równanie (9) przez $\mathbf{E}_2$. Otrzymujemy

$$\mathbf{E}_1 \times \mathbf{H}_2 = -T\sqrt{\frac{P_1}{P_2}}\mathbf{E}_2^* \times \mathbf{H}_2 + \mathbf{E}_d \times \mathbf{H}_2, \tag{10}$$

$$\mathbf{H}_1 \times \mathbf{E}_2 = T\sqrt{\frac{P_1}{P_2}}\mathbf{H}_2^* \times \mathbf{E}_2 + \mathbf{H}_d \times \mathbf{E}_2. \tag{11}$$

Równania te należy dodać stronami i scałkować po dowolnej gładkiej powierzchni S_2 otaczającej antenę odbiorczą a nie otaczającej anteny nadawczej. Prowadzi to do zależności całkowej o postaci

$$\int_{S_2} (\mathbf{E}_1 \times \mathbf{H}_2 + \mathbf{H}_1 \times \mathbf{E}_2) \cdot \mathbf{n}\, dS = -T\sqrt{\frac{P_1}{P_2}} \int_{S_2} (\mathbf{E}_2^* \times \mathbf{H}_2 + \mathbf{E}_2 \times \mathbf{H}_2^*) \cdot \mathbf{n}\, dS +$$

$$+ \int_{S_2} (\mathbf{E}_d \times \mathbf{H}_2 + \mathbf{H}_d \times \mathbf{E}_2) \cdot \mathbf{n}\, dS. \tag{12}$$

Oznaczono tu przez $\mathbf{n}$ jednostkowy wektor normalny do powierzchni S_2 skierowany na zewnątrz (rys. 1).

Z kolei przyjmuje się warunek ortogonalności dla pola $\mathbf{E}_d$, $\mathbf{H}_d$ w postaci

$$\int_{S_2} (\mathbf{E}_d \times \mathbf{H}_2 + \mathbf{H}_d \times \mathbf{E}_2) \cdot \mathbf{n} dS = 0. \tag{13}$$

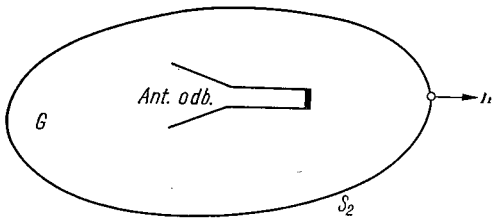
Nadto wiadomo, że

$$\int_{S_2} (\mathbf{E}_2^* \times \mathbf{H}_2 + \mathbf{E}_2 \times \mathbf{H}_2^*) \cdot \mathbf{n} dS = 4P_2. \tag{14}$$

Wykorzystując ostatnie dwie zależności otrzymuje się z (12) następujące wyrażenie dla współczynnika transmisji

$$T = -\frac{\int_{S_2} (\mathbf{E}_1 \times \mathbf{H}_2 + \mathbf{H}_1 \times \mathbf{E}_2) \cdot \mathbf{n} dS}{4(P_1 P_2)^{1/2}}. \tag{15}$$

Można wykazać, że wartość współczynnika transmisji nie zależy od wyboru powierzchni S_2.



Rys. 1. Powierzchnia S_2 otaczająca antenę odbiorczą

Fizyczna interpretacja warunku ortogonalności (13) sprowadza się do stwierdzenia, że prądy powierzchniowe płynące w antenie odbiorczej i będące źródłami jej pola własnego $\mathbf{E}_2$, $\mathbf{H}_2$ nie mogą być źródłami pola $\mathbf{E}_d$, $\mathbf{H}_d$. Aby to sprawdzić zakłada się, że powierzchnia S_2 pokrywa się z doskonale przewodzącą powierzchnią anteny odbiorczej. Wtedy na tej powierzchni

$$(\mathbf{H}_d \times \mathbf{E}_2) \cdot \mathbf{n} = \mathbf{H}_d \cdot (\mathbf{E}_2 \times \mathbf{n}) = 0 \tag{16}$$

i warunek ortogonalności sprowadza się do postaci

$$\int_{S_2} \mathbf{E}_d \cdot (\mathbf{n} \times \mathbf{H}_2) dS = 0. \tag{17}$$

Wektor $\mathbf{n} \times \mathbf{H}_2$ przedstawia gęstość prądów powierzchniowych będących źródłami pola własnego anteny odbiorczej. Z zależności (17) wynika, że prądy te nie mogą być źródłami pola $\mathbf{E}_d$, $\mathbf{H}_d$.

Znaleziona wartość współczynnika transmisji pozwala jednoznacznie określić pole $\mathbf{E}_d$, $\mathbf{H}_d$. Z równań (8), (9) wynika

$$\mathbf{E}_d = \mathbf{E}_1 + T \sqrt{\frac{P_1}{P_2}} \mathbf{E}_2^*, \tag{18}$$

$$\mathbf{H}_d = \mathbf{H}_1 - T \sqrt{\frac{P_1}{P_2}} \mathbf{H}_2^*. \tag{19}$$

Pole $\mathbf{E}_d$, $\mathbf{H}_d$ określone tymi wzorami spełnia układ Maxwella oraz warunek ortogonalności (13).

4. RÓWNANIE TRANSMISJI

Koncepcja podziału pola własnego anteny nadawczej na dwa składniki przedstawiona w poprzednim paragrafie toruje drogę do wyprowadzenia ogólnej postaci równania transmisji. Punktem wyjściowym będą w dalszym ciągu równania (8) i (9). Zastępując (9) równaniem sprzężonym otrzymuje się na wstępie układ o postaci

$$\mathbf{E}_1 = -T\sqrt{\frac{P_1}{P_2}}\,\mathbf{E}_2^* + \mathbf{E}_d,$$

$$\mathbf{H}_1^* = T^*\sqrt{\frac{P_1}{P_2}}\,\mathbf{H}_2 + \mathbf{H}_d^*.$$

Kolejnym krokiem jest pomnożenie wektorowe obu równań stronami przez siebie. Prowadzi to do zależności

$$\mathbf{E}_1 \times \mathbf{H}_1^* = -TT^*\frac{P_1}{P_2}\,\mathbf{E}_2^* \times \mathbf{H}_2 + \mathbf{E}_d \times \mathbf{H}_d^* + \sqrt{\frac{P_1}{P_2}}\,(T^*\mathbf{E}_d \times \mathbf{H}_2 - T\mathbf{E}_2^* \times \mathbf{H}_d^*).$$

Równanie to należy scałkować po powierzchni S_2 i zatrzymać jedynie połowę części rzeczywistej obu stron. W wyniku tych operacji uzyskuje się związek

$$\frac{1}{2}\operatorname{Re}\int_{S_2}(\mathbf{E}_1 \times \mathbf{H}_1^*)\cdot\mathbf{n}\,dS = -TT^*\frac{P_1}{P_2}\frac{1}{2}\operatorname{Re}\int_{S_2}(\mathbf{E}_2^* \times \mathbf{H}_2)\cdot\mathbf{n}\,dS +$$

$$+\frac{1}{2}\operatorname{Re}\int_{S_2}(\mathbf{E}_d \times \mathbf{H}_d^*)\cdot\mathbf{n}\,dS + \sqrt{\frac{P_1}{P_2}}\frac{1}{2}\operatorname{Re}\int_{S_2}(T^*\mathbf{E}_d \times \mathbf{H}_2 + T\mathbf{H}_d^* \times \mathbf{E}_2^*)\cdot\mathbf{n}\,dS. \tag{20}$$

Niech $P_1 = P_t$. Wtedy

$$\frac{1}{2}\operatorname{Re}\int_{S_2}(\mathbf{E}_1 \times \mathbf{H}_1^*)\cdot\mathbf{n}\,dS = -P_r, \tag{21}$$

$$\frac{1}{2}\operatorname{Re}\int_{S_2}(\mathbf{E}_2^* \times \mathbf{H}_2)\cdot\mathbf{n}\,dS = P_2, \tag{22}$$

$$\frac{1}{2}\operatorname{Re}\int_{S_2}(\mathbf{E}_d \times \mathbf{H}_d^*)\cdot\mathbf{n}\,dS = 0, \tag{23}$$

$$\frac{1}{2}\operatorname{Re}\int_{S_2}(T^*\mathbf{E}_d \times \mathbf{H}_2 + T\mathbf{H}_d^* \times \mathbf{E}_2^*)\cdot\mathbf{n}\,dS = 0. \tag{24}$$

Równania (21) i (22) nie wymagają uzasadnienia. Równanie (23) jest konsekwencją założenia, że pole $\mathbf{E}_d$, $\mathbf{H}_d$ nie wydziela energii czynnej w obszarze ograniczonym powierzchnią S_2. Równanie (24) wynika z warunku ortogonalności (13). Istotnie niech

$$\mathbf{E}_d \times \mathbf{H}_2 = \mathbf{a} + i\mathbf{b},$$
$$\mathbf{H}_d^* \times \mathbf{E}_2^* = \mathbf{c} - i\mathbf{d},$$
$$T = T_1 + iT_2,$$

gdzie **a**, **b**, **c**, **d**, T_1, T_2 są rzeczywiste. Wtedy

$$\operatorname{Re}\int_{S_2}(T^*\mathbf{E}_d\times\mathbf{H}_2+T\mathbf{H}_d^*\times\mathbf{E}_2^*)\cdot\mathbf{n}\,dS = T_1\int_{S_2}(\mathbf{a}+\mathbf{c})\cdot\mathbf{n}\,dS+T_2\int_{S_2}(\mathbf{b}+\mathbf{d})\cdot\mathbf{n}\,dS = 0,$$

ponieważ zgodnie z (13) można napisać

$$\int_{S_2}(\mathbf{a}+\mathbf{c})\cdot\mathbf{n}\,dS = 0,$$

$$\int_{S_2}(\mathbf{b}+\mathbf{d})\cdot\mathbf{n}\,dS = 0.$$

Uwzględniając (21), (22), (23) i (24) w (20) dochodzi się do równania transmisji w postaci (7)

$$\frac{P_r}{P_t} = TT^*.$$

Współczynnik transmisji nie zależy od tego, którą z anten I lub II obierze się jako nadawczą, tj. nie zależy od kierunku transmisji między antenami. Wynika stąd, że dla danej pary anten stosunek P_r/P_t nie zależy również od kierunku transmisji.

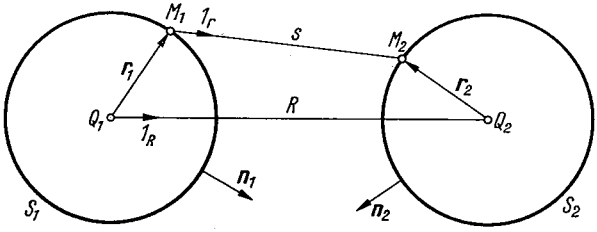
5. PRZEKSZTAŁCENIE WYRAŻENIA DLA WSPÓŁCZYNNIKA TRANSMISJI

Obliczenia numeryczne jak również konstruowanie wyrażeń przybliżonych dla współczynnika transmisji określonego przez (15) są znacznie łatwiejsze po odpowiednim przekształceniu tego wzoru.

Na wstępie otrzymuje się z (15)

$$T = \frac{1}{4(P_1P_2)^{1/2}}\int_{S_2}[\mathbf{E}_1\cdot(\mathbf{n}_2\times\mathbf{H}_2)+\mathbf{H}_1\cdot(\mathbf{n}_2\times\mathbf{E}_2)]\,dS_2, \tag{25}$$

gdzie $\mathbf{n}_2 = \mathbf{n}$ — wektor normalny zewnętrzny do powierzchni S_2.



Rys. 2. Szkic do przekształcenia wzoru (15)

Niech S_1 oznacza gładką powierzchnię otaczającą antenę nadawczą (rys. 2). Zakładając, że odległość między antenami jest duża w porównaniu z długością fali, można zapisać pole własne anteny nadawczej w postaci

$$\mathbf{E}_1(M_2) = \frac{ik}{4\pi}\int_{S_1}\mathbf{A}_1\frac{e^{iks}}{s}dS_1, \tag{26}$$

$$\mathbf{H}_1(M_2) = \frac{ik}{4\pi}\sqrt{\frac{\varepsilon}{\mu}}\int_{S_1}(\mathbf{1}_s\times\mathbf{A}_1)\frac{e^{iks}}{s}\,dS_1, \tag{27}$$

gdzie

$$\mathbf{A}_1 = \sqrt{\frac{\mu}{\varepsilon}}(\mathbf{n}_1 \times \mathbf{H}_1)^{tr} + \mathbf{1}_s \times (\mathbf{n}_1 \times \mathbf{E}_1), \tag{28}$$

$$(\mathbf{n} \times \mathbf{H}_1)^{tr} = [\mathbf{1}_s \times (\mathbf{n}_1 \times \mathbf{H}_1)] \times \mathbf{1}_s. \tag{29}$$

Znaczenie odległości s oraz wektorów jednostkowych $\mathbf{n}_1$, $\mathbf{1}_s$ widoczne jest z rys. 2.

Po podstawieniu (26) i (27) do (25) otrzymuje się

$$T = \frac{ik}{16\pi(P_1 P_2)^{1/2}} \int_{S_2} \int_{S_1} \left[\mathbf{A}_1 \cdot (\mathbf{n}_2 \times \mathbf{H}_2) + \sqrt{\frac{\varepsilon}{\mu}} (\mathbf{1}_s \times \mathbf{A}_1) \cdot (\mathbf{n}_2 \times \mathbf{E}_2) \right] \frac{e^{iks}}{s} dS_1 dS_2,$$

a następnie uwzględniając $\mathbf{A}_1 \cdot \mathbf{1}_s = 0$

$$T = \frac{ik}{16\pi(P_1 P_2)^{1/2}} \sqrt{\frac{\varepsilon}{\mu}} \int_{S_2} \int_{S_1} \mathbf{A}_1 \cdot \left[\sqrt{\frac{\mu}{\varepsilon}} (\mathbf{n}_2 \times \mathbf{H}_2)^{tr} - \mathbf{1}_s \times (\mathbf{n}_2 \times \mathbf{E}_2) \right] \frac{e^{iks}}{s} dS_1 dS_2. \tag{30}$$

Wyrażenie w nawiasie kwadratowym reprezentuje na powierzchni S_2 wektor $\mathbf{A}_2$ analogiczny do wektora A_1 na powierzchni S_1. W wyniku otrzymuje się wyrażenie dla T symetryczne względem obu anten:

$$T = \frac{ik}{16\pi(P_1 P_2)^{1/2}} \sqrt{\frac{\varepsilon}{\mu}} \int_{S_1} \int_{S_2} A_1 \cdot \mathbf{A}_2 \frac{e^{iks}}{s} dS_1 dS_2. \tag{31}$$

6. RÓWNANIE TRANSMISJI W STREFIE DALEKIEJ

Rozpatrzy się przypadek, gdy każda z anten znajduje się w strefie dalekiej drugiej anteny. Wtedy można w wyrażeniu (31) wprowadzić uproszczenia właściwe dla $R \to \infty$:

$$\frac{1}{s} = \frac{1}{R}, \tag{32}$$

$$\mathbf{1}_s = \mathbf{1}_R, \tag{33}$$

$$e^{iks} = e^{ik[R - \mathbf{1}_R \cdot (\mathbf{r}_1 - \mathbf{r}_2)]}. \tag{34}$$

Ostatnią z tych równości uzasadnia się jak następuje. Zgodnie z rys. 2 słuszna jest zależność wektorowa

$$s = \mathrm{R} - (\mathbf{r}_1 - \mathbf{r}_2). \tag{35}$$

Po podniesieniu obu stron do kwadratu i pierwiastkowaniu otrzymuje się

$$s = R \sqrt{1 - \frac{2\mathbf{1}_R \cdot (\mathbf{r}_1 - \mathbf{r}_2)}{R} + \frac{(\mathbf{r}_1 - \mathbf{r}_2)^2}{R^2}}.$$

Rozwijając pierwiastek na szereg potęgowy dostaje się

$$s = R - \mathbf{1}_R \cdot (\mathbf{r}_1 - \mathbf{r}_2) + \frac{[\mathbf{1}_R \times (\mathbf{r}_1 - \mathbf{r}_2)]^2}{2R} + \ldots \tag{36}$$

W strefie dalekiej wystarczy dla obliczenia e^{iks} zatrzymać pierwsze dwa wyrazy tego szeregu.

Podstawiając (32), (33) i (34) do (31) uzyskuje się

$$T = \frac{ik}{16\pi(P_1P_2)^{1/2}} \sqrt{\frac{\varepsilon}{\mu}} \frac{e^{ikR}}{R} \int_{S_1} \mathbf{A}_1(\mathbf{1}_R) e^{-ik\mathbf{1}_R\cdot\mathbf{r}_1} dS_1 \cdot \int_{S_2} \mathbf{A}_2(-\mathbf{1}_R) e^{ik\mathbf{1}_R\cdot\mathbf{r}_2} dS_2. \quad (37)$$

Wykorzysta się wyrażenia dla pola własnego obu anten w strefie dalekiej. Mają one postać

$$\mathbf{E}_1(Q_2) = \frac{ik}{4\pi} \frac{e^{ikR}}{R} \int_{S_1} \mathbf{A}_1(\mathbf{1}_R) e^{-ik\mathbf{1}_R\cdot\mathbf{r}_1} dS_1, \quad (38)$$

$$\mathbf{E}_2(Q_1) = \frac{ik}{4\pi} \frac{e^{ikR}}{R} \int_{S_2} \mathbf{A}_2(-\mathbf{1}_R) e^{ik\mathbf{1}_R\cdot\mathbf{r}_2} dS_2. \quad (39)$$

Uwzględniając te zależności otrzymuje się z (37)

$$T = \frac{\pi R e^{-ikR}}{ik(P_1P_2)^{1/2}} \sqrt{\frac{\varepsilon}{\mu}} \mathbf{E}_1(Q_2)\cdot\mathbf{E}_2(Q_1). \quad (40)$$

Pole anteny w strefie dalekiej można zapisać jak następuje

$$\mathbf{E} = \frac{e^{ikR}}{R} \mathbf{F}(\mathbf{1}_R); \quad (41)$$

$\mathbf{F}(\mathbf{1}_R)$ oznacza funkcję promieniowania anteny. Biorąc pod uwagę powyższą zależność uzyskuje się z (40)

$$T = \frac{\pi}{ik(P_1P_2)^{1/2}} \sqrt{\frac{\varepsilon}{\mu}} \frac{e^{ikR}}{R} \mathbf{F}_1\cdot\mathbf{F}_2. \quad (42)$$

Wyrażenie (42) pozwala napisać równanie transmisji dla strefy dalekiej:

$$\frac{P_r}{P_t} = TT^* = \frac{\pi^2}{k^2R^2P_1P_2} \frac{\varepsilon}{\mu} |\mathbf{F}_1\cdot\mathbf{F}_2|^2. \quad (43)$$

Aby doprowadzić to równanie do ogólnie znanej postaci wykorzysta się związki

$$|\mathbf{F}_1\cdot\mathbf{F}_2| = |\mathbf{F}_1||\mathbf{F}_2|\cos\alpha, \quad (44)$$

$$\sqrt{G(\mathbf{1}_R)} = \sqrt{\frac{2\pi}{P}} \sqrt{\frac{\varepsilon}{\mu}} |\mathbf{F}(\mathbf{1}_R)|, \quad (45)$$

gdzie $G(\mathbf{1}_R)$ oznacza zysk anteny dla kierunku $\mathbf{1}_R$, zaś P jest mocą wypromieniowaną. Wprowadzenie zależności (44) i (45) do (43) daje konwencjonalną postać równania transmisji

$$\frac{P_r}{P_t} = \left(\frac{\lambda}{4\pi R}\right)^2 G_1 G_2 \cos^2\alpha. \quad (46)$$

7. RÓWNANIE TRANSMISJI W OBSZARZE FRESNELA

Na wstępie rozpatrzy się najprostszy przypadek, gdy jedna z anten, np. odbiorcza, znajduje się w obszarze Fresnela drugiej anteny (nadawczej), natomiast antena nadawcza jest w strefie dalekiej anteny odbiorczej. Taka sytuacja zaistnieje np. wtedy, gdy antena nadaw-

cza jest duża, natomiast antena odbiorcza mała (np. sonda dipolowa). Biorąc pod uwagę (36) można wtedy przyjąć

$$e^{iks} = \exp ik\left[R - \mathbf{1}_R \cdot (\mathbf{r}_1 - \mathbf{r}_2) + \frac{(\mathbf{1}_R \times \mathbf{r}_1)^2}{2R}\right]. \tag{47}$$

Jeśli odległość między antenami nie jest zbyt mała, to podobnie jak dla strefy dalekiej

$$\frac{1}{s} = \frac{1}{R}; \quad \mathbf{1}_s = \mathbf{1}_R$$

Podstawiając te wartości do (31) otrzymuje się

$$T = \frac{ik}{16\pi(P_1P_2)^{1/2}} \sqrt{\frac{\varepsilon}{\mu}} \frac{e^{ikR}}{R} \int_{S_1} \mathbf{A}_1(\mathbf{1}_R) e^{-ik\left[\mathbf{1}_R \cdot \mathbf{r}_1 - \frac{(\mathbf{1}_R \times \mathbf{r}_1)^2}{2R}\right]} dS_1 \times$$

$$\times \int_{S_2} \mathbf{A}_2(-\mathbf{1}_R) e^{ik\mathbf{1}_R \cdot \mathbf{r}_2} dS_2. \tag{48}$$

W konsekwencji prowadzi to do wyrażenia analogicznego do (40)

$$T = \frac{\pi R e^{-ikR}}{ik(P_1P_2)^{1/2}} \sqrt{\frac{\varepsilon}{\mu}} \mathbf{E}_{1F}(Q_2) \cdot \mathbf{E}_2(Q_1). \tag{49}$$

Należy podkreślić, że pole anteny nadawczej $\mathbf{E}_{1F}$ jest polem w obszarze Fresnela. Postępując dalej podobnie jak w p. 6 dochodzi się do równania transmisji w postaci analogicznej do (46)

$$\frac{P_r}{P_t} = \left(\frac{\lambda}{4\pi R}\right)^2 G_{1F} G_2 \cos^2 \alpha. \tag{50}$$

Zysk anteny nadawczej G_{1F} jest zyskiem w obszarze Fresnela. Zysk taki oblicza się z zależności

$$G_F = \frac{|\mathbf{E}_F|^2}{|\mathbf{E}|^2} G, \tag{51}$$

gdzie $\mathbf{E}_F$ oznacza pole w obszarze Fresnela w odległości R od anteny, $\mathbf{E}$ oznacza pole obliczone w tej samej odległości R według zależności dla strefy dalekiej, zaś G jest zyskiem anteny dla strefy dalekiej.

Znaczenie parametru G_F jest dość ograniczone, gdyż — jak to wyniknie z dalszych rozważań — nie można równania transmisji uogólnić tak, aby wystąpił w nim iloczyn dwóch zysków w obszarze Fresnela.

Z kolei rozważy się bardziej ogólną sytuację, gdy każda z anten znajduje się w obszarze Fresnela drugiej anteny. Należy wtedy korzystać przy obliczaniu współczynnika transmisji T z wyrażenia (31) w pełnej postaci, przy czym całki po powierzchniach S_1, S_2 nie da się przedstawić w postaci iloczynu dwóch całek powierzchniowych. Stąd właśnie wynika niemożliwość skonstruowania w ogólnym przypadku równania transmisji, które zawierałoby iloczyn zysków anten. Pojęcie zysku anteny, tak ważne w strefie dalekiej, traci częściowo swą użyteczność w obszarze Fresnela.

Obliczenie całki (31) wymaga użycia szybko liczącej maszyny cyfrowej o odpowiednio dużej pamięci. W wielu przypadkach korzystne jest stosowanie do obliczenia całki wielokrotnej metody Monte Carlo [8]. W wyrażeniu podcałkowym (31) można — zależnie od odległości między antenami — stosować różne gradacje przybliżeń. „Najgrubsze" z tych przybliżeń polega na przyjęciu

$$\frac{1}{s} = \frac{1}{R}; \quad \mathbf{1}_s = \mathbf{1}_R \tag{52}$$

oraz zgodnie z (36)

$$e^{iks} = \exp ik\left\{R - \mathbf{1}_R \cdot (\mathbf{r}_1 - \mathbf{r}_2) + \frac{[\mathbf{1}_R \times (\mathbf{r}_1 - \mathbf{r}_2)]^2}{2R}\right\}. \tag{53}$$

Poza tym stosuje się konwencjonalnie następujące uproszczenia:

1) korzysta się z przybliżonych wartości pól własnych $\mathbf{E}_1$, $\mathbf{H}_1$ oraz $\mathbf{E}_2$, $\mathbf{H}_2$ na powierzchniach S_1 i S_2. Zaniedbuje się oddziaływanie wzajemne anten na te wartości. Wynika stąd potrzeba zachowania dostatecznie dużej odległości między antenami (rzędu 0,1 promienia strefy dalekiej). Dokładniejsze ustalenie minimalnej odległości, dla której można zaniedbać oddziaływanie wzajemne anten, może być przeprowadzone eksperymentalnie dla określonego typu anten;

2) całkuje się tylko po odpowiednio dobranych częściach powierzchni S_1 i S_2 (np. po „oświetlonych" częściach reflektorów antenowych lub po odpowiednio dobranych płaskich aperturach).

Należy również zwrócić uwagę na poczynione na wstępie założenie, że anteny umieszczone są w pustej przestrzeni. Założenia tego nie da się ściśle realizować — zawsze trzeba liczyć się z wpływem odbić od ziemi i przedmiotów otaczających.

Wpływ wymienionych uproszczeń rachunkowych na dokładność wyników jest trudny do ogólnej analizy. Z tego względu nieodzowna jest w wielu przypadkach konfrontacja wyników z eksperymentem.

W dalszym ciągu rozpatrzy się ważniejsze przypadki szczególne obliczania współczynnika transmisji.

A) Zakłada się, że obie anteny są antenami reflektorowymi (tj. każda antena składa się z reflektora i źródła oświetlającego). Niech S_1 i S_2 oznaczają powierzchnie reflektorów. Wtedy

$$\mathbf{A}_1 = \sqrt{\frac{\mu}{\varepsilon}}\,(\mathbf{n}_1 \times \mathbf{H}_1)^{tr},$$

$$\mathbf{A}_2 = \sqrt{\frac{\mu}{\varepsilon}}\,(\mathbf{n}_2 \times \mathbf{H}_2)^{tr}$$

i otrzymuje się

$$T = \frac{ik}{16\pi (P_1 P_2)^{1/2}} \sqrt{\frac{\mu}{\varepsilon}} \int_{S_1}\int_{S_2} (\mathbf{n}_1 \times \mathbf{H}_1)^{tr} \cdot (\mathbf{n}_2 \times \mathbf{H}_2)^{tr} \frac{e^{iks}}{s}\, dS_1\, dS_2. \tag{54}$$

B) Powierzchnie S_1 i S_2 są aperturami, pola $\mathbf{E}$ i $\mathbf{H}$ są styczne do powierzchni apertur, nadto w każdej aperturze zachodzi związek

$$\mathbf{H} = \sqrt{\frac{\varepsilon}{\mu}}\,\mathbf{n} \times \mathbf{E}.$$

Przy tych założeniach

$$\mathbf{A}_1 = \mathbf{1}_s \times [(\mathbf{n}_1 + \mathbf{1}_s) \times \mathbf{E}_1],$$
$$\mathbf{A}_2 = [(\mathbf{n}_2 - \mathbf{1}_s) \times \mathbf{E}_2] \times \mathbf{1}_s.$$

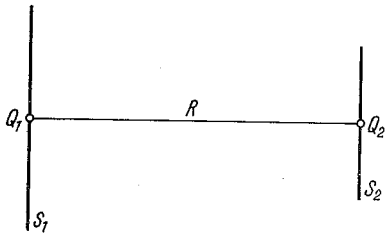
Iloczyn $\mathbf{A}_1 \cdot \mathbf{A}_2$ przybiera prostą postać, gdy można dodatkowo przyjąć

$$\mathbf{n}_1 = -\mathbf{n}_2 = \mathbf{1}_s.$$

Otrzymuje się wtedy

$$\mathbf{A}_1 = -2\mathbf{E}_1; \quad \mathbf{A}_2 = -2\mathbf{E}_2,$$
$$\mathbf{A}_1 \cdot \mathbf{A}_2 = 4\mathbf{E}_1 \cdot \mathbf{E}_2. \tag{55}$$

Taka sytuacja zachodzi np. dla układu dwóch płaskich apertur równoległych apertur promieniujących o wspólnej osi (rys. 3).



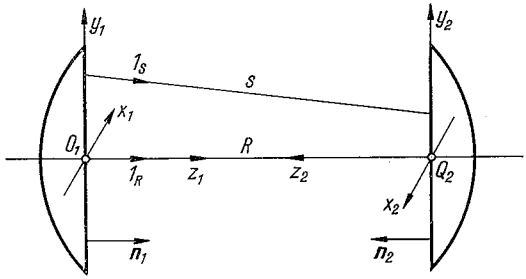
Rys. 3. Układ dwu płaskich apertur

Zgodnie z (55) współczynnik transmisji wyrazi się następująco

$$T = \frac{ik}{4\pi(P_1 P_2)^{1/2}} \sqrt{\frac{\varepsilon}{\mu}} \int_{S_1} \int_{S_2} \mathbf{E}_1 \cdot \mathbf{E}_2 \frac{e^{iks}}{s} \, dS_1 \, dS_2. \tag{56}$$

8. PRZYKŁAD ANTEN PARABOLICZNYCH

Rozpatrzy się układ dwu jednakowych anten parabolicznych o aperturach kołowych usytuowanych względem siebie jak na rys. 4 (osi symetrii reflektorów obu anten pokrywają się ze sobą). Dla uproszczenia przyjmie się $\mathbf{1}_s = \mathbf{n}_1 = -\mathbf{n}_2$. Wtedy współczynnik trans-



Rys. 4. Układ anten parabolicznych

misji wyrazi się wzorem (56). W aperturach anten wprowadzi się współrzędne prostokątne x_1, y_1, z_1 oraz x_2, y_2, z_2. Odpowiednie współrzędne biegunowe oznacza się przez r_1, ψ_1 oraz r_2, ψ_2. Przyjmie się, że pola w aperturach obu anten mają stałą zgodną polaryzację.

Niech np.

$$\mathbf{E}_1 = E_1 \mathbf{1}_{y1}; \quad \mathbf{E}_2 = E_2 \mathbf{1}_{y2}.$$

Wtedy

$$\mathbf{E}_1 \cdot \mathbf{E}_2 = E_1 E_2. \tag{57}$$

Dalsze rachunki będą przeprowadzone dla parabolicznego rozkładu pola:

$$E_1 = 1 - \left(\frac{2r_1}{D}\right)^2, \tag{58}$$

$$E_2 = 1 - \left(\frac{2r_2}{D}\right)^2. \tag{59}$$

Z kolei założy się w rachunku

$$\frac{1}{s} = \frac{1}{R}, \tag{60}$$

$$e^{iks} = \exp ik\left[R + \frac{(\mathbf{r}_1 - \mathbf{r}_2)^2}{2R}\right],$$

to jest

$$e^{iks} = \exp ik\left(R + \frac{r_1^2}{2R} + \frac{r_2^2}{2R} - \frac{\mathbf{r}_1 \cdot \mathbf{r}_2}{R}\right).$$

Wykorzystując współrzędne biegunowe otrzymuje się

$$\mathbf{r}_1 = r_1 \cos\psi_1 \mathbf{1}_{x1} + r_1 \sin\psi_1 \mathbf{1}_{y1},$$
$$\mathbf{r}_2 = r_2 \cos\psi_2 \mathbf{1}_{x2} + r_2 \sin\psi_2 \mathbf{1}_{y2},$$

co daje

$$\mathbf{r}_1 \cdot \mathbf{r}_2 = -r_1 r_2 \cos(\psi_1 + \psi_2)$$

i następnie

$$e^{iks} = \exp ik\left[R + \frac{r_1^2}{2R} + \frac{r_2^2}{2R} + \frac{r_1 r_2}{R}\cos(\psi_1 + \psi_2)\right]. \tag{61}$$

Moce P_1 i P_2 występujące w (56) oblicza się na podstawie wyrażenia

$$P = \frac{1}{2}\mathrm{Re}\int_S (\mathbf{E} \times \mathbf{H}^*) \cdot \mathbf{n} dS = \frac{1}{2}\sqrt{\frac{\varepsilon}{\mu}}\int_S \mathbf{E} \cdot \mathbf{E}^* dS.$$

Po podstawieniu

$$\mathbf{E} = \left[1 - \left(\frac{2r}{D}\right)^2\right]\mathbf{1}_y$$

dostaje się

$$P_1 = P_2 = \frac{1}{2}\sqrt{\frac{\varepsilon}{\mu}}\int_0^{2\pi} d\psi \int_0^{D/2} dr \left[1 - \left(\frac{2r}{D}\right)^2\right]^2 = \frac{\pi D^2}{24}\sqrt{\frac{\varepsilon}{\mu}}. \tag{62}$$

Wykorzystując równania (57)÷(62) w (56) otrzymuje się

$$T = \frac{12ie^{ikR}}{\lambda\pi D^2 R}\int_0^{D/2} dr_1 \int_0^{D/2} dr_2 \int_0^{2\pi} d\psi_1 \int_0^{2\pi} d\psi_2 \left[r_1 r_2 \left(1-\frac{4r_1^2}{D^2}\right)\left(1-\frac{4r_2^2}{D^2}\right) e^{i\varphi}\right], \tag{63}$$

gdzie

$$\varphi = k\left[\frac{r_1^2}{2R} + \frac{r_2^2}{2R} + \frac{r_1 r_2}{R}\cos(\psi_1+\psi_2)\right]. \tag{64}$$

Korzystnie jest wprowadzić znormalizowane zmienne całkowania u, v przyjmując

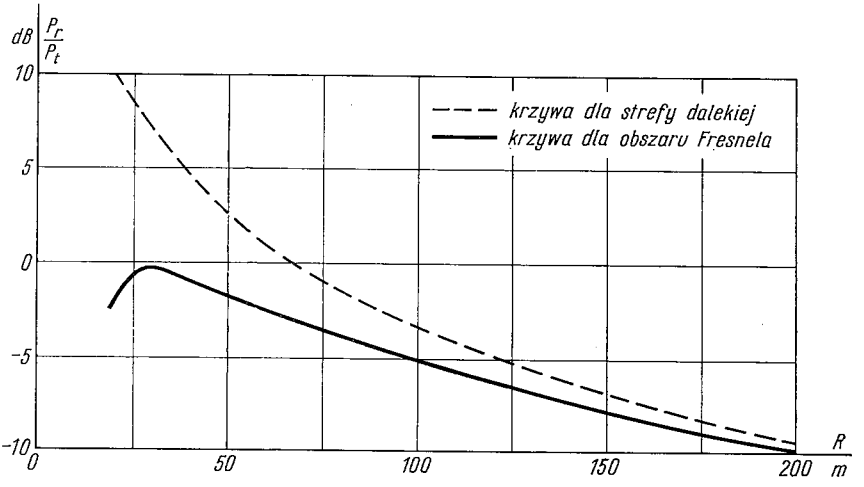
$$r_1 = \frac{Du}{2}; \quad r_1 = \frac{Dv}{2}. \tag{65}$$

Otrzymuje się wtedy równanie transmisji w postaci

$$\frac{P_r}{P_t} = TT^* = \left|\frac{3D^2}{4\pi\lambda R}\int_0^1 du \int_0^1 dv \int_0^{2\pi} d\psi_1 \int_0^{2\pi} d\psi_2 [uv(1-u^2)(1-v^2)e^{i\varphi}]\right|^2, \tag{66}$$

gdzie

$$\varphi = \frac{\pi D^2}{4\lambda R}[u^2+v^2+2uv\cos(\psi_1+\psi_2)]. \tag{67}$$



Rys. 5. Obliczenie P_r/P_t według równań dla strefy dalekiej i dla obszaru Fresnela

Wykonano rachunki podstawiając $D = 100$ cm oraz $\lambda = 8{,}6$ mm dla wartości $R \in (20\text{m}, 200\text{ m})$. Dla anteny o średnicy apertury 100 cm promień R_o strefy dalekiej obliczony według zależności

$$R_o = \frac{2D^2}{\lambda}$$

wynosi około 230 m. Dla porównania obliczono również wartości $\frac{P_r}{P_t}$ według wyrażenia dla strefy dalekiej. Ma ono postać

$$\frac{P_r}{P_t} = \left(\frac{3\pi D^2}{16\lambda R}\right)^2.$$

Wyniki w skali decybelowej pokazane są na rys. 5. Gdy odległość między antenami maleje, wartość stosunku $\frac{P_r}{P_t}$ obliczona wg (66) rośnie. Dla $R = 0{,}14\ R_o$ krzywa przechodzi przez maximum. Jest to najkorzystniejsza dla transmisji odległość między antenami w tym przypadku. Warto podkreślić, że krzywa obliczona według równania (69) dla strefy dalekiej jest w tym obszarze obarczona dużym błędem i nie może być praktycznie wykorzystywana.

9. ZAKOŃCZENIE

Podział pola anteny nadawczej na dwa składniki o odpowiednio dobranej polaryzacji umożliwił określenie i obliczenie współczynnika transmisji, a następnie przedstawienie w prostej postaci równania transmisji. Równanie obowiązuje dla wszystkich typów anten. Trudności rachunkowe zmuszają jednak do uproszczeń teorii: zakłada się, że odległość między antenami pozostaje na tyle duża, aby można było pominąć oddziaływanie wzajemne anten, dopuszcza się przybliżone wartości pola na powierzchniach otaczających anteny (apertury, powierzchnie reflektorów), pomija się wreszcie wpływ otoczenia. Okoliczności te sprawiają, że w niektórych przypadkach konieczna jest konfrontacja obliczeń z eksperymentem.

Mimo tych niedostatków wartość teorii jest duża. Pozwala ona znacznie dokładniej prześledzić wpływ odległości na poziom transmisji niż uproszczona formuła Friisa obowiązująca tylko dla strefy dalekiej. Ma to wielkie znaczenie w miernictwie charakterystyk promieniowania anten, gdzie często powstają wątpliwości, czy obrana odległość między antenami jest wystarczająca. Należy jednak podkreślić, że do proponowanych obliczeń konieczna jest maszyna cyfrowa pozwalająca na rachunek całek wielokrotnych.

Dodatek

POLA WEKTOROWE O ZGODNEJ TOPOGRAFII

Niech

$$\mathbf{A} = \mathbf{A}_1 + i\mathbf{A}_2,$$

gdzie $\mathbf{A}_1$, $\mathbf{A}_2$ są wielkościami rzeczywistymi, oznacza zespolone pole wektorowe określone w pewnym obszarze Ω.

Polem o zgodnej topografii z polem $\mathbf{A}$ będziemy nazywać każde pole $\mathbf{A}'$ o postaci

$$\mathbf{A}' = \alpha\mathbf{A} \quad \text{lub} \quad A' = \alpha\mathbf{A}^*,$$

gdzie $\alpha = \text{const} \neq 0$ jest dowolnym rzeczywistym współczynnikiem, zaś gwiazdka oznacza wielkość sprzężoną.

Biorąc pod uwagę powyższą definicję można konstruować również pole elektromagnetyczne o zgodnej topografii z danym polem **E**, **H**. Należy jednak zawsze sprawdzić, czy otrzymane nowe pole **E'**, **H'** spełnia równania Maxwella

$$\nabla \times \mathbf{H} = -i\omega\varepsilon \mathbf{E},$$
$$\nabla \times \mathbf{E} = i\omega\mu \mathbf{H}.$$

Tak np. pole $\mathbf{E}^*$, **H** mimo zgodności topografii z polem **E**, **H** nie spełnia układu Maxwella, nie jest więc polem elektromagnetycznym. Natomiast pole $-\mathbf{E}^*$, $\mathbf{H}^*$ ma zgodną topografię z polem **E**, **H** i jest polem elektromagnetycznym.

Pola elektromagnetyczne o zgodnej topografii cechuje również zgodność topografii uśrednionego w czasie wektora Poyntinga

$$\mathbf{S} = \frac{1}{2} \operatorname{Re}(\mathbf{E} \times \mathbf{H}^*).$$

Wymienione w charakterze przykładu pole $-\mathbf{E}^*$, $\mathbf{H}^*$ ma ciekawą własność odwracania kierunku uśrednionego wektora Poyntinga, tj. odwracania kierunku przepływu energii czynnej.

LITERATURA CYTOWANA W TEKŚCIE

1. M i n g-K u e i-H u, Near Zone Power Transmission Formulas, IRE National Conv. Record, 1958, Part C, s. 128—138.
2. J. R o b i e u x, Lois générales de la liaison entre radiateurs d'ondes, Annales de Radioélectricité, I. 1959 t. XIV, s. 187—229, II. 1960, t. XV, s. 28—77, III. 1960, t. XV, s. 331—390.
3. B. E. K i n b e r, K tieorii prijemnoj antienny, Radiotiechnika i elektronika, 1961, t. VI, s. 651—653.
4. E. J a c o b s, Fresnel Region Power Transfer, Electromagnetic Theory and Antennas — Proceedings of a symposium held at Copenhagen, June 1962, s. 1051—1074.
5. B. E. K i n b e r, W. B. C e j t l i n, O pogresznosti izmierenija koefficjenta naprawlennogo diejstwija i diagrammy naprawlennosti antien na bliskich rasstojanijach, Radiotiechnika i elektronika, 1964, t. IX, s. 1584—1593.
6. J. R. P a c e, Aysmptotic Formulus for Coupling Between Two Antennas in the Fresnel Region, IEEE Trans. on Antennas and Propagation, 1969, vol. AP-17, s. 285—291.
7. Z. C z y ż, Amplitudowe i mocowe przedstawienia parametrów anten i fal spolaryzowanych eliptycznie, Prace PIT, nr 63/1969, s. 11—22.
8. N. P. B u s l e n k o i inni, Metoda Monte Carlo, PWN, Warszawa 1967.

S. POGORZELSKI

POWER TRANSMISSION BETWEEN ANTENNAS IN FRESNEL REGION

S u m m a r y

There are two antennas considered, a transmitting and a receiving one, both of arbitrary structure and polarization, located in empty space at a certain distance to each other. The transmitting antenna field will be divided in two components:

$$\mathbf{E}_1 = -T\sqrt{\frac{P_1}{P_2}}\mathbf{E}_2^* + \mathbf{E}_d,$$

$$\mathbf{H}_1 = T\sqrt{\frac{P_1}{P_2}}\mathbf{H}_2^* + \mathbf{H}_d$$

The field topography of the first component is consistent with the topography of the field $\mathbf{E}_2$, $\mathbf{H}_2$ which would be radiated by the receiving antenna. The first component accounts for the received power. The second component $\mathbf{E}_d$, $\mathbf{H}_d$, orthogonal to $\mathbf{E}_2$, $\mathbf{H}_2$, accounts for the non received power. The complex coeffi-

cient $T = \text{const}$ will be called transmission factor. The factor $\sqrt{\frac{P_1}{P_2}}$ is introduced to make the value of T independent from the power levels P_1, P_2 radiated by the antennas.

Integral representations for T convenient for numerical computations are established. The transmission factor T enables to become a power transmission formula of a simple form

$$\frac{P_r}{P_t} = TT^*.$$

P_r and P_t represent received and transmitted powers correspondingly. There are two cases discussed. In the first case both antennas are situated in their far zones respectively. This yields the simple Friis formula. In the second case both antennas are situated in Fresnel regions of each other. Then the transmission formula becomes much more complicated, suitable for computer calculations only.

For illustration a particular example of two identical parabolic antennas with consistent polarizations is considered. Power transmission level P_r/P_t versus antenna distance is calculated.

S. POGORZELSKI

GLEICHUNG FÜR LEISTUNGSÜBERTRAGUNG ZWISCHEN ANTENNEN IM FRESNELGEBIET

Zusammenfassung

Zwei Antennen: eine Sende- und eine Empfangsantenne werden in Erwägung genommen, beide beliebiger Struktur und Polarisation, beide in leeren Raum in einer gewissen Entfernung von einander gesetzt. Das Feld der Sendeantenne $\mathbf{E}_1, \mathbf{H}_1$ wird in zwei Komponent gespaltet:

$$\mathbf{E}_1 = -T\sqrt{\frac{P_1}{P_2}}\,\mathbf{E}_2^* + \mathbf{E}_d,$$

$$\mathbf{H}_1 = T\sqrt{\frac{P_1}{P_2}}\,\mathbf{H}_2^* + \mathbf{H}_d.$$

Die Topographie des Feldes der ersten Komponente stimmt mit der Topographie des Feldes $\mathbf{E}_2, \mathbf{H}_2$ überein, das von der Empfangsantenne bestrahlt werden würde. Die erste Komponente stellt die empfangene Leistung, die zweite Komponente $\mathbf{E}_d$, $\mathbf{H}_d$, orthogonal zu $\mathbf{E}_2, \mathbf{H}_2$, die Streuleistung dar. Der komplexe Koeffizient $T = \text{const}$ wurde Übertragungsfaktor benannt. Der Faktor $\sqrt{\frac{P_1}{P_2}}$ wurde eingeführt, um den Wert von T von den Leistungsniveaus P_1, P_2 unabhänging zu machen, die entsprechend der ersten und der zweiten Antenne zugeordnet worden sind.

Man erhält Integralausdrücke für T, die für numerische Berechnungen geeignet sind. Der Übertragungsfaktor T gestattet es, die Gleichung der Leistungsübertragung in einer einfachen Form zu erhalten:

$$\frac{P_r}{P_t} = TT^*;$$

P_r und P_t stehen entsprechend für empfangene und ausgestrahlte Leistungen.

Zwei Fälle werden diskutiert. Im ersten Fall befinden sich beide Antennen in ihren entlegenen Fernzonen. Daraus ergibt sich die einfache Übertragungsformel von Friis. Im zweiten Fall wurden die Antennen im Fresnelbereich der zweiten Antenne untergebracht. Nun wird die Übertragungsformel komplizierter, geeignet nur für maschinelle Berechnungen.

Als Beispiel wird der Sonderfall von zwei identischen parabolischen Antennen mit gleichen Polarisationen betrachtet. Die Übertragungsleistung $\frac{P_r}{P_t}$ wird als Funktion der Entfernung beider Antennen berechnet.

С. ПОГОЖЕЛЬСКИ

УРАВНЕНИЕ ТРАНСМИССИИ МЕЖДУ АНТЕННАМИ В ОБЛАСТИ ФРЕСНЕЛЯ

Р е з ю м е

Рассматриваются две антенны: передаточная и приемная произвольной структуры и поляризации установленные в пустоте в некотором расстоянии от себя. Поле передаточной антенны разлагается на две составляющие:

$$\mathbf{E}_1 = -T\sqrt{\frac{P_1}{P_2}}\,\mathbf{E}_2^* + \mathbf{E}_d,$$

$$\mathbf{H}_1 = T\sqrt{\frac{P_1}{P_2}}\,H_2^* + \mathbf{H}_d.$$

Топография первой составляющей согласна с полем $\mathbf{E}_2$, $\mathbf{H}_2$, которое излучала бы приемная антенна. Эта составляющая представляет принимаемую энергию. Вторая составляющая в квадратуре по отношению к полю $\mathbf{E}_2$, $\mathbf{H}_2$ — она представляет непринимаемую энергию. Комплексный коэффициент $T = \text{const}$ называется коэффициентом трансмиссии. Член $\sqrt{\frac{P_1}{P_2}}$ обеспечивает независимость значения T от уровней мощностей P_1, P_2 излучаемых соответственно обеими антеннами.

Выведена интегральная формула для определения T удобная для численных (нумерических) расчетов. Коэффициент трансмиссии T позволяет свести уравнение t к простой форме:

$$\frac{P_r}{P_t} = TT^*;$$

P_r и P_t представляют соответственно принятую и излученную мощности. Продискутированы два случая. В первом — каждая из антенн расположена в далекой зоне второй антенны. В этом случае получается простая формула Фрииса. Во втором случае каждая из антенн находится в области Френеля второй антенны. В этом случае уравнение трансмиссии усложняется и может быть решено исключительно при помощи ЭЦВМ.

В качестве примера рассмотрен частный случай двух одинаковых параболических антенн с согласованными поляризациями. Расчитан уровень трансмиссии $\frac{P_r}{P_t}$ в функции расстояния между антеннами.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 1, ss. 123-141

621.317.76

Impulsowe detektory częstotliwości

ROMUALD NOWAK (WARSZAWA)

Instytut Radioelektroniki Politechniki Warszawskiej

Otrzymano 16.6.1973

W artykule omówiono fizyczne podstawy działania oraz charakterystyczne cechy impulsowych detektorów częstotliwości. Do detektorów impulsowych zaliczono układy, w których sygnał ciągły zostaje przetworzony na sygnał impulsowy w celu ograniczenia wpływu modulacji amplitudy i uproszczenia odtwarzania sygnału modulującego. Przetwarzanie takie umożliwia wyeliminowanie indukcyjności z układów detekcyjnych i zrealizowanie tych układów w wersji całkowicie scalonej. Podano podstawowe zależności opisujące własności detektorów i parametry ich podzespołów oraz przeprowadzono porównanie omawianych układów. Szczególną uwagę zwrócono na omówienie zjawisk fizycznych decydujących o tłumieniu modulacji amplitudy.

1. WSTĘP

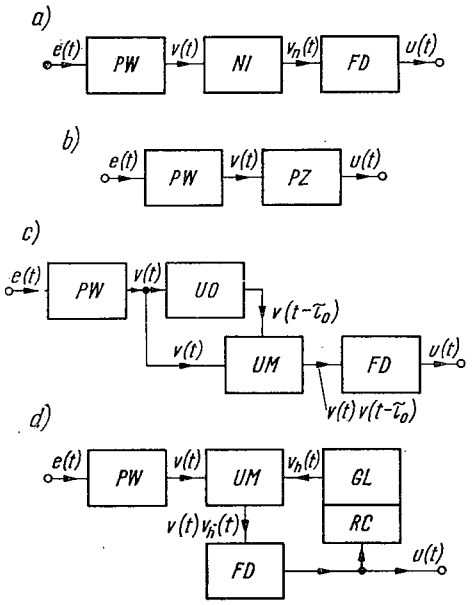
Do impulsowych detektorów częstotliwości zaliczono te układy detekcyjne, w których, w celu ograniczenia wpływu modulacji amplitudy i odtworzenia sygnału modulującego, realizowane jest przekształcenie sygnału ciągłego (analogowego) na sygnał impulsowy. Odpowiednie wykorzystanie własności sygnału impulsowego umożliwia osiągnięcie obu tych celów bez stosowania indukcyjności w obwodach detektora. Impulsowe detektory częstotliwości odznaczają się następującymi zaletami:

— małe zniekształcenia nieliniowe w szerokim zakresie dewiacji częstotliwości,
— skuteczne ograniczenie wpływu modulacji amplitudy,
— możliwość detekcji sygnałów o maksymalnej dewiacji tylko nieznacznie mniejszej od częstotliwości nośnej,
— znaczne uproszczenie operacji uruchamiania, dzięki wyeliminowaniu procesów strojenia obwodów,
— dostosowanie schematów do wymagań technologii układów scalonych.

Osiągnięcie tych zalet stało się możliwe kosztem odpowiedniej rozbudowy układu w porównaniu do detektorów analogowych. Wada ta traci stopniowo na znaczeniu w miarę rozwoju techniki i technologii układów scalonych. Powyższe okoliczności uzasadniają obserwowany obecnie wzrost zainteresowania techniką impulsową w dziedzinie detekcji sygnałów o modulacji kątowej.

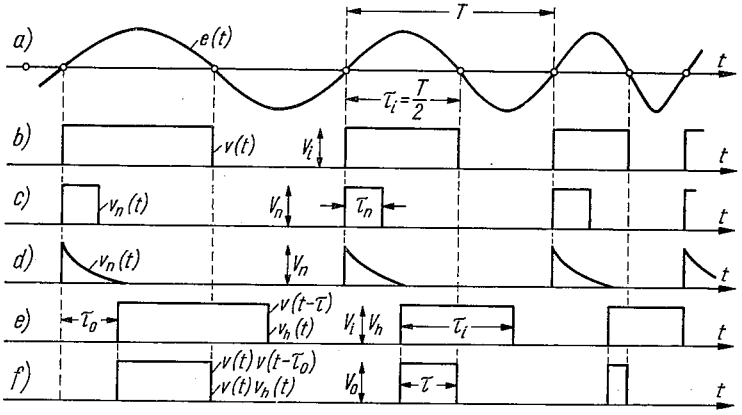
Znane są cztery podstawowe typy impulsowych detektorów częstotliwości, dla których przyjęto następujące nazwy robocze:

— detektor normalizujący impulsy,
— detektor zliczający impulsy,



Rys. 1. Schematy blokowe impulsowych detektorów częstotliwości: a) detektor normalizujący impulsy, b) detektor zliczający impulsy, c) detektor autokorelacyjny, d) detektor synchronizowany

PW — przetwornik wejściowy, *NI* — układ normalizacji impulsów, *FD* — filtr dolnoprzepustowy, *PZ* — układ diodowo-pojemnościowy, *UO* — układ opóźniający, *UM* — układ mnożący, *GL* — generator lokalny, *RC* — regulator częstotliwości

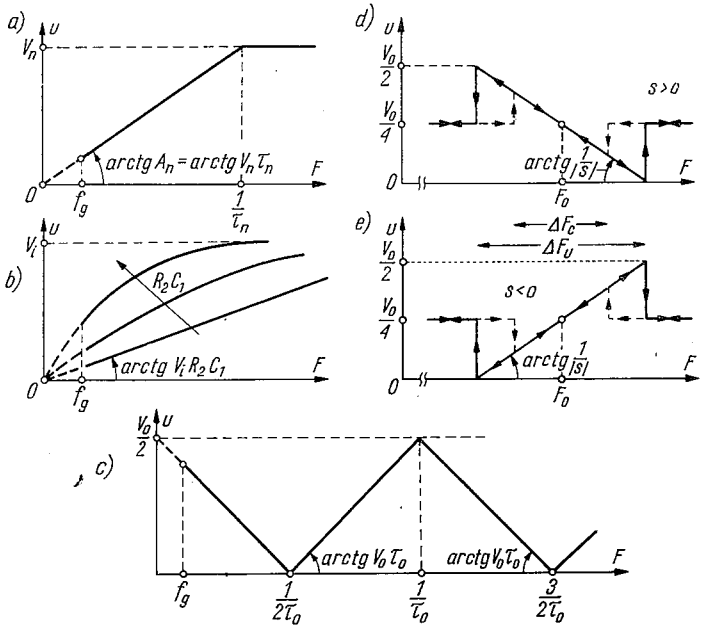


Rys. 2. Uproszczone przebiegi napięć w detektorach impulsowych przy zerowym poziomie odniesienia: a) napięcie wejściowe, b) sygnał impulsowy na wyjściu przetwornika *PW*, c) ciąg impulsów znormalizowanych przy formowaniu aktywnym, d) ciąg impulsów znormalizowanych przy formowaniu biernym, e) opóźniony sygnał impulsowy lub napięcie generatora lokalnego, f) napięcie wyjściowe układu mnożącego

— detektor autokorelacyjny,
— detektor synchronizowany.

Schematy blokowe detektorów impulsowych są przedstawione na rys. 1, a odpowiednie przebiegi napięć — na rys. 2. Na rys. 3 zestawiono uproszczone kształty charakterystyk statycznych detekcji $u(F)$ dla poszczególnych detektorów, gdzie u i F oznaczają odpowiednio chwilowe wartości napięcia wyjściowego i częstotliwości napięcia wejściowego detektora.

Każdy z detektorów zawiera przetwornik wejściowy PW, przetwarzający sygnał ciągły na impulsowy. Układy detektorów różnią się sposobem odtwarzania sygnału modulującego.



Rys. 3. Uproszczone charakterystyki statyczne detektorów impulsowych: a) detektora normalizującego przy normalizacji aktywnej, b) detektora zliczającego przy różnych wartościach stałej czasu R_2C_1, c) detektora autokorelacyjnego, d) i e) detektora synchronizowanego przy różnych znakach nachylenia charakterystyki przestrajania $\left(\Delta U > \frac{V_0}{2}\right)$

2. PRZETWARZANIE SYGNAŁU CIĄGŁEGO NA SYGNAŁ IMPULSOWY

Celem przetwarzania jest ograniczenie wpływu zmian amplitudy sygnału wejściowego detektora oraz uproszczenie operacji odtwarzania sygnału modulującego. W procesie przetwarzania ciągły (analogowy) sygnał wąskopasmowy[1] $e(t)$ zostaje przekształcony na sygnał impulsowy $v(t)$, zgodnie z następującą zależnością:

$$v(t) = \frac{1}{2} V_i\{\text{sign}[e(t) - E_0 + 1\} + v_p, \tag{1}$$

[1] Dla sygnału wąskopasmowego wartość względnej, skutecznej szerokości pasma jest znacznie mniejsza od jedności. Warunek wąskopasmowości sygnału wejściowego jest warunkiem koniecznym tylko dla detektora synchronizowanego.

gdzie:

V_i — wartość międzyszczytowa impulsów wyjściowych,

E_0 — poziom odniesienia dla sygnału wejściowego,

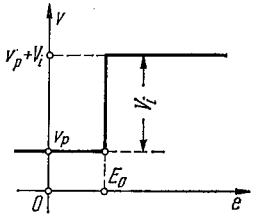
v_p — stały poziom tzw. piedestału,

sign(x)— znak wielkości x, tzn:

$$\text{sign}(x) = \begin{cases} 1 & \text{gdy } x > 0, \\ 0 & \text{gdy } x = 0, \\ -1 & \text{gdy } x < 0. \end{cases} \tag{2}$$

Zależność (1) jest równaniem charakterystyki statycznej idealnego przetwornika, przedstawionej na rys. 4.

Dla uproszczenia dalszych rozważań przyjęto zerowy poziom piedestału ($v_p = 0$). Założenie takie nie ogranicza ich ogólności, ponieważ poziom piedestału wpływa tylko na wartość składowej stałej sygnałów.



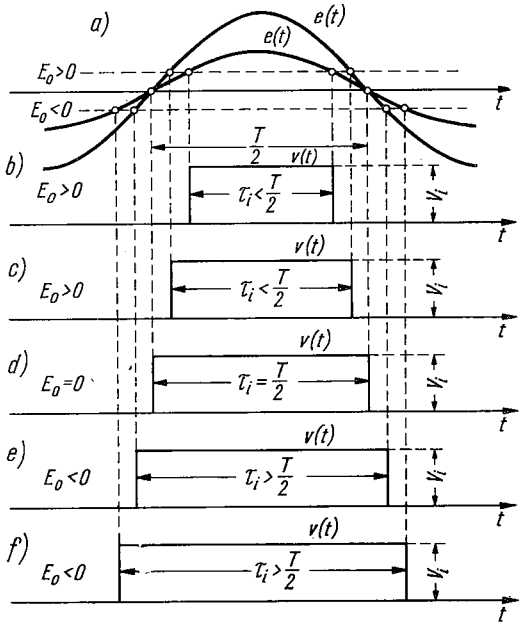
Rys. 4. Charakterystyka statyczna idealnego przetwornika wejściowego *PW*

Proces przetwarzania ilustruje rys. 2a i b. Przy powyższych założeniach sygnał wyjściowy $v(t)$ przetwornika PW jest ciągiem jednoimiennych, prostokątnych impulsów o stałej wartości międzyszczytowej V_i (równej wysokości impulsów). Czoła impulsów tworzą sygnał o szczególnej[2)] modulacji położenia impulsów. Odstęp czasu $T = \frac{1}{F}$ pomiędzy impulsami równa się okresowi napięcia wejściowego i odwrotności częstotliwości średniej F tego napięcia (wartość średnia za odstęp czasu T). Czas trwania (długość) impulsu τ_i zależy od okresu T i poziomu odniesienia E_0, natomiast współczynnik wypełnienia $q = \frac{\tau_i}{T}$ określony jest tylko przez poziom odniesienia. Wpływ poziomu odniesienia i amplitudy napięcia wejściowego na kształt impulsów wyjściowych przetwornika ilustruje rys. 5. Jak wynika z tego rysunku i dalszych rozważań, optymalnym poziomem odniesienia jest poziom zerowy. Przy takim poziomie odniesienia kształt impulsów nie zależy od amplitudy sygnału ciągłego, a sygnał impulsowy charakteryzuje się liniową modulacją długości impulsów $\left(\tau_i = \frac{T}{2}\right)$.

Przetwarzanie sygnału analogowego na sygnał impulsowy może być zrealizowane w dwojaki sposób. Sposób pierwszy — bierny — polega na nieliniowym wzmacnianiu sygnału ciągłego z dwustronnym (przy zerowym poziomie odniesienia — symetrycznym) ograni-

[2)] Przy $E_0 = 0$ czoła impulsów występują w chwilach, gdy chwilowa wartość fazy sygnału wąskopasmowego $e(t)$ przekracza poziomy odpowiadające kolejnym wielokrotnościom 2π rad.

czaniem poziomów. Sposób drugi — aktywny lub regeneracyjny — wykorzystuje do formowania impulsów procesy regeneracyjne zachodzące w układach o dodatnim sprzężeniu zwrotnym lub o ujemnej rezystancji (np. w przerzutniku Schmitta). Przy przetwarzaniu biernym położenie obu zboczy impulsów wyjściowych nie zależy od parametrów układu formującego, natomiast kształt zboczy zależy od chwilowych wartości sygnału wejściowego w otoczeniu poziomu odniesienia, a więc i od chwilowej wartości amplitudy tego sygnału.



Rys. 5. Kształt impulsów wyjściowych przetwornika *PW* przy różnych poziomach odniesienia i różnych amplitudach napięcia wejściowego: a) napięcia wejściowe, b) i c) impulsy wyjściowe przy $E_o > 0$, d) impuls wyjściowy przy $E_o = 0$, e) i f) impulsy wyjściowe przy $E_o < 0$

Przy przetwarzaniu aktywnym o stałym, zerowym poziomie odniesienia, kształt impulsów wyjściowych nie zależy od wartości chwilowych napięcia wejściowego i jego amplitudy. Jest on całkowicie określony przez parametry przerzutnika formującego i odstęp czasu pomiędzy różnoimiennymi przejściami przez zero wartości chwilowej sygnału analogowego. W realnych układach przerzutników formujących, poziom odniesienia nie jest dokładnie równy zeru, a ponadto jego wartość zależy od znaku pochodnej czasowej sygnału ciągłego (efekt histerezy). Uzależnia to w pewnym stopniu położenie obu zboczy impulsów wyjściowych od chwilowej wartości amplitudy sygnału wejściowego.

W przypadku idealnego przetwarzania (określonego zależnością (1)) o stałym zerowym poziomie odniesienia ($E_0 = 0$) i stałej wartości międzyszczytowej impulsów ($V_i = \text{const.}$), wyjściowy sygnał impulsowy przetwornika PW nie zależy od chwilowej wartości amplitudy sygnału wejściowego. Przetwornik wejściowy PW spełnia w tym przypadku rolę idealnego ogranicznika amplitudy. W rezultacie niedoskonałego działania realnych przetworników wejściowych modulacja amplitudy będzie się w ograniczonym stopniu przenosić na ciąg impulsów wyjściowych, wywołując odpowiednie zmiany napięcia wyjściowego detek-

tora. Jak łatwo zauważyć z powyższych rozważań, optymalne ograniczenie wpływu modulacji amplitudy daje się zrealizować przy odpowiednim wykorzystaniu w obwodach przetwornika wejściowego zalet obu sposobów przetwarzania — biernego i aktywnego. W takich rozwiązaniach przerzutnik formujący przetwornika PW jest poprzedzony nieliniowym wzmacniaczem z dwustronnym, symetrycznym ograniczeniem poziomów, który formuje wstępnie impulsy wyzwalające przerzutnik.

Przetwarzanie sygnału ciągłego na impulsowy, zachodzące w omawianych układach detekcyjnych, wymaga innego niż dla detektorów analogowych sposobu podejścia do określenia tłumienia modulacji amplitudy. W detektorach analogowych z wąskopasmowym ogranicznikiem amplitudy, jako miarę tłumienia modulacji amplitudy przyjmuje się stosunek głębokości modulacji przed i po ograniczeniu. Ten poglądowy sposób ujęcia wymaga jednak określenia wartości amplitudy fali nośnej na wyjściu ogranicznika. W detektorach impulsowych o nieliniowym, szerokopasmowym przetwarzaniu nie można dokonać odpowiedniego pomiaru. Przyjęcie natomiast amplitudy fali nośnej na wejściu przetwornika PW, jako podstawy do określenia tłumienia, wymagałoby wyznaczenia wzmocnienia napięciowego przetwornika, co w przypadku przetwarzania aktywnego jest niewykonalne, a w przypadku przetwarzania biernego — bardzo utrudnione. Wygodną miarą tłumienia modulacji amplitudy przez detektor impulsowy może być np. stosunek amplitudy napięcia wyjściowego detektora przy nominalnej dewiacji częstotliwości i połowy wartości międzyszczytowej[3] tego napięcia przy sinusoidalnej modulacji amplitudy napięcia wejściowego. Tak określone tłumienie zależy od głębokości modulacji amplitudy i poziomu sygnału wejściowego.

Proces przetwarzania jest operacją nieliniową. W procesie tym oprócz ograniczenia wpływu niepożądanej modulacji amplitudy sygnału wejściowego zachodzi jednocześnie próbkowanie sygnału ze zmiennym okresem próbkowania T. Jak wspomniano wyżej, okres T jest równy odstępowi czasu pomiędzy jednoimiennymi przejściami przez zero napięcia wejściowego i w ten sposób uzależniony od chwilowej wartości sygnału modulującego. Wartość częstotliwości $F = \frac{1}{T}$ jest więc wartością średnią próbki, odpowiadającą średniej wartości próbki sygnału modulującego. Dla sygnałów wąskopasmowych można z dobrym przybliżeniem traktować tak określoną wartość średnią częstotliwości jako częstotliwość chwilową sygnału zmodulowanego.

W wyniku próbkowania sygnał wąskopasmowy, o liniowej modulacji częstotliwości, zostaje przetworzony na szerokopasmowy sygnał o złożonej i w przypadku ogólnym nieliniowej modulacji impulsowej. W widmie takiego sygnału impulsowego występują bezpośrednio składowe sygnału modulującego, co umożliwia w zasadzie wydzielenie tego sygnału za pomocą obwodu liniowego w postaci filtru dolnoprzepustowego (selekcja częstotliwościowa sygnałów). Jednakże z uwagi na mały poziom i duże zniekształcenia nieliniowe składowych sygnału modulującego, zawartych w sygnale impulsowym $v(t)$, przed wydzieleniem sygnału modulującego należy przeprowadzić dodatkowe operacje w celu usunięcia powyższych niedogodności. Operacje takie polegają na przetworzeniu sygnału o złożonej, wieloparametrowej modulacji impulsowej na sygnał o modulacji prostszej — w przypadku ide-

[3] Napięcie wyjściowe detektora impulsowego, wywołane przez niepożądaną, sinusoidalną modulację amplitudy, jest zwykle silnie odkształcone.

alnym jednoparametrowej. W detektorze normalizującym, z uwagi na standaryzację kształtu impulsów (rys. 2c i d), będzie to przejście do „czystej" jednoparametrowej modulacji położenia impulsów, natomiast w detektorach autokorelacyjnym i synchronizowanym, w wyniku mnożenia dwóch przebiegów prostokątnych, następuje „pogłębienie" modulacji długości impulsów (rys. 2e). W detektorze zliczającym do skutecznego wydzielenia składowej modulującej stosuje się specjalny, diodowo-pojemnościowy obwód nieliniowy. Działanie tego obwodu da się w przybliżeniu porównać do działania przetwornika wydzielającego modulację częstotliwości powtarzania impulsów (por. p. 4).

Filtr dolnoprzepustowy FD (rys. 1), służący w detektorze impulsowym do ostatecznego wydzielenia sygnału modulującego, może pełnić jednocześnie rolę filtru korekcyjnego deemfazy. Dotyczy to również układu diodowo-pojemnościowego PZ, pełniącego podobną rolę w detektorze zliczającym.

Przy omawianiu przetwornika wejściowego detektora impulsowego założono (1) przetwarzanie sygnału analogowego na przebieg prostokątny. Możliwe są również realizacje detektorów impulsowych z przetwarzaniem na ciąg impulsów o innym kształcie, np. trójkątnym lub wykładniczym. Przetwarzanie na przebieg prostokątny może być zrealizowane jednak w sposób najbardziej zbliżony do idealnego i jak wyniknie z dalszych rozważań zapewnia optymalne wartości parametrów detektora. Szczególnie ważny jest kształt impulsów w detektorach synchronizowanym i autokorelacyjnym. Dla tych detektorów tylko przebiegi prostokątne zapewniają liniowość detekcji.

3. DETEKTOR NORMALIZUJĄCY IMPULSY

W układzie detektora normalizującego (rys. 1a) sygnał impulsowy $v(t)$ z przetwornika wejściowego PW jest dodatkowo formowany w układzie normalizującym NI. Dzięki normalizacji zostaje wytworzony ciąg impulsów $v_n(t)$ o stałej powierzchni A_n. Czoła impulsów znormalizowanych są zgodne w czasie z czołami ciągu $v(t)$, tworząc sygnał o jednoparametrowej, liniowej modulacji położenia impulsów (rys. rys. 2b, c i d).

Proces normalizacji impulsów, zachodzący w układzie NI, może być — podobnie jak przetwarzanie — procesem aktywnym lub biernym. W przypadku pierwszym przerzutnik monostabilny generuje pod wpływem przednich zboczy sygnału $v(t)$ impulsy prostokątne o stałej wartości międzyszczytowej V_n i stałej długości τ_n (rys. 2c). W przypadku drugim, z ciągu zróżniczkowanych zboczy sygnału $v(t)$ wybiera się impulsy jednoimienne odpowiadające czołom tego sygnału (rys. 2d).

Znormalizowany ciąg impulsów stanowi szczególny przypadek sygnału o liniowej modulacji impulsowej [1]. Widmo takiego sygnału zawiera m.in. pełne widmo niezniekształconego sygnału modulującego. Sygnał modulujący można więc wydzielić za pomocą filtru dolnoprzepustowego FD (rys. 1a). Transmitancja filtru winna zapewnić dostateczne stłumienie składowych, pochodzących od dolnej wstęgi bocznej sygnału wejściowego, a występujących również w widmie ciągu impulsów znormalizowanych. Jeżeli pominąć działanie korekcyjne deemfazy, to warunek na częstotliwość graniczną f_g filtru FD przyjmie postać nierówności:

$$f_m < f_g < F_0 - (m+1)f_m, \tag{3}$$

gdzie:

f_m — górna częstotliwość graniczna przenoszonego pasma sygnału modulującego,

F_0 — częstotliwość nośna sygnału wejściowego,

m — indeks modulacji kąta tego sygnału.

Przyjmując w przybliżeniu, że napięcie wyjściowe detektora u równa się wartości średniej poszczególnych próbek sygnału, tzn. wartości średniej sygnału $v_n(t)$ za odstęp czasu pomiędzy impulsami $T = \frac{1}{F}$, otrzymuje się następujące równanie charakterystyki statycznej detektora normalizującego (rys. 3a):

$$u = \frac{A_n}{T} = A_n F. \tag{4}$$

Równanie powyższe obowiązuje w ograniczonym przedziale częstotliwości F sygnału wejściowego:

$$f_g < F < \frac{1}{\tau_n}. \tag{5}$$

Górna granica tego przedziału odpowiada przypadkowi styczności kolejnych impulsów znormalizowanych, a dla impulsów prostokątnych — maksymalnej możliwej wartości średniej ciągu równej V_n. Dolna granica wynika z wymagań częstotliwościowej selekcji składowych sygnału modulującego.

Charakterystyka dynamiczna detektora, tzn. zależność amplitudy napięcia wyjściowego U od dewiacji częstotliwości ΔF sygnału wejściowego przy sinusoidalnej modulacji, będzie również liniowa, lecz nachylenie jej będzie nieznacznie zależało od częstotliwości modulującej f, nawet przy stałej wartości transmitancji filtru FD [1]:

$$U = S_n(f)\Delta F, \tag{6}$$

gdzie $S_n(f)$ — wartość modułu dwustronnej amplitudowej gęstości widmowej (transformaty Fouriera) pojedynczego impulsu znormalizowanego przy częstotliwości modulującej. Dla znormalizowanego ciągu impulsów prostokątnych:

$$U = A_n \frac{\sin(\pi f \tau_n)}{\pi f \tau_n} \Delta F \cong \mathrm{A}_n \Delta F. \tag{7}$$

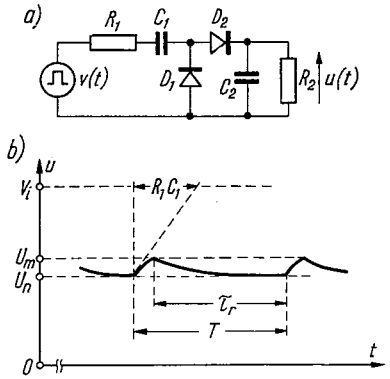
Przybliżenie jest słuszne gdy $2\pi f \tau_n \ll 1$.

Nachylenie obu charakterystyk jest proporcjonalne do powierzchni impulsu znormalizowanego. Jak łatwo zauważyć, maksymalną powierzchnię impulsu zapewnia w danych warunkach impuls prostokątny ($A_n = V_n \tau_n$).

Detektor normalizujący jest niewrażliwy na modulację amplitudy, jeżeli powierzchnia impulsów znormalizowanych nie zależy od chwilowej wartości amplitudy sygnału wejściowego detektora. W związku z powyższym optymalnymi układami formowania impulsów w przetworniku wejściowym i układzie normalizującym dla detektora tego typu są układy aktywne.

4. DETEKTOR ZLICZAJĄCY IMPULSY

W detektorze zliczającym (rys. 1b) sygnał impulsowy $v(t)$ z przetwornika wejściowego PW jest doprowadzony bezpośrednio do wejścia nieliniowego układu diodowo-pojemnościowego PZ. Układ ten, stosowany m.in. w wychyłowych miernikach częstotliwości, przy spełnieniu określonych warunków, reaguje tylko na liczbę impulsów w jednostce czasu, czyli na częstotliwość powtarzania impulsów periodycznych — stąd nazwa detektora. Schemat podstawowego układu diodowo-pojemnościowego przedstawiono na rys. 6a, a przebieg napięcia wyjściowego tego układu, które jest jednocześnie napięciem wyjściowym detektora, na rys. 6b [2] [6].



Rys. 6. Podstawowy schemat układu diodowo-pojemnościowego *PZ* (a) i przebieg chwilowej wartości napięcia wyjściowego (b)

Zasada działania układu diodowo-pojemnościowego polega na impulsowym doładowywaniu przez obwód nieliniowy pojemności wyjściowej C_2. Wartości elementów dobiera się w taki sposób, aby były spełnione warunki:

$$C_1 \ll C_2, \quad FR_2C_1 \ll 1, \quad FR_2C_2 \gg 1, \quad R_1C_1 \ll R_2C_2. \tag{8}$$

Przy dodatnim skoku napięcia wejściowego pojemności C_1 i C_2 są szybko doładowywane przez rezystancję wewnętrzną źródła R_1 i diodę D_2, ze stałą czasu równą w przybliżeniu R_1C_1. Obie pojemności zyskują w ten sposób prawie jednakowe ładunki o wartości $C_1(V_i - U_m)$, gdzie U_m oznacza maksymalną wartość chwilową napięcia wyjściowego detektora (rys. 6b). W przerwie między impulsami wejściowymi pojemność C_1 jest całkowicie rozładowywana przez rezystancję R_1 i diodę D_1. Ładunek tracony przez pojemność C_2, w czasie rozładowywania τ_r poprzez rezystancję obciążenia R_2, ma wartość określoną przez następujące zależności:

$$(U_m - U_n)C_2 = U_mC_2\left[1 - \exp\left(-\frac{\tau_r}{R_2C_2}\right)\right] \cong$$

$$\cong U_mC_2\left[1 - \exp\left(-\frac{1}{FR_2C_2}\right)\right] \cong \frac{U_m}{FR_2}, \tag{9}$$

gdzie U_n oznacza minimalną wartość napięcia wyjściowego (rys. 6b). W równaniu (9) przyjęto, wynikające z warunków (8), założenie upraszczające: $\tau_r \cong T = \frac{1}{F}$.

Pomijając tętnienia napięcia wyjściowego, tzn. zakładając $u \cong U_m$, otrzymuje się z bilansu ładunków następujące, przybliżone równanie charakterystyki statycznej detektora zliczającego (rys. 3b)[4]:

$$u \cong \frac{VR_2C_1F}{1+R_2C_1F} \cong V_iR_2C_1F, \tag{10}$$

słuszne w przedziale częstotliwości: $\frac{1}{R_2C_2} \ll F \ll \frac{1}{R_1C_1}$.

Jak wynika z tego równania, napięcie wyjściowe detektora zliczającego zależy tylko od wysokości impulsów V_i i częstotliwości F, nie zależy natomiast, przy spełnieniu omówionych wyżej warunków, od kształtu impulsów. W detektorze zliczającym, oprócz pełnego ograniczania wpływu modulacji amplitudy związanego z przetwarzaniem sygnału ciągłego na impulsowy, zmniejsza się ten wpływ dodatkowo, wykorzystując omówione wyżej właściwości nieliniowego układu wydzielania składowej modulującej. Resztkowy wpływ modulacji amplitudy może uzewnętrznić się jedynie przez zmianę czasu rozładowania τ_r (niepożądana modulacja położenia impulsów pochodząca od zmian amplitudy sygnału wejściowego). Z tego względu zalecanym sposobem przetwarzania sygnału ciągłego na impulsowy jest dla detektora zliczającego przetwarzanie bierne.

5. DETEKTOR AUTOKORELACYJNY

W detektorze autokorelacyjnym następuje mnożenie dwóch identycznych, lecz przesuniętych w czasie o odstęp τ_o, sygnałów impulsowych, powstałych w wyniku przetworzenia tego samego sygnału analogowego. Do wejść układu mnożącego UM (rys. 1c) dołączone są napięcia: wyjściowe przetwornika PW — $v(t)$ i wyjściowe układu opóźniającego UO — $v(t-\tau_o)$. Przebiegi wejściowe i wyjściowy układu mnożącego przedstawiają odpowiednio rysunki 2b, e i f.

Przebieg wyjściowy układu mnożącego $v(t)v(t-\tau_o)$ jest sygnałem impulsowym o jednoczesnej modulacji okresu i pogłębionej modulacji długości impulsów. Napięcie wyjściowe detektora otrzymuje się po scałkowaniu tego sygnału za pomocą filtru dolnoprzepustowego FD.

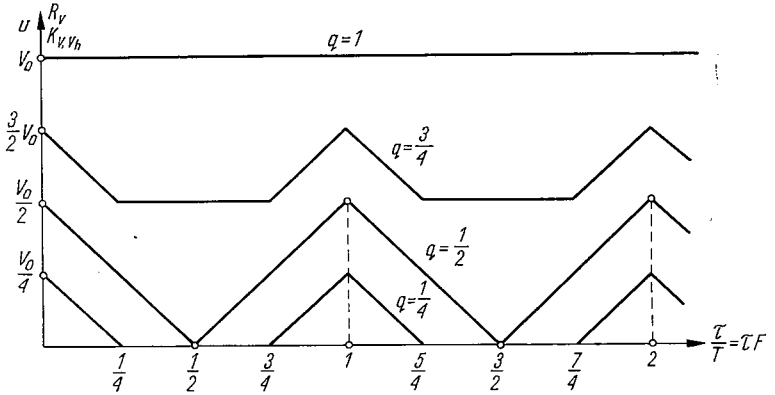
Napięcie wyjściowe detektora autokorelacyjnego jest więc proporcjonalne do wartości funkcji autokorelacji $R_v(\tau)$ sygnału impulsowego $v(t)$ [3]:

$$R_v(\tau) = \lim_{\Theta\to\infty} \frac{1}{\Theta} \int_0^{\Theta} v(t)v(t-\tau_o)dt = \frac{1}{T} \int_0^{T} v(t)v(t-\tau_o)dt. \tag{11}$$

Funkcja autokorelacji takiego przebiegu, przedstawiona na rys. *b* dla różnych wartości współczynnika wypełnienia q, zależy w sposób odcinkowo liniowy od iloczynu $F\tau = \frac{\tau}{T}$,

[4] Przybliżenie słuszne dla $FR_2C_1 \ll 1$.

gdzie τ — opóźnienie między napięciami wejściowymi układu mnożącego. Przy stałej wartości opóźnienia, $\tau = \tau_o = \text{const}$, napięcie wyjściowe detektora autokorelacyjnego zależy liniowo od częstotliwości F sygnału wejściowego. Charakterystyka statyczna detektora, przy współczynniku wypełnienia $q = \frac{1}{2}$, ma postać symetrycznej łamanej (rys. 3c) i jest periodyczną funkcją częstotliwości F. Maksymalne wartości napięcia wyjściowego występują przy wartościach odciętej $\frac{n-1}{\tau_o}$, natomiast wartości minimalne przy odciętej $\frac{2n-1}{\tau_o}$, gdzie n — liczba naturalna. Różnica ekstremalnych wartości napięcia wyjściowego detektora równa się połowie wartości międzyszczytowej V_o impulsów wyjściowych układu mnożącego. (Wartość międzyszczytowa V_o jest proporcjonalna do kwadratu wartości międzyszczytowej V_i impulsów wejściowych układu mnożącego).



Rys. 7. Rodzina funkcji autokorelacji $R_v\left(\frac{\tau}{T}\right)$ przebiegu prostokątnego $v(t)$ lub funkcji korelacji wzajemnej $K_{v,v_h}\left(\frac{\tau}{T}\right)$ dwóch synchronicznych przebiegów prostokątnych $v(t)$ i $v_h(t)$ przy różnych wartościach współczynnika wypełnienia q

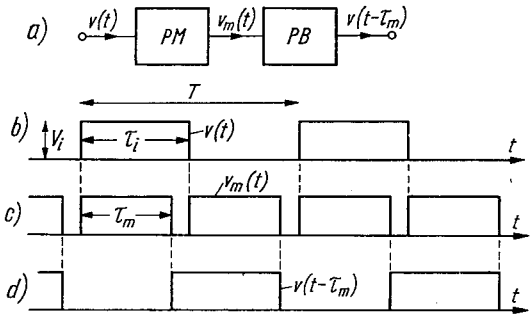
Z uwagi na kształt charakterystyki statycznej detektora, częstotliwość nośna F_o sygnału oraz dewiacja częstotliwości ΔF winny spełniać następujące warunki:

$$F_o \cong \frac{2n-1}{4\tau_o}, \qquad \Delta F \leqslant \frac{1}{4\tau_o}. \tag{12}$$

Ponadto chwilowa wartość częstotliwości sygnału wejściowego w ujemnym szczycie modulacji winna być dostatecznie większa od maksymalnej częstotliwości sygnału modulującego, aby zapewnić prawidłową selekcję częstotliwościową sygnału modulującego przez filtr dolnoprzepustowy FD ($F_o - \Delta F \gg f_m$).

Rolę układu mnożącego dla przebiegów impulsowych prostokątnych dobrze spełnia bramka logiczna typu *and* lub *nand*. Wymagane opóźnienie daje się zrealizować za pomocą odcinka linii opóźniającej lub za pomocą zespołu przerzutników monostabilnych

i bistabilnego. Schemat blokowy układu opóźniającego z przerzutnikami oraz przebiegi napięć w takim układzie przedstawiono na rys. 8. W układzie tym oba zbocza każdego z impulsów sygnału $v(t)$ wyzwalają przerzutnik monostabilny PM, generujący impuls prostokątny o czasie trwania τ_m. Tylne zbocze wytworzonego impulsu odpowiada opóźnionemu w czasie o τ_m zboczu wyzwalającemu. Opóźnione zbocza sygnału impulsowego



Rys. 8. Układ opóźniający *UO* z przerzutnikami: a) schemat blokowy układu, *PM* — przerzutnik mono stabilny, *PB* — przerzutnik bistabilny, *b*) wejściowy sygnał impulsowy, c) napięcie wyjściowe przerzutnika monostabilnego *PM*, d) opóźniony sygnał impulsowy na wyjściu przerzutnika bistabilnego *PB*

pobudzają symetrycznie przerzutnik bistabilny PB, na którego wyjściu zostaje odtworzony opóźniony sygnał impulsowy $v(t-\tau_m)$. Opóźnienie τ_m, wprowadzane przez pojedyńczy przerzutnik monostabilny, nie może przekraczać połowy minimalnego odstępu czasu pomiędzy impulsami sygnału $v(t)$. Wynika stąd warunek:

$$\tau_m > \frac{1}{2(F_o+\Delta F)}.$$

Zwiększając N-krotnie liczbę przerzutników monostabilnych w łańcuchu można zrealizować N-krotnie większe opóźnienie wypadkowe $\tau_o = N\tau_m$.

Wpływ modulacji amplitudy na pracę detektora autokorelacyjnego ograniczony jest jedynie przez przetwarzanie sygnału analogowego na impulsowy. Przy regeneracyjnym sposobie przetwarzania, zmiany amplitudy na wejściu detektora będą powodować tylko niewielkie zmiany współczynnika wypełnienia q ciągu impulsów $v(t)$.

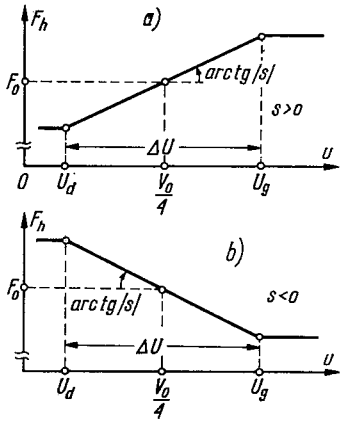
Wpływ zmian współczynnika wypełnienia ilustruje rys. 7. Jak wynika z tego rysunku, przy zmianach współczynnika wypełnienia zmienia się długość liniowego odcinka charakterystyki statycznej i wartość chwilowa napięcia wyjściowego detektora, natomiast nachylenie charakterystyki pozostaje stałe. Własność taka jest równoważna automatycznej regulacji wzmocnienia odbiornika, nie daje jednak dodatkowego ograniczenia wpływu modulacji amplitudy sygnału wejściowego. Przy biernym sposobie przetwarzania zmiany kształtu impulsów wyjściowych przetwornika PW, spowodowane modulacją amplitudy sygnału wejściowego, będą wywoływały nie tylko zmiany nachylenia charakterystyki statycznej detektora, lecz również będą wpływały na liniowość tej charakterystyki (tylko funkcja autokorelacji fali prostokątnej jest funkcją liniową).

Detektor autokorelacyjny może być wykonany w wersji analogowej bez przetwarzania sygnału na sygnał impulsowy. W tym przypadku jednak liniowość charakterystyki sta-

tycznej detektora jest znacznie gorsza (np. funkcją autokorelacji sinusoidy jest również sinusoida). Powszechnie znane detektory analogowe, jak detektor różnicowy czy stosunkowy, są szczególnymi przypadkami detektorów autokorelacyjnych o opóźnieniu równym nachyleniu charakterystyki fazowej obwodu dyskryminatora fazy, czyli równym ćwiartce okresu fali nośnej.

6. DETEKTOR SYNCHRONIZOWANY

W impulsowym detektorze synchronizowanym generator lokalny GL napięcia prostokątnego $v_h(t)$, stanowiący integralną część detektora, jest synchronizowany za pomocą pętli fazowej regulacji częstotliwości przez sygnał impulsowy $v(t)$[5]. Schemat blokowy detektora przedstawia rys. 1d, a odpowiednie przebiegi napięć zamieszczono na rysunkach 2b, e i f. Pętla automatyki składa się z fazowego detektora błędu, którego rolę w omawianym układzie pełni układ iloczynu logicznego UM, z filtru dolnoprzepustowego FD



Rys. 9. Charakterystyki przestrajania $F_h(u)$ generatora lokalnego o dodatnim (a) i ujemnym (b) nachyleniu s.

oraz z regulatora częstotliwości RC, umożliwiającego napięciowe przestrajanie generatora lokalnego GL[6]. Napięcie wyjściowe u detektora synchronizowanego, występujące na wyjściu filtru FD, jest jednocześnie napięciem błędu pętli automatyki. Służy ono do automatycznego dostarajania generatora lokalnego, zgodnie z charakterystyką przestrajania $F_h(u)$, gdzie F_h — częstotliwość powtarzania impulsów napięcia $v_h(t)$. Przykłady charakterystyki przestrajania zamieszczono na rys. 9. Dla uproszczenia dalszych rozważań założono, że moduł transmitancji napięciowej filtru FD w paśmie przepustowym jest równy jedności.

Warunki pracy pętli fazowej regulacji częstotliwości są dobierane w taki sposób, aby w całym zakresie zmian częstotliwości chwilowej sygnału wymuszany był warunek synchronizmu w postaci:

$$F(t) = F_o + \delta F(t) = F_h(t), \tag{13}$$

[5] Z tego powodu detektor synchronizowany jest często w literaturze nazywany detektorem z pętlą fazowej regulacji częstotliwości.

[6] Jako generator lokalny z regulatorem częstotliwości może w detektorze impulsowym służyć np. przestrajany napięciowo multiwibrator.

gdzie $\delta F(t) = F(t) - F_o$ oznacza chwilową wartość odchylenia częstotliwości sygnału F od częstotliwości nośnej F_o, proporcjonalną do chwilowej wartości sygnału modulującego[7]. Zależność (1*d*) jest spełniona dla odchyleń częstotliwości $\delta F(t)$ o widmie przenoszonym przez filtr FD, tylko bowiem w takich warunkach częstotliwość chwilowa generatora lokalnego $F_h(t)$ nadąża za zmianami częstotliwości sygnału wejściowego. W zakresie liniowej części charakterystyki przestrajania $F_h(u)$ napięcie wyjściowe detektora będzie liniową funkcją chwilowej wartości sygnału modulującego.

Detektor synchronizowany ma wiele cech wspólnych z detektorem autokorelacyjnym. Jak wynika z rys. 1d napięcie wyjściowe detektora synchronizowanego jest liniowo zależne od wartości funkcji korelacji wzajemnej $K_{v,v_h}(\tau)$ dwóch przebiegów prostokątnych: $v(t)$ i $v_h(t) = v(t-\tau)$, gdzie τ — przesunięcie czasowe obu przebiegów[8]

$$K_{v,v_h}(\tau) = \lim_{\Theta\to\infty} \frac{1}{\Theta} \int_0^{\Theta} v(t)\cdot v_h(t)dt = \frac{1}{T}\int_0^{T} v(t)\cdot v(t-\tau)dt. \tag{14}$$

Funkcja korelacji wzajemnej $K_{v,v_h}(\tau)$ ma więc taki sam kształt jak funkcja autokorelacji $R_v(\tau)$ określona równaniem (11) i przedstawiona na rys. 7. Można wykazać, że w stanie synchronizmu przy liniowej charakterystyce przestrajania $F_h(u)$ pętla automatyki wymusza stałe przesunięcie w czasie ($\tau = \tau_o = \text{const}$) pomiędzy ciągami impulsów $v(t)$ i $v_h(t)$. Wartość tego opóźnienia można wyznaczyć z warunku:

$$\tau_o |s| V_o = 1, \tag{15}$$

gdzie $s = \dfrac{dF_h(u)}{du}$ — nachylenie charakterystyki przestrajania.

Detektory autokorelacyjny i synchronizowany różnią się więc tylko sposobem stabilizacji opóźnienia τ_o. W detektorze autokorelacyjnym stałość opóźnienia jest zapewniona przez parametry układu opóźniającego UO, natomiast w detektorze synchronizowanym opóźnienie będzie stałe przy liniowej charakterystyce przestrajania $F_h(u)$ i spełnieniu warunku synchronizmu (13).

W detektorze synchronizowanym synchronizacja będzie zachodzić tylko na odcinkach funkcji korelacji wzajemnej o jednoimiennym znaku nachylenia. Przy danej charakterystyce przestrajania, niezbędne dla prawidłowej regulacji, ujemne częstotliwościowe sprzężenie zwrotne będzie występować tylko dla określonego znaku nachylenia charakterystyki detektora błędu, tzn. dla określonego znaku pochodnej funkcji korelacji. Np. przy $s > 0$ punkt pracy w stanie synchronizmu będzie się poruszał po opadającej części charakterystyki $K_{v,v_h}(F\tau_o)$ (rys. 7).

---

[7] Podobnie jak poprzednio, z uwagi na impulsowy charakter przebiegów, chwilowe wartości częstotliwości i sygnału modulującego należy traktować jako wartości pojedyńczych próbek (tzn. wartości średnie za okres $T = \dfrac{1}{F} = \dfrac{1}{F_h}$).

[8] Przy założeniu upraszczającym $V_i = V_h$ oraz $q = 0{,}5$ dla obu przebiegów prostokątnych $v(t)$ i $v_h(t)$.

Charakterystyka statyczna detektora synchronizowanego (rys. 3d i e) jest fragmentem charakterystyki przestrajania (rys. 9). Przy liniowej charakterystyce przestrajania będzie to jednocześnie fragment funkcji korelacji wzajemnej (rys. 7). Jeżeli zakres liniowego odcinka charakterystyki przestrajania ΔU (rys. 9) jest szerszy od międzyszczytowej wartości $\frac{V_o}{2}$ funkcji korelacji wzajemnej, to krańcowe punkty charakterystyki statycznej detektora będą wyznaczone przez wartości ekstremalne funkcji korelacji. W przypadku przeciwnym, krańcowe punkty charakterystyki statycznej odpowiadają granicznym punktom zakresu przestrajania (U_g, U_d) charakterystyki $F_h(u)$. Przy dewiacjach częstotliwości wykraczających poza określone wyżej granice, wyznaczające tzw. zakres utrzymywania synchronizmu ΔF_u (rys. 3d i e), pętla wypadnie z synchronizmu i napięcie wyjściowe detektora nie będzie jednoznaczną funkcją częstotliwości sygnału wejściowego. Wartość średnia napięcia wyjściowego będzie — w tym przypadku — równa w przybliżeniu średniej arytmetycznej ekstremalnych wartości napięć wyjściowych układu mnożącego $\left(u \cong \frac{V_o}{4}\right)$, ponadto wystąpi na wyjściu detektora odkształcona składowa zmienna napięcia o częstotliwości podstawowej równej częstotliwości dudnień $|F-F_h|$. Aby wymusić powrót pętli do stanu synchronizmu należy dostroić generator lokalny w taki sposób, aby częstotliwość dudnień była mniejsza od połowy wartości tzw. zakresu chwytania ΔF_c pętli. Zakres chwytania pętli zależy od transmitancji filtru FD i jest tym mniejszy od zakresu utrzymywania im niższa jest częstotliwość graniczna f_g tego filtru. (Przy transmitancji filtru równej jedności w całym zakresie częstotliwości, zakres chwytania jest równy zakresowi utrzymywania.) Odpowiednie przejścia, związane z utrzymywaniem i chwytaniem synchronizmu w pętli automatyki, zaznaczono na rys. 3d i e strzałkami oraz pomocniczymi liniami przerywanymi.

Detektor synchronizowany na ogół odznacza się gorszą liniowością w porównaniu do detektorów normalizującego i autokorelacyjnego. Liniowość charakterystyki statycznej detektora synchronizowanego zależy bowiem dodatkowo od kształtu charakterystyki przestrajania $F_h(u)$.

Unikalną zaletą detektora synchronizowanego jest bardzo mała wartość tzw. progu poprawy. Próg poprawy jest parametrem charakterystyk szumowych modulacji, określonym przez stosunek sygnału do szumu na wejściu detektora, odpowiadający punktowi zmiany nachylenia tych charakterystyk.

Charakterystyczną cechą modulacji kątowych jest znaczne polepszenie stosunku sygnału do szumu po przejściu przez detektor, pod warunkiem, że indeks modulacji jest duży a stosunek sygnału do szumu na wejściu detektora ma wartość większą od progu poprawy. Poniżej progu poprawy wyjściowy stosunek sygnału do szumu znacznie maleje. Dla detektorów niesynchronizowanych poziom szumów wejściowych określony jest m.in. przez skuteczną szerokość pasma przenoszonego przez wzmacniacze selektywne poprzedzające detektor. Przy wartościach indeksu modulacji większych od jedności sygnał zajmuje pasmo o szerokości rzędu $2(m+1)f_m$, znacznie szersze od pasma zajmowanego przez sygnał modulujący, co wymaga stosowania odpowiednio szerokopasmowych wzmacniaczy selektywnych. Sygnał poddawany w takich warunkach detekcji charakteryzuje się inten-

sywnymi losowymi odchyleniami fazy i częstotliwości o widmie rozciągającym się od zera do częstotliwości rzędu $(m+1)f_m$[9]. Pętla automatycznej regulacji fazy w detektorze synchronizowanym, dzięki wtrąconemu filtrowi dolnoprzepustowemu FD, działa dla fluktuacji fazy również jak filtr dolnoprzepustowy o takiej samej częstotliwości granicznej. (Jak wspomniano wyżej, częstotliwość generatora lokalnego nie będzie w stanie nadążać za szybkimi fluktuacjami częstotliwości sygnału i fluktuacje te nie wystąpią w napięciu wyjściowym detektora.) Dla przeniesienia sygnału pożądanego wystarczy aby pasmo przepustowe filtru obejmowało widmo sygnału modulującego (tzn. aby $f_g \cong f_m$). Wobec tego pętla automatyki działa jak skuteczny eliminator zakłóceń sygnału poddawanego detekcji (przed detekcją!), przesuwając próg poprawy w stronę mniejszych (w przybliżeniu m-krotnie) stosunków sygnału do szumu na wejściu detektora [7, 8].

Dla impulsowych detektorów synchronizowanych produkowany[10] jest specjalny obwód scalony. W skład obwodu scalonego wchodzą: impulsowy, przestrajany napięciem generator lokalny, układ mnożący jako detektor fazy i filtr dolnoprzepustowy z zewnętrznym kondensatorem. W takiej postaci obwód jest wykorzystywany m.in. do bezpośredniej detekcji częstotliwości wąskopasmowych sygnałów analogowych. Aby zrealizować w oparciu o powyższy obwód impulsowy detektor synchronizowany należy uzupełnić układ wejściowy przetwornikiem analogowo impulsowym.

Detektor synchronizowany może być również zrealizowany w wersji analogowej z analogowym detektorem fazy w roli detektora błędu i generatorem lokalnym napięcia sinusoidalnego. W tym przypadku jednak do nieliniowości charakterystyki statycznej detektora, związanych z nieliniowością charakterystyki przestrajania, dochodzi dodatkowo wpływ nieliniowej funkcji korelacji wzajemnej sygnału i heterodyny (nieliniowa charakterystyka detektora błędu).

7. PODSUMOWANIE

Na podstawie przeprowadzonych wyżej rozważań można porównać w sposób bardziej ogólny omówione układy detekcyjne. Porównanie takie dotyczy głównie potencjalnych możliwości, które mogą być osiągnięte przy odpowiednim zaprojektowaniu i wykonaniu detektora.

Opisane wyżej cztery podstawowe układy impulsowych detektorów częstotliwości dadzą się, na podstawie analogii procesów fizycznych, zgrupować w dwie pary o zbliżonych właściwościach. Działanie detektorów normalizującego i zliczającego oparte jest na standaryzacji impulsów napięcia lub ładunku, natomiast podstawowym procesem zachodzącym w detektorach autokorelacyjnym i synchronizowanym jest mnożenie dwóch przebiegów impulsowych.

Detektor normalizujący jest najbardziej liniowym z detektorów impulsowych, warunek liniowości jest tu bowiem spełniony przy stałej powierzchni impulsów znormalizo-

[9] Próg poprawy związany jest ze zjawiskiem występowania na wyjściu detektora losowych impulsów towarzyszących losowym odchyleniom fazy sygnału wejściowego o wartość większą od 2π rad. W takim przypadku wypadkowy wskaz zakłóconego sygnału wykonuje dodatkowy obrót wokół początku układu współrzędnych.

[10] Np. typu NE 565 produkowany przez firmy Signetics i Motorola.

wanych. Detektor zliczający odznacza się największą odpornością na niepożądaną modulację amplitudy. Warunek maksymalnego tłumienia jest dla tego detektora spełniony przy zapewnieniu stałej wartości międzyszczytowej impulsów, niezależnie od ich kształtu. Przy spełnieniu odpowiedniego z wymienionych wyżej warunków, detektory normalizujący i zliczający mogą pracować liniowo z formowaniem impulsów o innym kształcie niż prostokątny.

Z przeprowadzonych pomiarów wynika, że dla detektorów normalizującego i zliczającego, zawierających po pięć tranzystorów (trzy dla przetwarzania wejściowego i dwa dla normalizacji lub zliczania), pracujących w zakresie $100 \div 400$ kHz, uzyskano bez specjalnych zabiegów następujące wartości parametrów:

— nachylenie charakterystyki statycznej: 9 V/MHz,
— zniekształcenia nieliniowe przy dewiacji częstotliwości 50 kHz: $0{,}2 \div 0{,}04\%$,
— tłumienie modulacji amplitudy, określone w sposób opisany w p. 2: $50 \div 60$ dB.

Specyficzną cechą detektora autokorelacyjnego jest periodyczna charakterystyka statyczna. Ta unikalna własność może mieć znaczenie w specjalnych zastosowaniach, umożliwiając np. odbiór sygnałów o różnych częstotliwościach nośnych. Detektor synchronizowany zapewnia najmniejszą wartość progu poprawy i z tego powodu nadaje się specjalnie do odbioru sygnałów przy dużym poziomie szumów. Liniowość detektora autokorelacyjnego jest na ogół lepsza niż synchronizowanego, ponieważ zależy tylko od kształtu impulsów, podczas gdy w detektorze synchronizowanym wymagana jest dodatkowo liniowość charakterystyki przestrajania. Zdolność do tłumienia modulacji amplitudy jest w obu detektorach jednakowa, określona przez parametry przetwornika wejściowego.

Zasada działania detektorów autokorelacyjnego i synchronizowanego jest wykorzystywana również do bezpośredniej (tzn. bez wstępnego przetwarzania sygnału na przebieg impulsowy) detekcji ciągłych sygnałów wąskopasmowych o modulacji częstotliwości. W rozwiązaniach takich staje się jednak niezbędne stosowanie obwodów rezonansowych, a z uwagi na sinusoidalny charakter przebiegów liniowość detekcji jest znacznie gorsza. Również i inne parametry analogowych detektorów autokorelacyjnych i synchronizowanych będą z reguły gorsze od parametrów ich impulsowych odpowiedników.

W tablicy 1 zestawiono wzory określające nachylenie charakterystyki statycznej $\frac{du}{dF}$ detektorów oraz warunki dotyczące maksymalnej wartości dewiacji częstotliwości ΔF sygnału wejściowego i wartości częstotliwości nośnej F_o tego sygnału. Zakładając, że wartości międzyszczytowe wszystkich sygnałów impulsowych są w przybliżeniu równe (tzn. $V_i \cong V_n \cong V_o$) oraz, że częstotliwość graniczna filtru dolnoprzepustowego jest znacznie mniejsza od dewiacji częstotliwości (tzn. $f_g \cong f_m \ll \Delta F$), łatwo zauważyć, iż największą osiągalną w tych warunkach wartością nachylenia charakterystyki $\left(\frac{V_n}{2\Delta F}\right)$ odznacza się detektor normalizujący. Odpowiednie wartości dla detektorów autokorelacyjnego i synchronizowanego są dwukrotnie mniejsze $\left(\frac{V_o}{4\Delta F}\right)$, a dla detektora zliczającego co najmniej pięciokrotnie mniejsze.

Tablica 1

Porównanie parametrów charakterystyk statycznych detektorów impulsowych

Typ detektora	$\left\lvert\frac{du}{dF}\right\rvert$	ΔF	F_0
normalizujący	$A_n = V_n\tau_n$ [1]	$f_m+2\Delta F < \frac{1}{\tau_n}$	$f_m+\Delta F < F_0 < \frac{1}{\tau_n}-\Delta F$
zliczający	$V_i R_2 C_1$	$f_m+2\Delta F < \frac{0{,}1}{R_2C_1}$	$f_m+\Delta F < F_0 < \frac{0{,}1}{R_2C_1}-\Delta F$
autokorelacyjny	$V_0\tau_0$	$4\tau_0\Delta F < 1$ [2]	$F_0 \cong \frac{2n-1}{4\tau_0}$ n — liczba naturalna
synchronizowany	$\frac{1}{\lvert s\rvert}$	$2\Delta F < \Delta F_u = \min\left(\lvert s\rvert\Delta U,\ \lvert s\rvert\frac{V_0}{2}\right)$ [3]	$F_0 \cong F_h$

[1] dla prostokątnych impulsów znormalizowanych,
[2] dla współczynnika wypełnienia $q = 0{,}5$
[3] dla modułu transmitancji filtru dolnoprzepustowego w paśmie przenoszenia równym jedności

Jak wynika z kształtów charakterystyk statycznych zamieszczonych na rys. 3, maksymalne, osiągalne nachylenie tych charakterystyk można zrealizować dla detektorów normalizującego i zliczającego tylko przy pracy w zakresie niskich częstotliwości dla sygnałów o względnej szerokości pasma większej od jedności. Będzie to zachodzić przy spełnieniu następujących warunków: $F_0 \cong \Delta F$ oraz $F_0 \cong \frac{1}{2\tau_n}$ dla detektora normalizującego i $F_0 \cong \frac{0{,}1}{R_2C_1}$ dla detektora zliczającego. Dla detektorów autokorelacyjnego i synchronizowanego osiągalna wartość nachylenia charakterystyki statycznej nie zależy w zasadzie od zakresu częstotliwości pracy. Dla tych detektorów natomiast pewną rolę odgrywa ograniczenie maksymalnej wartości dewiacji częstotliwości: $\Delta F < \frac{1}{4\tau_0}$ dla detektora autokorelacyjnego i $\Delta F < \Delta F_u$ dla detektora synchronizowanego. Detektor autokorelacyjny przy odpowiednio dużej wartości opóźnienia τ_0 może być wykorzystywany do odbioru szerokopasmowych sygnałów o niskiej częstotliwości nośnej. Dla detektora synchronizowanego występuje w tym przypadku ograniczenie natury technicznej, związane z osiągalnymi wartościami zakresu względnego przestrajania generatora lokalnego i zakresu utrzymywania synchronizmu ΔF_u.

Schematy elektryczne impulsowych detektorów częstotliwości są oczywiście bardziej rozbudowane i zawierają większą liczbę elementów w porównaniu z detektorami analogowymi. Ta cecha charakterystyczna, silnie ograniczająca dotychczas rozwój zastosowań detektorów impulsowych, traci obecnie coraz bardziej na znaczeniu z uwagi na rozpowszechnianie się technologii układów scalonych. Dzięki tej technologii stosowanie detektorów impulsowych staje się ekonomicznie uzasadnione nawet w tanim sprzęcie masowego użytku.

LITERATURA CYTOWANA W TEKŚCIE

1. P. A. Agadżanow, B. M. Groszkow, G. D. Smirnow, Osnowy radiotelemetrii, Woj. Izdad., Moskwa 1971.
2. E. D. Frost, Pulse-Counting F. M. Tuner, Wireless World, nr 12, 1965.
3. P. I. Jewdokimow, Algoritm i schiema idealnogo czastotnogo detektora, Praca zbiorowa, Metody pomiechoustoiczi wego prijema CzM i FM, Sow. Radio, Moskwa 1970.
4. W. Rotkiewicz, P. Rotkiewicz, B. Zalewski, Technika odbioru radiowego. Podstawowe układy wielkiej częstotliwości, WNT, Warszawa 1973.
5. F. Balik, Ogólna charakterystyka i klasyfikacja detektorów całkujących FM (w przygotowaniu).
6. J. B. Earnshaw, The Diode Pump Integrator, El. Engineering, nr 1, 1956.
7. A. J. Viterbi, Principles of Coherent Communication, McGraw-Hill Book Company, lub tłumaczenie rosyjskie Sow. Radio, Moskwa 1970.
8. A. S. Winickij, Modulirowannyje filtry i sliediaszczij pijem CzM, Sow. Radio, Moskwa 1969.

R. NOWAK

FM PULSE DETECTORS

Summary

The review of FM pulse detectors is presented. The pulse FM detector has a continuous — to — pulse signal converter for amplitude limitation and for restoration of a modulating signal. Pulse *FM* detectors have no inductors and may by realised simply as integrated circuits.

Four basic circuits of pulse *FM* detectors are described and compared: a detector with pulse normalising unit, a pulse counting detector, an autocorelation detector and a detector with phase-locked loop. The paper is concerned mainly with nonlinear distortion and *AM* attenuation in *FM* pulse detectors.

R. NOWAK

FM-IMPULSFREQUENZDETEKTOREN

Zusammenfassung

Es wird eine Übersicht über die Impulsfrequenzdetektoren dargestellt. Zu den Impulsfrequenzdetektoren gehören Schaltungen mit Impulsbildung. Die Impulsbildung begrenzt die Amplitudenmodulation und vereinfacht die Abbildung des Modulationssingals. Die Impulsfrequenzdetektoren enthalten keine Induktivitäten und deshalb können sie leicht als integrierte Schaltungen realisiert werden.

Es werden vier Grundschaltungen der Detektoren beschrieben: 1) mit Impulsnormalisierungsstufe, 2) mit Diodenpumpenstufe, 3) mit Autokorelatorstufe und 4) mit automatischer Frequenznachstimmung. Die physikalischen Grundlagen von Verzerrungen und Amplitudenmodulationsdämpfung sind ausführlich erörtert worden.

Р. НОВАК

ИМПУЛЬСНЫЕ ДЕТЕКТОРЫ ЧМ СИГНАЛОВ

Резюме

В статье рассматриваются системы детекторов с преобразованием непрерывных сигналов в импульсные. Это преобразование одновременно ограничивает влияние амплитудной модуляции и дает возможность воспроизвести модулирующий сигнал с помощью фильтра низкой частоты. Из схем импульсных детекторов исключена индуктивность, благодаря чему они могут быть без затруднений исполнены как интегральные схемы.

Рассмотрены четыре основные схемы импульсных детекторов: нормирующие импульсы, считающие импульсы, автокорреляционные и с фазовой автоподстройкой частоты.

Особое внимание уделено линейности детекторных характеристик и подавлению амплитудной модуляции.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 1, ss. 143-152

621.3.089

Sterowany cyfrowo stabilizator napięciowo-prądowy

JANUSZ SOSNOWSKI (WARSZAWA)

Instytut Maszyn Matematycznych Politechniki Warszawskiej

Otrzymano 21.10.1972

W pracy przedstawiono własności i parametry sterowanych cyfrowo (tzn. z możliwością cyfrowego ustawiania wartości parametrów wyjściowych) stabilizatorów napięciowo-prądowych.

Ponadto podano oryginalną koncepcję realizacji układu. Wartości parametrów wyjściowych (napięcie, prąd) ustawiane są za pomocą półprzewodnikowych konwerterów cyfrowo-analogowych, dzięki czemu układ odznacza się dużą szybkością przy zachowaniu dużej dokładności. Możliwości funkcjonalne układu (stabilizator napięciowy, prądowy, komparator napięcia, układ pomiaru napięcia i prądu) oraz jego własności (szybkość i dokładność) predysponują go do systemów automatycznego testowania i pomiarów.

1. WSTĘP

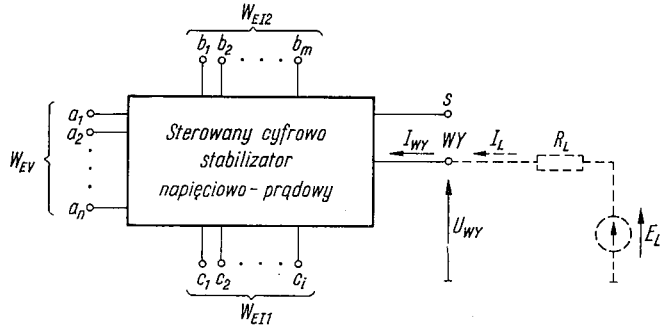
Wraz z rozwojem systemów cyfrowych wzrasta zapotrzebowanie na urządzenia cyfrowo-analogowe umożliwiające systemowi cyfrowemu bezpośrednie oddziaływanie na współpracujący z nim obiekt rzeczywisty lub też śledzenie jego pracy. Do grupy tych urządzeń należą sterowane cyfrowo stabilizatory napięcia i prądu [8÷10], [12] (najczęściej stosowane w automatycznych systemach testowania i pomiarów [2], [11]). Wartość napięcia wyjściowego stabilizatora napięcia lub wartość prądu wyjściowego stabilizatora prądu zadawana jest w postaci cyfrowej (proces ustawiania tych wartości nazywa się programowaniem stabilizatora [12]).

W pracy przedstawiono ogólne własności oraz parametry stabilizatorów napięciowo-prądowych (sterowanych cyfrowo). Ponadto podano koncepcję układu stabilizatora charakteryzującego się dużą szybkością programowania (rozumianą jako odwrotność czasu upływającego od zmiany sygnału na wejściu cyfrowym do chwili ustalenia się sygnału wyjściowego z założoną dokładnością [12]) przy zachowaniu dużej dokładności (co wynika z zastosowania półprzewodnikowych konwerterów cyfrowo-analogowych).

2. OGÓLNE WŁASNOŚCI STEROWANEGO CYFROWO STABILIZATORA NAPIĘCIOWO-PRĄDOWEGO

Ogólny schemat sterowanego cyfrowo stabilizatora napięciowo-prądowego przedstawiono na rys. 1. W układzie można wyróżnić 2 grupy wejść cyfrowych: napięciowe ($W_{EV}-a_1, a_2, \ldots, a_n$) i prądowe ($W_{EI1}-b_1, b_2, \ldots, b_m$, $W_{EI2}-c_1, c_2, \ldots, c_i$), jedno wyj-

ście analogowe WY (napięciowo-prądowe) i jedno wyjście cyfrowe S (sygnalizujące stan układu — stabilizacja prądu lub napięcia). Sygnały podawane na wejścia $a_1 \div a_n$, $b_1 \div b_m$, $c_1 \div c_i$ określają wartości programowanych parametrów[1]. Na wejściu W_{EV} programowana jest wartość (V) napięcia wyjściowego, na wejściu W_{EI1} wartość (I_1) maksymalnego prądu jaki może wpłynąć do stabilizatora (prąd absorbowany: $I_{WY} > 0$) i na wejściu W_{EI2} wartość (I_2) maksymalnego prądu jaki może wypływać ze stabilizatora (prąd emitowany: $I_{WY} < 0$). Wartości napięcia wyjściowego (U_{WY}) i prądu wyjściowego (I_{WY}) są funkcją V, I_1, I_2 oraz parametrów obciążenia.



Rys. 1. Schemat ogólny sterowanego cyfrowo stabilizatora napięciowo-prądowego

Przy obciążeniu w postaci szeregowego połączenia oporności R_L i źródła napięcia E_L (jak na rys. 1) zależność tę opisują równania:

dla $R_L > \max\left(\frac{V-E_L}{I_2}, \frac{E_L-V}{I_1}\right)$

$$U_{WY} = V,$$
$$I_{WY} = I_L = \frac{E_L-V}{R_L}; \tag{1a}$$

dla $R_L \leqslant \frac{V-E_L}{I_2}$

$$U_{WY} = E_L - R_L I_L,$$
$$I_{WY} = I_L = -I_2; \tag{1b}$$

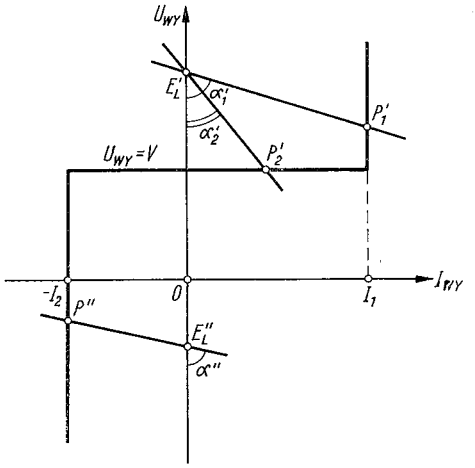
dla $R_L \leqslant \frac{E_L-V}{I_1}$

$$U_{WY} = E_L - R_L I_L,$$
$$I_{WY} = I_L = I_1. \tag{1c}$$

[1] Programowanie to polega na zakodowaniu wartości parametru (napięcie, prąd) przy pomocy odpowiadających mu sygnałów cyfrowych. W praktyce najczęściej stosuje się kody binarne (reprezentowane przez sygnały cyfrowe przyjmujące wartości „0" i „1") — naturalny i BCD [3].

Równania (1) ilustruje wykres z rys. 2, na którym przedstawiono charakterystykę wyjściową stabilizatora i charakterystyki obciążenia. Punkt przecięcia obu charakterystyk wyznacza parametry punktu pracy $P(U_{WY}, I_{WY})$.

Układ o podanych wyżej własnościach może być wykorzystywany jako stabilizator napięciowy z ograniczeniem prądu wyjściowego lub stabilizator prądowy z ograniczeniem napięcia wyjściowego [12]. Ograniczenie prądu i napięcia wyjściowego zabezpiecza stabilizator i obciążenie przed ewentualnym zniszczeniem. Wartość ograniczenia powinno



Rys. 2. Charakterystyka wyjściowa stabilizatora napięciowo-prądowego z wrysowanymi charakterystykami obciążenia ($\alpha'_1 = \text{arctg}\,(1/R'_{L1})$, $\alpha'_2 = \text{arctg}\,(1/R'_{L2})$ i $\alpha'' = \text{arctg}\,(1/R''_L)$)

się ustalać na podstawie pewnych informacji o obciążeniu. Tak na przykład w stabilizatorze napięciowym wartość ograniczenia prądu ustala się przez oszacowanie maksymalnego prądu jaki może płynąć przez obciążenie dla przewidywanego zakresu napięć wyjściowych.

Stosując układ z rys. 1 jako stabilizator napięcia w zakresie napięć $[\underline{V}, \overline{V}]$ należy na jego wejściach W_{EI1} i W_{EI2} zaprogramować odpowiednio wartości $I_1 = \overline{I_{L\max}}$ i $I_2 = \underline{I_{L\max}}$ (gdzie: $\overline{I_{L\max}}$ — maksymalny prąd jaki może wypłynąć z obciążenia dla $U_{WY} \in [\underline{V}, \overline{V}]$, $\underline{I_{L\max}}$ — maksymalny prąd jaki może wpłynąć do obciążenia dla $U_{WY} \in [\underline{V}, \overline{V}]$). Na wejściu W_{EV} programowana jest wartość ($V \in [\underline{V}, \overline{V}]$) stabilizowanego napięcia.

Wykorzystując układ z rys. 1 jako stabilizator prądu absorbowanego ($I_{WY} > 0$) w zakresie $[\underline{I}, \overline{I}]$ należy na wejściu W_{EV} zaprogramować wartość $V = \underline{U_{WY\max}}$ ($\underline{U_{WY\max}}$ — minimalne napięcie U_{WY} jakie może powstać na obciążeniu przy $I_{WY} \in [\underline{I}, \overline{I}]$). Na wejściu W_{EI1} programowana jest wartość ($I_1 \in [\underline{I}, \overline{I}]$) stabilizowanego prądu.

Analogicznie dla stabilizatora prądu emitowanego ($I_{WY} < 0$) na wejściu W_{EV} należy zaprogramować $V = \overline{U_{WY\max}}$ ($\overline{U_{WY\max}}$ — maksymalne napięcie U_{WY} jakie może powstać na obciążeniu przy $I_{WY} \in [\underline{I}, \overline{I}]$). Wartość stabilizowanego prądu ($I_2 \in [\underline{I}, \overline{I}]$) programowana jest w tym przypadku na wejściu W_{EI2}.

W zależności od zastosowań potrzebne są stabilizatory o różnych własnościach: z dokładnym programowaniem napięcia wyjściowego i mniej dokładnym prądu (stabilizatory

napięcia) lub odwrotnie (stabilizatory prądu), bardzo szybkie o średniej dokładności (systemy automatycznego testowania), wolne o bardzo dużej dokładności (systemy pomiarowe) itd. Wymagania te oczywiście mają istotny wpływ na konstrukcję układową. Wygodnie jest dysponować pewnym zestawem parametrów, które pozwoliłyby na jakościową i ilościową ocenę własności układu. Poniżej proponuje się zestaw parametrów zewnętrznych charakteryzujących te własności statyczne i dynamiczne, które są najbardziej istotne przy wykorzystywaniu sterowanych cyfrowo stabilizatorów napięciowo-prądowych.

Parametry statyczne

a) Zakres programowania (napięcia, prądu) — parametr ten określa maksymalne i minimalne wartości napięcia (V_{min}, V_{max}) lub prądu (I_{min}, I_{max}) jakie mogą być stabilizowane na wyjściu.

b) Rozdzielczość programowania (napięcia, prądu) — najmniejsza wartość (kwant) o jaką można zmienić wartość stabilizowanego napięcia lub prądu. Rozdzielczość określona jest jednoznacznie przez przyjęty sposób kodowania sygnałów cyfrowych. Dla układu z rys. 1 przy kodzie naturalnym rozdzielczość programowania przyjmuje następujące wartości:

dla wejścia $W_{EV} \div \frac{V_{max}-V_{min}}{2^n}$,

dla $W_{EI1} \div \frac{I_{1max}-I_{1min}}{2^m}$,

dla $W_{EI2} \div \frac{I_{2max}-I_{2min}}{2^i}$.

c) Dokładność programowania (napięcia, prądu) — dokładność bezwzględna równa jest modułowi różnicy między wartością stabilizowanego napięcia lub prądu na wyjściu a wartością na odpowiednim wejściu cyfrowym, natomiast dokładność względna jest dokładnością bezwzględną odniesioną do zakresu programowania. Obie dokładności są funkcją obciążenia (związane jest to z impedancją wyjściową stabilizatora) oraz wartości stabilizowanych wielkości. Dokładność względna powinna być mniejsza od połowy wartości rozdzielczości programowania.

d) Histereza reakcji na przeciążenie (napięciowe, prądowe) — różnica między wartością prądu lub napięcia na wyjściu, po przekroczeniu której układ przechodzi od stabilizacji napięciowej do prądowej, a wartością po przekroczeniu której układ przechodzi od stabilizcji prądowej do napięciowej.

Parametry dynamiczne

a) Czas programowania (napięcia, prądu) — czas jaki upływa od momentu podania na wejście cyfrowe (napięciowe W_{EV} lub prądowe W_{EI}) nowej wartości do chwili ustalenia się jej (z założoną dokładnością) na wyjściu. Wartość tego parametru jest funkcją para-

metrów obciążenia (pojemność) oraz amplitudy zmiany wartości na wejściu. Ponadto zależy ona od wartości ograniczeń; tak np. czas programowania napięcia wyjściowego zależy od ustawionego ograniczenia prądu wyjściowego (które określa maksymalny prąd ładowania pojemności obciążenia).

b) Czas odpowiedzi na zmianę obciążenia — czas jaki upływa od momentu skokowej zmiany obciążenia (wywołującej zakłócenie w stabilizacji) do momentu ustalenia się wartości stabilizowanego parametru (napięcia, prądu).

c) Czas reakcji na przeciążenie — czas jaki upływa od momentu skokowej zmiany obciążenia wywołującej przeciążenie do chwili zadziałania ograniczenia (innymi słowy czas przejścia od stabilizacji napięciowej do prądowej lub odwrotnie pod wpływem skokowej zmiany obciążenia). Jest to dość istotny parametr pozwalający (dla danej wartości amplitudy przeciążenia) ocenić niebezpieczeństwo uszkodzenia stabilizatora lub obciążenia.

d) Przesłuch z wejść cyfrowych na wyjście — czas i amplituda impulsów zakłócających pojawiających się na wyjściu układu przy zmianie parametrów wejściowych (patrz rozdz. 3).

3. KONCEPCJA UKŁADU

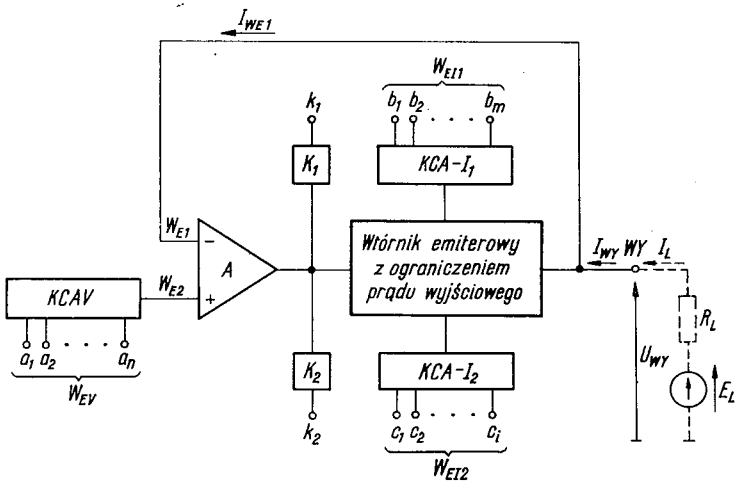
Sterowany cyfrowo stabilizator napięciowo-prądowy można zrealizować w oparciu o znane układy stabilizatorów z regulacją ciągłą (analogową) [6] przez zastąpienie w nich elementów o parametrach regulowanych w sposób ciągły, oporniki i potencjometry, elementami o parametrach zadawanych w sposób cyfrowy, sterowane cyfrowo oporności. Te ostatnie tworzą zbiór oporników połączonych szeregowo lub równolegle. Z każdym opornikiem o ustalonej wartości oporności związany jest przekaźnik elektromechaniczny sterowany sygnałem cyfrowym, który w zależności od wartości tego sygnału włącza dany opornik do tworzonego zbioru lub nie. W ten sposób oporność wypadkowa tworzonego zbioru oporników jest funkcją sygnałów sterujących przekaźnikami (dokładniej opisano to w [4]).

Stabilizatory ze sterowanymi cyfrowo opornikami są produkowane przez niektóre firmy (np. [7]) i charakteryzują się dużą dokładnością programowanych parametrów. Wadą ich jest mała szybkość programowania, ograniczona szybkością przełączania przekaźników elektromechanicznych. Typowe dla przekaźników zjawiska: odbicia i sklejanie się zestyków oraz duży czas przełączania mogą spowodować pojawienie się na wyjściu chwilowych przebiegów znacznie różniących się od zaprogramowanych wartości. Przykładowo, przy zmianie na wejściu W_{EV} wartości sygnałów cyfrowych $a_1, a_2, \ldots, a_n$ z 0, 1, 0, 1, ..., 1, na 1, 0, 1, 0, ..., 0, na wyjściu układu mogą się pojawić przebiegi nieprawidłowe (przesłuch z wejść cyfrowych na wyjście) odpowiadające dowolnym wartościom sygnałów $a_1, a_2, \ldots, a_n$ (np. 1, 1, 1, ... 1; 0, 0, 0, ..., 0). Wynika to stąd, że przedstawiona zmiana wartości sygnałów cyfrowych pociąga za sobą zmiany stanu wszystkich przekaźników. W rzeczywistości niemożliwa jest jednoczesna zmiana stanu wszystkich przekaźników i to jest przyczyną wyżej opisanego zjawiska. Czas trwania przebiegów nieprawidłowych jest rzędu czasu przyłączania przekaźnika, a więc stosunkowo duży, co niekiedy może być nawet przyczyną uszkodzenia obciążenia. Rozbudowując odpowiednio układ sterowania

cyfrowego można uniknąć pojawienia się przebiegów nieprawidłowych. Jedno z rozwiązań podano w [7].

Stosowanie przekaźników elektromechanicznych (od kilkunastu do kilkudziesięciu) nie pozwala na uzyskanie dostatecznie dużej niezawodności układu. W związku z tym układy takie nie są najodpowiedniejsze w zastosowaniach, gdzie dokonuje się częstych zmian wartości programowanych parametrów (np. automatyczne systemy testowania o dużej wydajności [11]).

W celu uniknięcia wad przedstawionych rozwiązań opracowano układ stabilizatora, do programowania którego wykorzystywane są szybkie (półprzewodnikowe) konwertery cyfrowo-analogowe [3]. Schemat blokowy układu przedstawiono na rys. 3.



Rys. 3. Koncepcja układu sterowanego cyfrowo stabilizatora napięciowo-prądowego

Do realizacji układu wykorzystano opracowany przez autora wtórnik emiterowy z ograniczeniem prądu wyjściowego [5] (ze względu na prosty i wygodny sposób programowania wartości ograniczeń). Wtórnik ten pracuje w pętli ujemnego sprzężenia zwrotnego wzmacniacza operacyjnego (A), którego nieodwracające fazy wejście (W_{E2}) jest sterowane z konwertera c-a o wyjściu napięciowym ($KCAV$). Na wejściu W_{EV} tego konwertera programowana jest wartość napięcia V. Rolę źródeł prądowych zasilających wtórnik emiterowy pełnią konwertery c-a o wyjściu prądowym, na których wejściach W_{EI1} i W_{EI2} programowana jest wartość ograniczenia prądowego odpowiednio: I_1 i I_2.

Charakterystyka układu jest analogiczna, jak podana na rys. 2. Wynika to stąd, że dla $-I_2 < I_L < I_1$ wtórnik z ograniczeniem prądu wyjściowego zachowuje się podobnie jak zwykły wtórnik (bez ograniczenia) i że jego prąd wyjściowy ograniczony jest na poziomie I_1 i $-I_2$. Tak więc dla $-I_2 < I_L < I_1 \left[\text{tj. } R_L > \max\left(\frac{V-E_L}{I_2}, \frac{E_L-V}{I_2}\right)\right]$

$$U_{WY} = \frac{VA_1A_2}{1+A_1A_2}, \tag{2}$$

gdzie:

A_1 — wzmocnienie wzmacniacza operacyjnego (A) w układzie otwartym,

A_2 — wzmocnienie wtórnika z ograniczeniem prądu.

We wzorze pominięto wpływ napięcia i prądu dryftu.

Zależność (2) przy spełnionym w praktyce założeniu $A_1 \gg 1$ i $A_2 = 1$ sprowadza się do równości $U_{WY} = V$.

Przy obciążeniu o parametrach jak w (1b) i (1c) wtórnik działa jako stabilizator prądu. W tym przypadku jego wejście jest odseparowane od wyjścia spolaryzowanym w kierunku zaporowym złączem emiter-baza tranzystora lub diody [5], co jednocześnie oznacza przerwanie pętli sprzężenia zwrotnego wzmacniacza (A). Prąd wyjściowy stabilizowany jest na poziomie I_1 lub $-I_2$, a napięcie wyjściowe osiąga wartości różne od V. Przy stabilizacji prądu wyjściowego na poziomie I_1 wzmacniacz operacyjny może wejść w stan nasycenia i na jego wyjściu pojawi się minimalny poziom napięcia $\underline{E_{nas}}$. Stan taki ma miejsce dla $U_{WY} - V > \frac{\underline{E_{nas}}}{A_1}$. Analogicznie przy stabilizacji na poziomie $-I_2$ i $V - U_{WY} > \frac{\overline{E_{nas}}}{A_1}$ na wyjściu wzmacniacza ustala się maksymalny poziom napięcia $\overline{E_{nas}}$.

Wykrywanie pojawienia się na wyjściu wzmacniacza (A) poziomów $\underline{E_{nas}}$ i $\overline{E_{nas}}$ pozwala określić aktualny reżim pracy stabilizatora (prądowy lub napięciowy). Na rys. 3 zadanie to spełniają komparatory, K_1 i K_2, które na swych wyjściach k_1 i k_2 sygnalizują odpowiednio ustalenie się na wyjściu wzmacniacza (A) poziomu $\overline{E_{nas}}$ lub $\underline{E_{nas}}$. Dodanie komparatorów zwiększa możliwości funkcjonalne stabilizatora (patrz rozdz. 4).

Należy zwrócić uwagę, że wartości parametrów rzeczywistego układu różnią się nieco od wartości zaprogramowanych, co wynika z błędów wnoszonych przez poszczególne elementy układu. Tak też po zaprogramowaniu na wejściach W_{EI1}, W_{EI2} wartości I_1 i I_2 prąd wyjściowy układu będzie w rzeczywistości ograniczany na poziomach

$$\begin{aligned} I_1' &= I_1 + \Delta I_1' + \Delta I_1'' + I_{WE1}, \\ I_2' &= I_2 + \Delta I_2' + \Delta I_2'' - I_{WE1}, \end{aligned} \tag{3}$$

gdzie:

$\Delta I_1'$, $\Delta I_2'$ — błąd wnoszony przez konwertery KCA-I_1 i KCA—I_2,

$\Delta I_1''$, $\Delta I_2''$ — błąd wnoszony przez wtórnik [5],

I_{WE1} — prąd wejściowy wzmacniacza operacyjnego.

W praktyce błędy wnoszone przez konwertery i wtórnik są rzędu ułamków μA, natomiast błąd pochodzący od prądu I_{WE1} w przypadku popularnych wzmacniaczy (np. μA 709) rzędu kilku μA. Ten ostatni można zmniejszyć stosując dodatkowy stopień wejściowy (np. na tranzystorach unipolarnych) we wzmacniaczu.

Analogicznie przy zaprogramowaniu na wejściu W_{EV} wartości V na wyjściu otrzymamy (przy spełnieniu warunków na stabilizację napięciową)

$$U_{WY} = \frac{(V + \Delta V' + \Delta V'')A_1 A_2}{1 + A_1 A_2}, \tag{4}$$

gdzie:

$\Delta V'$ — błąd wnoszony przez konwerter KCA-V,

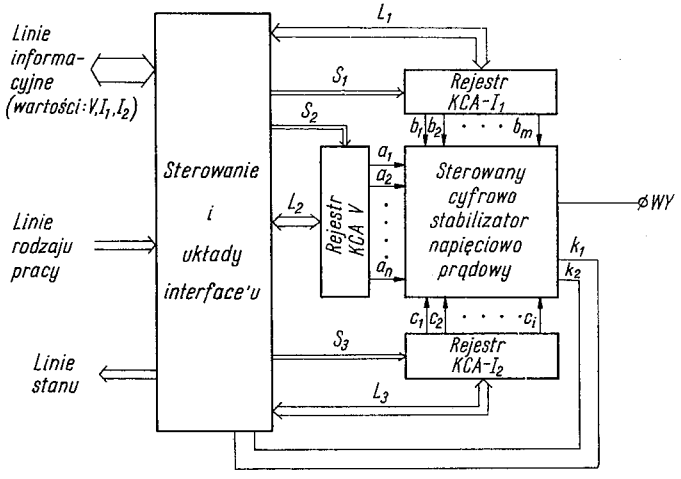
$\Delta V''$ — błąd wnoszony przez wzmacniacz operacyjny (pochodzący od napięcia i prądu dryftu).

Ze względu na duże wartości wzmocnienia (A_1) produkowanych wzmacniaczy operacyjnych dokładność programowania napięcia wyjściowego jest głównie ograniczona błędami $\Delta V'$ i $\Delta V''$ (w popularnych układach rzędu mV). Jest to słuszne dla $V \in [(\underline{E_{nas}} + |\Delta U_{WY}|_{max}), (\overline{E_{nas}} - |\Delta U_{WY}|_{max})]$, gdzie $|\Delta U_{WY}|_{max}$ — maksymalna wartość przesunięcia napięcia na wtórniku między wejściem i wyjściem dla $-I_2 < I_{WY} < I_1$, tzn. przy założeniu, że wzmacniacz nie wchodzi w nasycenie.

W przypadku pracy układu z obciążeniem o dużym poborze prądu celowym może być rozbudowanie stabilizatora o tzw. układ zdalnego sterowania [6], pozwalający na wyeliminowanie błędów pochodzących od spadków napięć na przewodach łączących obciążenie ze stabilizatorem.

4. UWAGI KOŃCOWE

Przedstawiony układ stabilizatora cechuje się dużą dokładnością i szybkością programowania, co wynika z parametrów zastosowanych podzespołów. Istotną jego zaletą jest to, że składa się on z typowych podzespołów elektronicznych (konwertery c/a, wzmacniacz operacyjny) produkowanych w szerokim asortymencie (różne szybkości i dokładności), co znacznie ułatwia realizację techniczną układu. Pewną wadą podanego rozwiązania są



Rys. 4. Schemat blokowy stabilizatora napięciowo-prądowego ze sterowaniem
S_1, S_2, S_3 — linie sterujące (odczyt, zapis do rejestru, strob zapisu itd.), L_1, L_2, L_3 — linie przesyłania informacji do z rejestrów

trudności w uzyskiwaniu dużych wartości napięcia wyjściowego. Trudności te wynikają z ograniczeń na wartość napięcia wyjściowego w konwerterach c/a oraz wartości poziomów $\underline{E_{nas}}$ i $\overline{E_{nas}}$ wzmacniacza operacyjnego.

Stabilizatory napięciowo-prądowe z dodatkowymi komparatorami (jak na rys. 3) sygnalizującymi rodzaj pracy mogą być wykorzystane jako komparatory napięcia o progra-

mowanym progu komparacji V (przy $I_1 = I_2 = 0$), układy do pomiaru prądu przy określonym wymuszeniu napięciowym V lub do pomiaru napięcia przy określonym wymuszeniu prądowym (I_1, $-I_2$). W dwu ostatnich przypadkach układ wymaga specjalnego sterowania. Przykładowo w celu zmierzenia prądu (w obciążeniu) przy wymuszeniu napięciowym V należy na wejściu W_{EV} zaprogramować wartość V, a następnie zmieniać wartość ograniczenia prądu wyjściowego tak, aby stabilizator przeszedł od stabilizacji napięciowej do prądowej (co zostanie zasygnalizowane przez jeden z komparatorów K_1, K_2).

Zmiana ta może być realizowana według analogicznych algorytmów jak w konwerterach analogowo-cyfrowych [3]. Wartość prądu, po przekroczeniu której układ zmienia rodzaj pracy (ze stabilizacji napięcia na stabilizację prądu), jest wynikiem pomiaru.

Szerokie możliwości funkcjonalne układu predysponują go do systemów automatycznego testowania i pomiarów, gdzie stosuje się sterowane cyfrowo stabilizatory napięcia, prądu i układy pomiarowe [1, 2]. Zastępując te układy stabilizatorami o podanych możliwościach zwiększa się uniwersalność systemu (w zależności od realizowanej w systemie czynności możliwe jest dysponowanie zmienną liczbą stabilizatorów napięcia, prądu i układów pomiarowych).

Dla umożliwienia wykonywania przez stabilizator wielu funkcji należy go wyposażyć w odpowiedni układ sterowania [12]. Układ ten przyjmuje informacje od współpracującego urządzenia cyfrowego i w zależności od rodzaju pracy różnie ją interpretuje (przesłanie do rejestru KCA V, KCA-I_1, lub KCA I_2, pomiar U_{WY}, pomiar I_{WY}, komparacja itd.). Ponadto sterowanie sygnalizuje urządzeniu współpracującemu stan stabilizatora (K_1, K_2, koniec wykonywania operacji, np. pomiaru lub komparacji) lub też podaje wynik pomiaru bądź komparacji. Ogólny schemat blokowy stabilizatora ze sterowaniem podano na rys. 4.

LITERATURA CYTOWANA W TEKŚCIE

1. D. Mactaggart, Computer, Software, p-c Cards Match up to Cut Costs in an Automatic Test System, Electronics, July, 6, 1970, s. 71-75.
2. G. C. Padwick, IC Test Equipment, Microelectronics, January, 1968, s. 22—25.
3. H. Schmid, Electronic Analog-Digital Conversion, Van Nostrand Reinhold Company, New York, 1970.
4. W. B. Smołow, Przetworniki dyskretno-analogowe i ich zastosowanie, WNT, 1963.
5. J. Sosnowski, Ograniczanie prądu wyjściowego we wtórniku emiterowym, Rozprawy Elektrotechniczne, z. 4, 1973.
6. M. Starbowski, Stabilizatory sterowane napięcia i prądu, WNT, 1972.
7. KEPCO Talks Power Supply Technology, How to Teach an Analog Voltage Digital Tricks, Electronics, July, 6, 1970, s. 94.
8. New Bipolar Power-Dac Solves Five Major System Problems in Automatic Test Equipment, Electronics, April, 24, 1972, s. 15.
9. Programable Supplies Adapt to Automatic Systems, Electronics, January, 4, 1971, s. 85—86.
10. Supplies are Programable Power Sources Keep Voltage Current Levels Constant for up to 10 000 Test a Second, Electronics, November, 8, 1971, s. 125.
11. J-259 Computer Operated Circuit System, Teradyne, Inc., March 1968.
12. DC Power Supply Handbook, Hewlett Packard, 1970, Application Note 90A.

J. SOSNOWSKI

DIGITALY PROGRAMMED VOLTAGE-CURRENT SUPPLY

Summary

In the paper there are presented properties and parameters of the digitaly programmed voltage-current supplies. There is given an original realization of a digitaly programmed supply. Voltage and current levels are digitaly programmed by means of binary-coded electrical signals which are directed to fast solid-state digital analog converters. Because of the functional possibilities (voltage supply with output current limitation, current supply with output voltage limitation, voltage comparator, current and voltage measuring circuit) and its properties (speed and accuracy) the circuit is very usefull in automatic test and measuring systems.

J. SOSNOWSKI

STABILISATEUR VOLTAGE-COURANT AVEC PROGRAMMATION DIGITALE

Résumé

Dans l'article on décrit les propriétés et les paramètres des stabilisateurs voltage-courant avec programmation digitale.

On donne une réalisation originale du circuit du stabilisateur. Les valeurs des paramètres de sortie (voltage et courant) sont programmables avec les converteurs digitale-analogic (en semiconducteur) alors la vitesse du circuit est très grande. A cause des possibilités foctionnelles (stabilisation voltage, stabilisation courant, comparaison, mesures du voltage et du courant) et des propriétés (vitesse et precission) du circuit, ces stabilisateurs sont très utiles dans les systèmes automatiques de mesure.

J. SOSNOWSKI

ZAHLGESTEUERTER STROMSPANNUNGSSTABILISATOR

Zusammenfassung

In dieser Bearbeitung wurden die Eigenschaften und Kennwerte eines zahlengesteuerten Stromspannungsstabilisator dargestellt, d.h. eines Stabilisators, dessen Ausgangsparameterwerte zahlenmäßig eingestellt werden können.

Außerdem wurde eine eigene neue Realisierungskonzeption des Systems angegeben. Die Ausgangsparameterwerte (Spannung, Strom) werden mittels zahlenanalogen Halbleiterkonvertern eingestellt, weshalb sich das System durch besonders schnelles Arbeiten beim gleichzeitigen Beibehalten großer Genauigkeit auszeichnet. Dank seinen Funktionsmöglichkeiten (Stromspannungsstabilisator, Spannungskompatator, Spannungs- und Strommeßsystem) sowie Eigenschaften (Schnelligkeit und Genauikeit) eignet sich das System zum Einsatz für Systeme automatischen Testens und Messens.

Я. СОСНОВСКИ

ЦИФРОВО ПРОГРАММИРОВАННЫЙ СТАБИЛИЗАТОР НАПРЯЖЕНИЯ И ТОКА

Резюме

В работе описаны свойства и параметры цифрово програмированных стабилизаторов напряжение-ток.

Представлена оригинальная схема стабилизатора, в которой для программирования использованы полупроводниковые преобразователи код-аналог. Функциональные возможности схемы (стабилизатор напряжения с ограничением исходного тока, стабилизатор тока с ограничением исходного напряжения, компаратор напряжения, система для измерений напряжения и тока) и параметры (скорость и точность программирования) предрасполагают его для систем автоматического тестирования и автоматических измерений.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 1, ss. 153-162

621.317.724

Badania modelowe uderzeń pioruna w czynne kominy lub inne źródła ciepła

JERZY JANUSZ ZIELIŃSKI (WARSZAWA)
Instytut Elektrotechniki, Warszawa

Otrzymano 4.8.1972

Analiza warunków termicznych panujących nad czynnymi kominami. Propozycja odtworzenia tych warunków przy modelowych badaniach piorunowych przez zastąpienie ciągu spalin kominowych ciągiem spalin płomyka gazu świetlnego doprowadzonego do elektrod rurowych. Badania modelowe rozkładu temperatur oraz wybiorczości pioruna w elektrody zimne i gorące. Wykazanie, że słup gorących spalin wywołuje taką zmianę wybiórczości pioruna, jak podwyższenie elektrody modelującej komin do poziomu odpowiadającego temperaturze spalin 60—80°C. Wnioski dotyczące celowości uzupełnienia dotychczasowej oceny stref osłonowych wysokich budowli na podstawie obserwacji kominów.

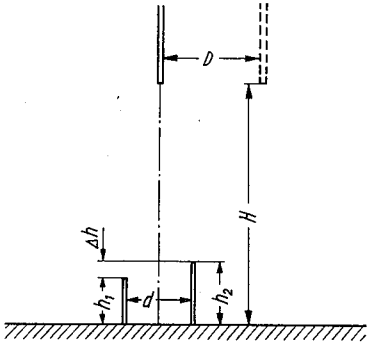
W załączniku przeprowadzono test losowości lokalizacji wyładowań wynikający z prawa serii dla rozkładu dwupunktowego uzasadniając tym prawidłowość doboru liczności i częstości prób.

Badania parametrów wyładowań piorunowych oraz statystyczne badania lokalizacji wyładowań i skuteczności stref osłonowych zwodów prowadzone są w znacznej części na wysokich budowlach. Przyczyną jest prawdopodobieństwo uzyskania liczniejszych wyników, gdyż liczba uderzeń pioruna jest w przybliżeniu proporcjonalna do powierzchni przekroju poziomego strefy osłonowej. Powierzchnia ta jak wiadomo [1] zależy od wysokości zwodu, dla budowli niezbyt wysokich w przybliżeniu proporcjonalnie do kwadratu wysokości, a dla bardzo wysokich w przybliżeniu proporcjonalnie do wysokości. Praktyka wykazała [2, 3], że szczególnie dogodne jest wykorzystywanie do tych celów kominów fabrycznych. Wysokość ich sięga bowiem paruset metrów, prosty kształt i łatwa do określenia droga prądowa ułatwiają rejestrację parametrów, ich ocenę (m.in. w Polsce badania sztabkami magnetycznymi), a duże rozpowszechnienie podobnych obiektów w różnych rejonach kraju umożliwia uzyskanie licznych wyników i ich porównywanie.

Przy ocenie liczby prawdopodobnych wyładowań, jako wynikającej z burzliwości terenu i z wymiarów trafionego obiektu oraz przy określaniu stref osłonowych przyjmuje się dla kominów podobnie jak dla innych budowli (np. drapaczy chmur, masztów radiostacji, słupów linii elektroenergetycznych) rzeczywiste wymiary geometryczne. Odmienność czynnego komina polega jednak na rozwijaniu się nad nim słupa gorących spalin o zwiększonym stopniu zjonizowania w stosunku do otaczającego powietrza.

Zagadnienie wpływu lokalnych ruchów ciepłego powietrza na wybiorczość pioruna jest na ogół znane pod względem jakościowym. Przytacza się przykłady lokalizacji tzw. gniazd burzowych na granicy obszarów leśnego, wodnego i piaszczystego o różnych stopniach

nagrzania przed burzą. Dokonywano we Francji doświadczeń polegających na wywoływaniu wyładowań piorunowych z chmury burzowej przez nagłe zapalenie dużej liczby palników mazutowych skupionych na ograniczonym terenie płaskim. Również w znanym przypadku uderzenia pioruna w rakietę kosmiczną Apollo 12 strumień gorących spalin mógł stanowić czynnik sprzyjający rozwinięciu się wyładowania i jego skierowaniu ku rakiecie. Warto dodać, że przypadek rakiety Apollo można interpretować podobnie jak dla obiektów nieruchomych, gdyż prędkość rakiety — około 10^4 m/s — jest mała w porównaniu z prędkością średnią wstępnego wyładowania piorunowego rzędu $10^5...10^6$ m/s [1, 4], a już zupełnie pomijalna w porównaniu z maksymalną prędkością strzały wyładowania schodkowego rzędu $10^7...10^8$ m/s [1, 4]. Można przypuszczać, że długość ostatniego wyła-



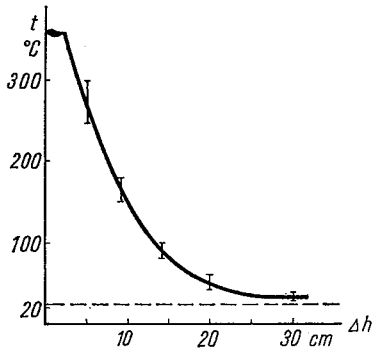
Rys. 1. Szkic przestrzenny układu elektrod modelujących ostrze piorunowe i obiekty naziemne (kominy)

dowania schodkowego przebiegającego w nagrzanych silnie zjonizowanych spalinach silników rakietowych była szczególnie duża, znacznie większa od przeciętnie obserwowanych w powietrzu atmosferycznym schodków o długości kilkunastu lub kilkudziesięciu metrów.

W celu zbadania wpływu strumienia gorących spalin na lokalizację wyładowania przeprowadzono cykl prób modelowych w laboratorium Zakładu Wysokich Napięć Instytutu Elektrotechniki. Zastosowano układ probierczy analogiczny do używanego przez wielu badaczy przy określaniu stref osłonowych zwodów [4, 5, 7]. Źródłem napięcia był generator udarowy 2800 kV, 32 kJ, wytwarzający tzw. udary normalne piorunowe o kształcie 1,2/50 μs i o biegunowości dodatniej, preferowanej do badań modelowych [5, 7], jako lepiej odtwarzającej rozwój wyładowania. Wartości szczytowe udarów podczas prób wynoszące 1100—1200 kV tylko nieznacznie przekraczały tzw. 100-procentowe napięcie przeskoku przerwy iskrowej między modelem ostrza piorunowego a modelami obiektów naziemnych. Szkic układu modelowego podano na rys. 1. Wysokość elektrody prętowej modelującej ostrze piorunowe nad płytą uziemioną była stała i wynosiła $H = 200$ cm. Ustawienie pionowych elektrod uziemionych w postaci uciętych poprzecznie rur o średnicy 2 cm było regulowane; wysokość h_1 i h_2 nie przekraczała 50 cm. Zmniejszenie modelu w stosunku do wymiarów kominów fabrycznych wynosiło więc kilkaset.

Słup gorących spalin wytwarzano przez doprowadzenie kolejno do poszczególnych elektrod rurowych gazu świetlnego i zapalenie płomyka o wysokości ok. 2 cm. Stwierdzono przy tym, że umieszczenie na elektrodzie zakończenia metalowego o wymiarach płomyka nie wpływa w sposób istotny na zmianę lokalizacji wyładowań. Badania prowadzono w tem-

peraturze otoczenia 22...24°C, ograniczając w miarę możności ruchy poziome powietrza w laboratorium (aby uzyskać zbliżony do pionowego kierunek kowekcji spalin). Rozkład temperatur nad elektrodą mierzono termometrem w zakresie od temperatury otoczenia do ok. 150°C. Wyższe temperatury oceniano na podstawie zmian stanu bibułki zawieszonej nad płomykiem, przyjmując jako temperaturę wystąpienia pierwszych zmian koloru bibułki 150–180°C, a jako temperaturę zapalenia 250—300°C. Wykres średnich temperatur w zależności od wysokości Δh nad elektrodą z płomykiem podano na rys. 2.



Rys. 2. Wykres średnich temperatur w zależności od wysokości nad modelową elektrodą gorącą

Badania prowadzono porównując prawdopodobieństwo trafienia wyładowania w poszczególne elektrody w przypadku obu elektrod zimnych i w przypadku zapalonego płomyka na jednej z elektrod. Szczególną uwagę poświęcono racjonalnemu doborowi liczby udarów i częstości udarów (odstępom czasu między kolejnymi udarami). Stosowanie zbyt licznych serii zwiększyłoby pracochłonność badań, a ponadto uniemożliwiłoby przeprowadzenie cyklu serii porównawczych w praktycznie jednakowych warunkach otoczenia. Stosowanie zbyt dużej częstości udarów wprawdzie skróciłoby cykl badań, ale mogłoby wprowadzić istotny wpływ drogi wyładowania poprzedniego udaru na kierunek wyładowania przy udarze następnym (wskutek pozostałości ładunku przestrzennego na tej drodze). W celu uniknięcia takich niepożądanych zjawisk oraz sprawdzenia dopuszczalności proponowanej liczby udarów w serii zastosowano test wynikający z tzw. prawa serii. Szczegółowe omówienie tej metody i jej wyników podano w załączniku. Test powyższy pozwolił na stwierdzenie na dużym poziomie istotności (np. $\alpha = 0{,}025$, $\alpha = 0{,}005$ [6, 8]), że w przypadku stosowania serii po 20 udarów z odstępami między udarami 30...40 sekund, nie występuje wpływ wybiorczości wyładowania przy poprzednim udarze na następny. A zatem można przypuszczać, że prawdopodobieństwo określone takimi seriami jest dostatecznie dokładne, gdyż zwiększenie liczby udarów w serii nie wprowadziłoby systematycznej zmiany wyniku.

Próby przeprowadzono dla następujących zakresów zmian parametrów geometrycznych (p. rys. 1):

a) zmiany odstępu między elektrodami modelowymi w granicach $d = 50...100$ cm (zmiany co 25 cm), przy stałej wysokości obu elektrod $h_1 = h_2 = 50$ cm;

b) zmiany wysokości elektrody z dopływem gazu w granicach $h_1 = 25...50$ cm (zmiany co 5 cm), przy niezmiennej wysokości drugiej elektrody $h_2 = 50$ cm, odstępie $d = 50$ cm

Przykładowe wyniki badań wybiorczości pioruna na modelu wg cyklu b)

Tablica 1

Kolejny udar	Obydwie elektrody A i B zimne			Elektroda A gorąca, elektroda B zimna				
	$h_1 = 50$ cm	$h_1 = 45$ cm	$h_1 = 40$ cm	$h_1 = 50$ cm	$h_1 = 45$ cm	$h_1 = 40$ cm	$h_1 = 35$ cm	$h_1 = 30$ cm
1	A	B	B	A	A	A	B	B
2	B	B	B	A	A	A	A	B
3	A	B	B	A	A	A	A	B
4	A	B	B	A	A	A	A	B
5	A	B	B	A	A	A	B	A
6	A	B	B	A	A	A	B	B
7	B	B	B	A	A	A	A	A
8	B	B	B	A	A	A	A	B
9	B	A	B	A	A	A	A	B
10	A	B	B	A	A	A	B	B
11	A	B	B	A	A	A	A	B
12	B	B	B	A	A	A	A	B
13	B	B	B	A	A	A	A	B
14	B	B	B	A	A	B	B	B
15	B	B	B	A	A	A	A	A
16	A	B	B	A	A	A	A	B
17	B	B	B	A	A	A	A	B
18	B	B	B	A	A	A	A	B
19	A	B	B	A	A	A	A	B
20	A	B	B	A	A	B	B	B
$n = 20$	$p = 0{,}5$	$p = 0{,}05$	$p = 0$	$p = 1{,}0$	$p = 1{,}0$	$p = 0{,}9$	$p = 0{,}7$	$p = 0{,}15$

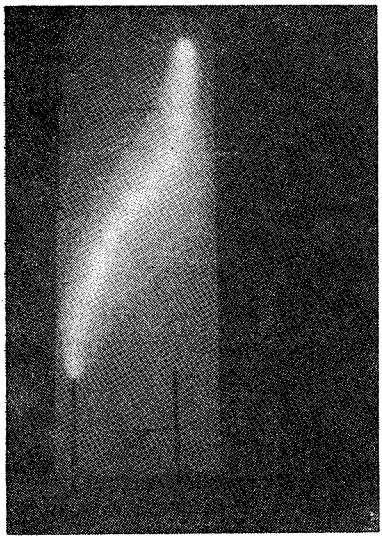
oraz dla elektrody modelującej ostrze piorunowe zawieszonej symetrycznie między elektrodami dolnymi,

c) zmiany położenia elektrody zawieszonej od $D = 0$ (zawieszenie symetryczne) do $D = 100$ cm (zmiany co 10 cm) przy jednakowej wysokości obu elektrod dolnych $h_1 = h_2 = 50$ cm i stałej ich odległości $d = 50$ cm.

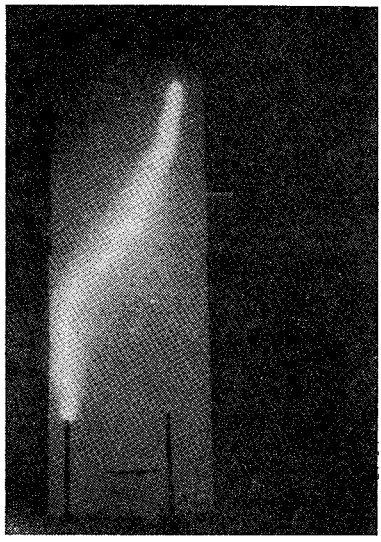
We wszystkich przypadkach stwierdzono, że wytworzenie słupa spalin nad elektrodą zwiększa prawdopodobieństwo wyładowania w tę elektrodę. Poszczególne cykle pomiarów rozpoczynano od położenia, w którym prawdopodobieństwo trafienia w każdą z dwóch zimnych elektrod było jednakowe, a kończono ustalające warunki jednakowego prawdopodobieństwa trafienia w bliższą ostrza piorunowego elektrodę zimną i w dalszą — gorącą. Przykładowe wyniki jednego z cykli pomiarów ze zmianami parametrów wg b) podano w tablicy 1.

Wobec identyczności elektrod jednakowe prawdopodobieństwo trafienia w każdą z nich w stanie zimnym obserwowano przy jednakowej odległości od ostrza piorunowego. Po zapaleniu płomyka wyrównanie prawdopodobieństwa trafienia każdej z elektrod uzyskiwano przez obniżenie elektrody gorącej lub przesunięcie wzajemne tej elektrody lub elektrody zawieszonej imitującej ostrze piorunowe. Przykładowe fotografie wyładowań do bardziej oddalonej elektrody gorącej przedstawiono na rys. 3, 4, 5, 6. Następnie po uzyskaniu takiego krytycznego położenia gaszono płomyk i starano się uzyskać wyrównanie prawdopodobieństwa trafienia w każdą z zimnych elektrod przez wydłużenie (podwyższenie) elektrody, która uprzednio była gorąca. We wszystkich cyklach badań wg zakresu a), b) i c) osiągano to przez wydłużenie elektrody o 15 ... 20 cm. Jak wynika z pomiarów temperatury nad elektrodą gorącą (p. rys. 2) odpowiada to poziomowi, na którym słup spalin osiąga temperaturę 60 ... 80°C.

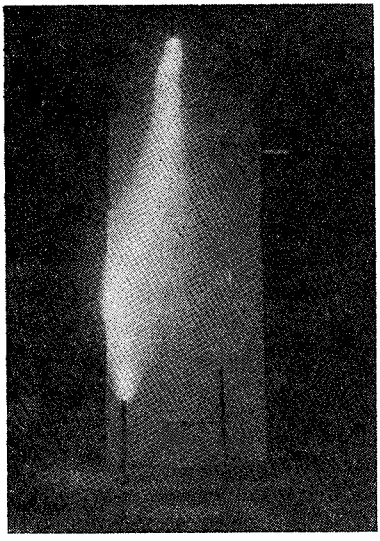
Jak wiadomo warunki prób modelowych wybiorczości wyładowań piorunowych odbiegają od występujących w rzeczywistości i trudno stąd wyciągać wnioski ilościowe co do zmiany strefy osłonowej komina czynnego w stosunku do zimnego obiektu o tych samych wymiarach geometrycznych. Jednak można niewątpliwie sformułować wniosek jakościowy, że działanie cieplne czynnego komina tak wpływa na lokalizację wyładowań piorunowych, jak zwiększenie jego wysokości. Umowne podwyższenie jest tym większe im wyższa jest temperatura spalin wychodzących z komina. Na podstawie danych dotyczących kominów wielkich elektrowni cieplnych w Polsce ocenia się temperaturę spalin wychodzących z komina na 150 ... 200°C, a dla najnowszych konstrukcji o silnie wzmożonym ciągu kominowym spotyka się nawet wyższe wartości. Można spodziewać się więc, że spadek temperatury spalin do 60 ... 80°C występuje na dość znacznej wysokości nad wylotem komina. Wysokość ta zależy niewątpliwie od szeregu parametrów, zbadanie wpływu których w laboratoryjnych warunkach modelowych jest niemożliwe — dotyczy to np. prędkości wiatru, wilgotności powietrza, rodzaju i intensywności opadów, składu chemicznego spalin, prędkości ciągu. Jednak zbadanie rozkładu temperatur nad kominem w warunkach rzeczywistych w zależności od różnych czynników eksploatacyjnych jest przy obecnym stanie wyposażenia technicznego możliwe. Na przykład można tu zaproponować metodę termowizyjną, mającą zastosowanie w szeregu innych badań eksploatacyjnych w elektroenergetyce, do której wyposażenie pomiarowo-rejestrujące produkowane jest m.in. przez szwedzką firmę AGA.



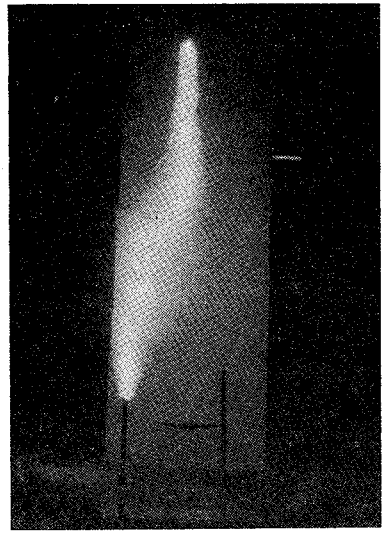
Rys. 3



Rys. 4



Rys. 5



Rys. 6

Rys. 3. Trafienie wyładowań w elektrodę gorącą bardziej odległą od ostrza piorunowego niż zimna. Przypadek stopniowej zmiany kierunku wyładowania

Rys. 4. Trafienie wyładowania w elektrodę gorącą bardziej odległą od ostrza piorunowego niż zimna. Przypadek nagłej zmiany kierunku wyładowania z prawdopodobnym wpływem wyładowań oddolnych

Rys. 5. Trafienie wyładowań w elektrodę gorącą niższą niż zimne. Przypadek stopniowej zmiany kierunku wyładowania

Rys. 6. Trafienie wyładowania w elektrodę gorącą niższą niż zimna. Przypadek skokowej zmiany kierunku wyładowania

W każdym razie uwzględnienie omówionych wyżej zjawisk przy interpretacji wyników badań piorunowych na czynnych kominach wydaje się celowe, szczególnie w przypadku użytkowania tych wyników do oceny rozmiarów stref osłonowych oraz do analizy mechanizmu wyładowań piorunowych.

Załącznik

ZASTOSOWANIE PRAWA SERII DO TESTU LOSOWOŚCI LOKALIZACJI WYŁADOWAŃ

Wyniki prowadzonych badań nad wybiorczością modelowego wyładowania piorunowego w jeden z dwóch obiektów naziemnych należą do populacji generalnej o rozkładzie dwupunktowym: trafienie w A lub trafienie w B. Przy ocenie wyników takiej statystyki jest istotne, czy populacja generalna jest losowa, tzn. czy wynik każdej kolejnej próby nie zależy od wyników poprzednich, od kolejności przeprowadzanych doświadczeń czy prawdopodobieństwo uzyskania wyniku nie zmienia się w miarę przebiegu doświadczenia. Jako test losowości dla statystyki o rozkładzie dwupunktowym zastosowano sprawdzenie za pomocą prawa serii. Prawo to określa prawdopodobną liczbę przemian[1] w serii wyników rozkładu dwupunktowego, przy której (na pewnym poziomie istotności) można wnioskować o losowości wyników.

Ten zakres liczby przemian zależy naturalnie zarówno od liczby doświadczeń jak i od prawdopodobieństwa wyniku (wyznaczanego doświadczalnie stosunkiem liczby określonych wyników do liczby wszystkich doświadczeń).

Wzory dla tych zależności przedstawiają się następująco:

$$P(K=k, N_1=n_1, N_2=n_2) = \begin{cases} 2\binom{n_1-1}{\frac{k}{2}-1}\binom{n_2-1}{\frac{k}{2}-1} p^{n_1}(1-p)^{n_2} & \frac{k}{2}\in N \\ \left[\binom{n_1-1}{\frac{k-3}{2}}\binom{n_2-1}{\frac{k-1}{2}}+\binom{n_1-1}{\frac{k-1}{2}}\binom{n_2-1}{\frac{k-3}{2}}\right] p^{n_1}(1-p)^{n_2} & \frac{k+1}{2}\in N \end{cases}$$

$$P(k>k_0)=\alpha=1-\sum_{k=1}^{k=K_0} P(K=k, N_1=n_1, N_2=n_2),$$

gdzie (w przypadku doświadczenia z wybiorczością wyładowań):

n_1 — liczba trafień w 1 obiekt naziemny,

n_2 — liczba trafień w 2 obiekt naziemny,

k — liczba przemian w serii doświadczalnej,

$n=n_1+n_2$ — liczba wyładowań w serii doświadczalnej,

N_1, N_2, K, N — odpowiednie parametry populacji generalnej,

$p=\frac{N_1}{N}$, $q=1-p=\frac{N_2}{N}$ — prawdopodobieństwo trafienia w poszczególne obiekty (z populacji generalnej).

W praktyce zastępuje się je w przybliżeniu wartością uzyskaną ze statystyki doświadczalnej

$$p\approx\frac{n_1}{n}, \quad q=1-p\approx\frac{n_2}{n},$$

α — poziom istotności.

---

[1] Liczba przemian jest to liczba grup jednakowych wyników w serii. Np. w seriach ABABAB i AAABBB mimo tego samego procentowego udziału poszczególnych wyników w pierwszej było 6 przemian, w drugiej tylko 2.

Pewne uproszczenie powyższych wzorów zachodzi wówczas, gdy przedmiotem testu są tylko serie pomiarowe o 50%-owym prawdopodobieństwie trafienia w każdy z obiektów. Wówczas $p = q = 0{,}5$, $n_1 = n_2 = \frac{n}{2}$.

Wobec skomplikowania obliczeń korzysta się na ogół z tablic [6], z których wyciąg użyteczny w rozpatrywanym przypadku podano w tablicy 2.

Graniczne liczby przemian k w serii n udarów o 50%-owym prawdopodobieństwie trafienia w każdy obiekt

Tablica 2

n	$\alpha = 0{,}025$	$\alpha = 0{,}05$	$\alpha = 0{,}95$	$\alpha = 0{,}975$
10	2	3	8	9
12	3	3	10	10
14	3	4	11	12
16	4	5	12	13
18	5	6	13	14
20	6	6	15	15
22	7	7	16	16
24	7	8	17	18
26	8	9	18	19
28	9	10	19	20
30	10	11	20	21
32	11	11	22	22
34	11	12	23	24
36	12	13	24	25
38	13	14	25	26
40	14	15	26	27

Zakresy liczby przemian w serii świadczące o losowości na poziomie $\alpha = 0{,}05$

Tablica 3

n	$p = 0{,}5$	$p = 0{,}4$ lub $p = 0{,}6$	$p = 0{,}3$ lub $p = 0{,}7$
10	3—8	3—8	2—7
20	6—15	6—14	5—12
30	10—20	11—20	9—18

W tablicy 3 zestawiono zakresy liczby przemian na poziomie $\alpha = 0{,}05$ dla różnych prawdopodobieństw trafienia w poszczególne obiekty.

Liczby przemian niższe od podanych w tablicy 3 świadczyłyby o wpływie sprzyjającym poprzednich wyładowań (rozkład pozostałego ładunku przestrzennego sprzyja wyładowaniu w tym samym kierunku co poprzednio). Liczby przemian wyższe od podanych w tablicy 3 świadczyłyby o niesprzyjającym wpływie poprzednich pomiarów (zaporowe działanie ładunku przestrzennego).

Wszystkie wyniki przeprowadzonych prób o prawdopodobieństwie w zakresie $p = 0{,}3...0{,}7$ mieściły się dla wybranej liczności $n = 20$ udarów wewnątrz podanych przedziałów zmian i to przeważnie w pobliżu środka przedziału. Por. np. przykłady z tablicy 1:

$k = 9$ przy $p = 0{,}5$ i dopuszczalnym zakresie 6—15,

$k = 9$ przy $p = 0{,}3$ i dopuszczalnym zakresie 5—12.

Stąd można wnioskować, że dobrana technika wykonywania prób i liczność prób w serii była dostateczna dla obiektywnej (opartej na losowości) oceny wybiorczości wyładowania.

LITERATURA CYTOWANA W TEKŚCIE

1. H. Ryżko, Podstawy ochrony budowli przed piorunami, PWN, 1957.
2. S. Szpor, Ochrona odgromowa, PWT, 1953 i 1955.
3. J. Suchocki, E. Wasilenko, 20 lat badań piorunowych w Polsce, Zesz. Nauk. Pol. Gdańskiej, 1970, s. 103.
4. I. S. Stiekolnikow, Priroda dlinnoj iskry, Izd. AN SSSR, 1960.
5. H. Prinz, H. Heindl, 1 MV Blitzmodellanlage in Deutschen Museum München, Elektrizitätswirtschaft, 1954, s. 266.
6. T. Czechowicz, M. Fisz, Tablice statystyczne, PWN, 1957.
7. F. Rühling, Modelluntersuchungen über den Schutzraum und ihre Bedeutung für Gebäudeblitzschutz, Bull. SEV, 1972, z. 10, s. 522.
8. PN-58/N-01051, Rachunek prawdopodobieństwa i statystyka matematyczna. Terminy, określenia i symbole, Wyd. II. Wyd. Norm., Warszawa 1964.

J. J. ZIELIŃSKI

MODEL TESTS OF LIGHTNING STROKES TO ACTIVE CHIMNEYS OR OTHER HEAT SOURCES

Summary

Analysis of thermal conditions prevailing over active chimneys. Suggested simulation of these conditions in model lightning tests by substituting the flow of combustion gases with flow of combustion gases from a flame of urban gas fed into tube electrodes. Model investigation of the temperature distribution and the localisation of lightning strokes to cold and hot electrodes. It is displayed that the flow of hot combustion gases causes the same change of this localisation as a heightning of the electrode simulating the chimney, to a level corresponding to a combustion gas temperature of 60—80°C. Conclusions concerning the need to complete up-to-date evaluations of the protection zone of high buildings basing on chimney investigations. The appendix gives a test of the random character of the localisation of discharges resulting from the law of series for a bi-point distribution, thus, justifying the correctness of the choice of the number and of the repetition rate of voltage applications.

J. J. ZIELIŃSKI

ESSAIS MODELLANT LES COUPS DE FOUDRE SUR HAUTES-CHEMINÉES ACTIVES OU SUR AUTRES SOURCES DE CHALEURS

Résumé

Analyse des conditions thérmiques au-dessus des cheminées industrielles actives. Suggestion de réproduire ces conditions lors des essais de laboratoire en remplacant le courant des gaz de combustion par celui d'une flamme de gaz d'éclairage alimentant des électrodes tubulaires. Essais de modèle de la répartition des

températures et de la localisation des coups de foudre á des electrodes froides et chaudes. On démontre que le courant des gaz de combustion chauds provoque le même changement de la localisation des coups de foudre qu'une élévation de l'électrode modèle au niveau correspondant á une température des gaz de combustion de 60—80°C. Conclusions concernant la nécessité de compléter la présente évaluation de zones de protection de haute batiments á base des études des cheminées. L'annexe présente un test du hasard de la localisation des décharges résultant de la loi de série pour des distributions bi-points, justifiant de cette manière le choix correcte du nombre et de la fréquence des applications de la tension.

J. J. ZIELIŃSKI

MODELLUNTERSUCHUNGEN VON BLITZSCHLÄGEN IN TÄTIGE SCHORNSTEINE ODER IN ANDERE WARMQUELLEN

Zusammenfassung

Eine Analyse thermischer, über tätigen Schornsteinen herrschender Bedingungen. Ein Vorschlag zur Abbildung dieser Bedingungen bei Blitzschlag-Modelluntersuchungen durch Auswechslung des Verbrennungsgasflusses durch den Verbrennungsfluß eines Lichtgasflämmchens auf Rohrelektroden. Modelluntersuchungen von Temperaturverteilung und Lokalisierung der Blitzstöße in kalten und heißen Elektroden. Beweisführung, daß heiße Verbrennungsgassäulen die gleiche Änderung von Blitzstoßlokalisierung ausüben, wie die Erhöhung der Modellelektrode des Schornsteines zu einem der Verbrennungsgastemperatur von 60—80°C entsprechenden Niveau.

Vorschläge betreffend die Zweckmäßigkeit einer Ergänzung der bisherigen Auswertung der Abschirmzone hoher Gebäude auf Grund von Schornsteinbeobachtungen.

In der Anlage wurde ein auf Basis des Serienprinzips durchgeführter Test über die Zufälligkeit der Blitzschlaglokalisierung gebracht. Dieser Test begründet die Häufigkeit und Zahl der nötigen Untersuchungen.

Е. Я. ЗЕЛИНЬСКИ

МОДЕЛЬНЫЕ ОПЫТЫ УДАРОВ МОЛНИИ В ДЕЙСТВУЮЩИЕ ДЫМООТВОДЫ ИЛИ ДРУГИЕ ИСТОЧНИКИ ТЕПЛА

Резюме

Статья содержит: анализ термических условий над действующими дымоотводами. Предложение отображения этих условий при модельных опытах молниевых разрядов путем замещения тяги дымовых газов тягой газов сгорания пламени светильного газа подведенного к трубчатым электродам. Модельные опыты распределения температур, а также избирательности молниевых разрядов на накаленных и холодных электродах. Обнаружено, что столб накаленных газов вызывает такое изменение избирательности молниевых разрядов, как повышение высоты электрода моделирующего дымоотвод до высоты соответствующей температуре газов 60—80°C. Заключения касающиеся целесообразности дополнения оценок зон защиты высоких построек с точки зрения учета влияния дымоотводов. В приложении проведен тест случайности локализации разрядов вытекающей из закона серии для двоеточечного распределения доказывающий правильность избирания количества и частоты испытаний.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 1, ss. 163-170

621.365.55

Model cyfrowy dielektrycznego suszenia ścianki cylindrycznej

TADEUSZ NIEWIEROWICZ (ŁÓDŹ)

Centralne Laboratorium Przemysłu Lniarskiego, Łódź

Otrzymano 29.9.1972

W artykule przekształcono równania wymiany masy i ciepła (1) sformułowane przez A. W. Łykowa oraz uwzględniono w nich obecność wewnętrznych źródeł ciepła. Następnie sprowadzono te równania do postaci wektorowej (9) i przestawiono sposób ich rozwiązania dla przypadku dielektrycznego suszenia ścianki cylindrycznej.

W końcu artykułu umieszczono praktyczne wskazania odnośnie możliwości wykorzystania przedstawionego modelu, a także zwrócono uwagę na jego zalety i wady.

1. WSTĘP

W wielu przypadkach konieczna jest znajomość zjawisk zachodzących wewnątrz materiałów higroskopijnych kapilarno-porowatych intensywnie suszonych prądami wielkiej częstotliwości. Wtedy procesy wymiany masy i ciepła należy rozpatrywać razem, ponieważ na wymianę masy ma istotny wpływ wymiana ciepła (gradienty temperatur osiągają znaczne wartości). Równania ogólne dotyczące omawianych zjawisk dla przypadku suszenia konwekcyjnego sformułował A. W. Łykow [3].

W niniejszej pracy przekształcono równania A. W. Łykowa uwzględniając w nich obecność wewnętrznych źródeł ciepła i przedstawiono sposób ich rozwiązania dla przypadku suszenia ścianki cylindrycznej. W rozważaniach przyjęto, że wszystkie punkty powierzchni ograniczającej ściankę cylindryczną znajdują się w jednakowych warunkach oraz że zewnętrzny opór wymiany wilgoci jest nieznaczny w porównaniu do oporów wymiany wilgoci w ciele suszonym. W końcu artykułu umieszczono praktyczne wskazania dotyczące możliwości wykorzystania przedstawionego modelu, a także zwrócono uwagę na jego zalety i wady.

2. RÓWNANIA PRZENOSZENIA MASY I CIEPŁA

Wewnętrzną wymianę masy i ciepła w przypadku bezźródłowego pola temperatur i wilgoci opisać można następującym układem równań różniczkowych [3]:

$$\begin{cases} \dfrac{\partial t(r,\tau)}{\partial \tau} = a_q\left[\dfrac{\partial^2 t(r,\tau)}{\partial r^2} + \dfrac{\Gamma}{r}\dfrac{\partial t(r,\tau)}{\partial r}\right] + \dfrac{\varepsilon\varrho}{c_q}\dfrac{\partial u(r,\tau)}{\partial \tau}, \\ \dfrac{\partial u(r,\tau)}{\partial \tau} = a_m\left[\dfrac{\partial^2 u(r,\tau)}{\partial r^2} + \dfrac{\Gamma}{r}\dfrac{\partial u(r,\tau)}{\partial r}\right] + a_m\delta\left[\dfrac{\partial^2 t(r,\tau)}{\partial r^2} + \dfrac{\Gamma}{r}\dfrac{\partial t(r,\tau)}{\partial r}\right], \end{cases} \tag{1}$$

gdzie:

t — temperatura,
τ — czas,
u — zawartość wody,
r — współrzędna uogólniona,
C_q — ciepło właściwe,
a_q — współczynnik dyfuzji cieplnej,
a_m — współczynnik dyfuzji masy (wilgoci),
δ — współczynnik termogradientny,
ϱ — ciepło właściwe przekształcenia fazowego,
ε — współczynnik przemiany fazowej,
Γ — współczynnik kształtu ciała suszonego (dla płyty $\Gamma = 0$, dla cylindra $\Gamma = 1$, dla kuli $\Gamma = 2$).

Po przekształceniu równań (1) dla ścianki cylindrycznej otrzymamy układ równań o postaci:

$$\begin{cases} \dfrac{\partial t}{\partial \tau} = \dfrac{a_{11}}{r}\dfrac{\partial}{\partial r}\left(r\dfrac{\partial t}{\partial r}\right) + \dfrac{a_{12}}{r}\dfrac{\partial}{\partial r}\left(r\dfrac{\partial u}{\partial r}\right), \\ \dfrac{\partial u}{\partial \tau} = \dfrac{a_{21}}{r}\dfrac{\partial}{\partial r}\left(r\dfrac{\partial t}{\partial r}\right) + \dfrac{a_{22}}{r}\dfrac{\partial}{\partial r}\left(r\dfrac{\partial u}{\partial r}\right), \end{cases} \tag{2}$$

gdzie a_{ik} — współczynniki przewodności:

$$a_{11} = a_q + \frac{\varepsilon\varrho}{c_q} a_m \delta; \quad a_{12} = \frac{\varepsilon\varrho}{c_q} a_m;$$

$$a_{21} = a_m \delta; \quad a_{22} = a_m.$$

Równania (2) dotyczą bezźródłowego pola temperatur. W przypadku ogólnym podczas suszenia dielektrycznego każdy punkt ścianki cylindrycznej jest źródłem ciepła o objętościowej gęstości mocy $g(r, \tau)$. Do dalszych rozważań przyjmujemy założenie równomierności rozkładu gęstości mocy grzejnej w rozpatrywanej ściance cylindrycznej, czyli $g = g(\tau)$. Założenie to ma sens, gdyż przy doborze odpowiednich parametrów dielektrycznego układu grzejnego nierównomierność gęstości mocy będzie praktycznie tak mała, że można ją pominąć. Uwzględniając dodatkowe źródło ciepła w równaniach (2) otrzymamy układ równań opisujący przenoszenie ciepła i wilgoci w ściance cylindrycznej w przypadku suszenia pojemnościowego [1] [4]:

$$\begin{cases} \dfrac{\partial t}{\partial \tau} = \dfrac{a_{11}}{r}\dfrac{\partial}{\partial r}\left(r\dfrac{\partial t}{\partial r}\right) + \dfrac{a_{12}}{r}\dfrac{\partial}{\partial r}\left(r\dfrac{\partial u}{\partial r}\right) + a_{13} g, \\ \dfrac{\partial u}{\partial \tau} = \dfrac{a_{21}}{r}\dfrac{\partial}{\partial r}\left(r\dfrac{\partial t}{\partial r}\right) + \dfrac{a_{22}}{r}\dfrac{\partial}{\partial r}\left(r\dfrac{\partial u}{\partial r}\right), \end{cases} \tag{3}$$

gdzie:

$a_{13} = \dfrac{1}{C_q \gamma}$; γ — gęstość ciała.

W celu uproszczenia zagadnienia załóżmy, że:

1. współczynniki przewodności a_{ik} mają wartość stałą,
2. zewnętrzny opór wymiany wilgoci jest nieznaczny w porównaniu do oporów wymiany wilgoci w ciele suszonym,
3. zewnętrzne warunki suszenia nie ulegają zmianie.

Ponadto przyjmujemy, że funkcje $t(r, \tau)$ oraz $u(r, \tau)$ opisujące stan ciała suszonego spełniają następujące warunki początkowe i brzegowe:

$$\begin{aligned} &u(r, 0) = u_p; \\ &t(r, 0) = t_p; \end{aligned} \quad \text{dla} \quad r \in (r_w, r_z); \tag{4}$$

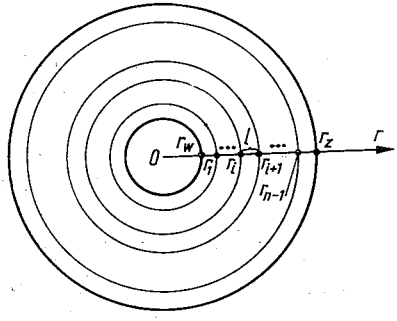
$$\begin{aligned} &u(r_w, \tau) = \varkappa_1 u_r, \\ &u(r_z, \tau) = u_r, \\ &\left(\frac{\partial t}{\partial r}\right)_{r=r_w} = \varkappa_2 h[t(r_w, \tau) - t_p], \quad \text{dla} \quad \tau \in (0, \infty) \\ &\left(\frac{\partial t}{\partial r}\right)_{r=r_z} = -h[t(r_z, \tau) - t_p], \end{aligned} \tag{5}$$

gdzie:

$\varkappa_1$ i $\varkappa_2$ — współczynniki korekcyjne uwzględniające utrudnienie wymiany ciepła i wilgoci wewnętrznej ścianki cylindrycznej,

h — współczynnik wymiany ciepła,

u_r — równoważna zawartość wilgoci.



Rys. 1. Podział ścianki cylindrycznej na warstwy

Równania (3) rozwiązujemy metodą różnicową [1] [2]. W tym celu podzielmy grubość ścianki cylindrycznej na n części. Współrzędnymi punktów podziału są:

$$r_i = r_w + il \quad \text{dla} \quad i = 0, 1, 2, \ldots, n,$$

przy czym $l > 0$ jest przyrostem zmiennej niezależnej r

$$l = \frac{r_z - r_w}{n}.$$

Przechodzimy od równań różniczkowych (3) do równań różnicowych, zastępując pochodne cząstkowe odpowiednimi ilorazami różnicowymi drugiego rzędu:

$$\frac{\partial}{\partial r}\left(r\frac{\partial t}{\partial r}\right) \approx \frac{1}{l}\left[\left(r_i+\frac{l}{2}\right)\frac{t_{i+1}-t_i}{l} - \left(r_i-\frac{l}{2}\right)\frac{t_i-t_{i-1}}{l}\right],$$
$$\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) \approx \frac{1}{l}\left[\left(r_i+\frac{l}{2}\right)\frac{u_{i+1}-u_i}{l} - \left(r_i-\frac{l}{2}\right)\frac{u_i-u_{i-1}}{l}\right]. \tag{6}$$

Podstawiając (6) do (3) nasz układ równań przyjmie postać

$$\frac{dt_i}{d\tau} \approx \frac{a_{11}}{l^2}\left(\frac{d_i}{r_i}t_{i-1}-2t_i+\frac{d_{i+1}}{r_i}t_{i+1}\right) + \frac{a_{12}}{l^2}\left(\frac{d_i}{r_i}u_{i-1}-2u_i+\frac{d_{i+1}}{r_i}u_{i+1}\right)+a_{13}g,$$
$$\frac{du_i}{d\tau} \approx \frac{a_{21}}{l^2}\left(\frac{d_i}{r_i}t_{i-1}-2t_i+\frac{d_{i+1}}{r_i}t_{i+1}\right) + \frac{a_{22}}{l^2}\left(\frac{d_i}{r_i}u_{i-1}-2u_i+\frac{d_{i+1}}{r_i}u_{i+1}\right), \tag{7}$$

gdzie:

$$d_i = r_i-\frac{l}{2}, \qquad \text{dla} \quad i = 1, 2, \ldots, n-1.$$

Wielkości temperatur i zawartości wody w wewnętrznej warstwie ścianki cylindrycznej (dla $i = 0$) oraz zewnętrznej (dla $i = n$) zależą od zadanych warunków brzegowych. W naszym przypadku

$$u_0 = \varkappa_1 u_r, \qquad u_n = u_r,$$
$$t_0 = \frac{t_1+\varkappa_2 hlt_p}{1+\varkappa_2 hl}, \qquad t_n = \frac{t_{n-1}+hlt_p}{1+hl}. \tag{8}$$

Uwzględniając (8) w (7) otrzymamy $n-1$ równań opisujących zmianę temperatury oraz $n-1$ równań opisujących zmianę zawartości wody podczas procesu suszenia w poszczególnych warstwach ścianki cylindrycznej. Otrzymane równania możemy zapisać w postaci wektorowej, przypisując każdej warstwie funkcję jednej zmiennej $V_i(\tau) = V(r_i, \tau)$, gdzie V_i jest współrzędną wektora stanu procesu suszenia. Wektor ten zawiera współrzędne opisujące temperaturę i zawartość wody w poszczególnych warstwach ścianki cylindrycznej

$$\mathbf{v} = \begin{bmatrix}\mathbf{t}\\ \mathbf{u}\end{bmatrix},$$

gdzie

$$\mathbf{t} = \begin{bmatrix}t_1\\ t_2\\ \vdots\\ t_{n-1}\end{bmatrix}, \qquad \mathbf{u} = \begin{bmatrix}u_1\\ u_2\\ \vdots\\ u_{n-1}\end{bmatrix}.$$

Uwzględniając powyższe oznaczenia, układ $2n-2$ równań procesu suszenia przedstawić możemy w postaci:

$$\frac{d\mathbf{v}}{d\tau} = \mathbf{A}\mathbf{v}+\mathbf{m}+\mathbf{n}, \tag{9}$$

gdzie:

$\mathbf{v}$ — wektor stanu procesu,

$\mathbf{m}$ — wektor sterowania,

$\mathbf{n}$ — wektor zakłóceń,

$\mathbf{A}$ — macierz procesu.

Wektor sterowania zawiera zmienne niezależne $g(\tau)$ i u_r, które są zmiennymi wejściowymi sterującymi procesem suszenia. Zmianę $g(\tau)$ dokonuje się np. zmianą napięcia generatora wielkiej częstotliwości, natomiast u_r możemy zmieniać za pośrednictwem zmiany zewnętrznych warunków suszenia, tj. wilgotności i temperatury otoczenia, prędkości przedmuchu powietrza itp.

Wektor zakłóceń uwzględnia wpływ temperatury otoczenia poprzez wewnętrzną i zewnętrzną warstwę na proces suszenia ścianki cylindrycznej, nie uwzględnia natomiast zakłóceń przypadkowych.

W naszym przypadku wektory sterowania oraz zakłóceń przyjmują postać

$$\mathbf{m} = \begin{bmatrix} a_{12}b_1u_r+a_{13}g \\ a_{13}g \\ \vdots \\ a_{13}g \\ a_{12}b_2u_r+a_{13}g \\ a_{22}b_1u_r \\ 0 \\ \vdots \\ 0 \\ a_{22}b_2u_r \end{bmatrix}, \qquad \mathbf{n} = \begin{bmatrix} a_{11}b_3 \\ 0 \\ \vdots \\ 0 \\ a_{11}b_4 \\ a_{21}b_3 \\ 0 \\ \vdots \\ 0 \\ a_{21}b_4 \end{bmatrix}, \tag{10}$$

gdzie:

$$b_1 = \frac{d_1\varkappa_1}{r_1l^2}, \qquad b_3 = \frac{d_1\varkappa_2ht_p}{r_1l(1+\varkappa_2hl)},$$

$$b_2 = \frac{d_n}{r_{n-1}l^2}, \qquad b_4 = \frac{d_nht_p}{r_{n-1}l(1+hl)},$$

natomiast macierz procesu będzie quasi-macierzą składającą się z czterech macierzy Jakobiego (trójdiagonalnych) stopnia $n-1$

$$\mathbf{A} = \frac{1}{l^2}\begin{bmatrix} \mathbf{A}_1\mathbf{A}_2 \\ \mathbf{A}_3\mathbf{A}_4 \end{bmatrix}, \tag{11}$$

gdzie:

$$\mathbf{A}_1 = \begin{bmatrix} a_{11}b_5 \dfrac{a_{11}d_2}{r_1} & & & & \\ \dfrac{a_{11}d_2}{r_2} - 2a_{11}\dfrac{a_{11}d_3}{r_2} & & & & \\ & \dfrac{a_{11}d_3}{r_3} - 2a_{11}\dfrac{a_{11}d_4}{r_3} & \cdot & & \\ & & \cdot \quad \cdot & \cdot & \\ & & \cdot \quad \cdot & \cdot & \cdot \\ & & & \dfrac{a_{11}d_{n-1}}{r_{n-1}} & a_{11}b_6 \end{bmatrix} \begin{matrix} b_5 = \dfrac{d_1}{r_1(1+\varkappa_2 hl)} - 2 \\ b_6 = \dfrac{d_n}{r_{n-1}(1+hl)} - 2 \end{matrix}$$

$$\mathbf{A}_2 = \begin{bmatrix} -2a_{12}\dfrac{a_{12}d_2}{r_1} & & & & \\ \dfrac{a_{12}d_2}{r_2} - 2a_{12}\dfrac{a_{12}d_3}{r_2} & & & & \\ & \dfrac{a_{12}d_3}{r_3} - 2a_{12}\dfrac{a_{12}d_4}{r_3} & \cdot & & \\ & & \cdot \quad \cdot & \cdot & \\ & & \cdot \quad \cdot & \cdot & \cdot \\ & & & \dfrac{a_{12}d_{n-1}}{r_{n-1}} - 2a_{12} \end{bmatrix},$$

$$\mathbf{A}_3 = \begin{bmatrix} a_{21}b_5 \dfrac{a_{21}d_2}{r_1} & & & & \\ \dfrac{a_{21}d_2}{r_2} - 2a_{21}\dfrac{a_{21}d_3}{r_2} & & & & \\ & \dfrac{a_{21}d_3}{r_3} - 2a_{21}\dfrac{a_{21}d_4}{r_3} & \cdot & & \\ & & \cdot \quad \cdot & \cdot & \\ & & \cdot \quad \cdot & \cdot & \cdot \\ & & & \dfrac{a_{21}d_{n-1}}{r_{n-1}} & a_{21}b_6 \end{bmatrix},$$

$$\mathbf{A}_4 = \begin{bmatrix} -2a_{22} & \frac{a_{22}d_2}{r_1} & & & & \\ \frac{a_{22}d_2}{r_2} & -2a_{22} & \frac{a_{22}d_3}{r_2} & & & \\ & \frac{a_{22}d_3}{r_3} & -2a_{22} & \frac{a_{22}d_4}{r_3} & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot \\ & & & & \frac{a_{22}d_{n-1}}{r_{n-1}} & -2a_{22} \end{bmatrix}.$$

Wektor stanu procesu suszenia spełnia warunek początkowy $\mathbf{v}(0) = \mathbf{v}_0$. W naszym przypadku rozwiązanie równania wektorowego (9) ma postać [5]:

$$\mathbf{v}(\tau) = \int_0^{\tau} \exp[\mathbf{A}(\tau-\theta)][\mathbf{m}(\theta)+\mathbf{n}]\,d\theta + \mathbf{v}_0 \exp[\mathbf{A}\tau]. \tag{12}$$

3. ZAKOŃCZENIE

Przedstawiony wyżej (9) model cyfrowy dielektrycznego suszenia ścianki cylindrycznej w postaci układu równań różniczkowo-różnicowych zaprogramowano na maszynę cyfrową ODRA 1013, przyjmując jako sposób rozwiązania metodę Rungego i Kutty.

Zaletą opisanego modelu cyfrowego jest jego znaczna elastyczność pozwalająca na zmianę współczynników przewodzenia w czasie poszczególnych kroków obliczeniowych, a tym samym jeszcze dokładniejsze zbliżenie do stanu rzeczywistego oraz możliwość modelowania procesów o różnym początkowym rozkładzie wilgoci i temperatur panujących wewnątrz ciała suszonego. Przedstawiony model wykorzystać można również do optymalizacji procesów suszenia przy zadanym kryterium [4], jak też do badania wpływu ograniczeń nakładanych na sygnał sterujący lub składowe wektora stanu na czas suszenia.

Zasadniczą wadą modelu jest długi czas obliczeń w przypadku podziału suszonej ścianki cylindrycznej na dużą (> 10) ilość warstw.

LITERATURA CYTOWANA W TEKŚCIE

1. E. Kącki, Termokinetyka, WNT, Warszawa 1966.
2. E. Kącki, Równania różniczkowe cząstkowe w elektrotechnice, WNT, Warszawa 1971.
3. A. Łykow i J. A. Michajłow, Teoria pierenosa energii i wieszczestwa, Mińsk 1959.
4. T. Niewierowicz, Optymalizacja procesu pojemnościowego suszenia przędzy, Arch. Budowy Maszyn, nr 3, 1972.
5. J. T. Tou, Nowoczesna teoria sterowania, WNT, Warszawa 1967.

T. NIEWIEROWICZ

DIGITAL MODEL OF DIELECTRIC DRYING OF CYLINDRICAL WALL

Summary

In the article were transformed equations of mass and heat (1) exchange, formulated by A. W. Łykow. In these equations the presence of internal heat sources has been taken into consideration. Further the bespoken equations were reduced to vector form (9) and manner of their solution was represented for the case of dielectric drying of cylindrical wall.

In the end of article were located practical indications concerning possibilities of represented model's using and it was taken into account its advantages as well as defects.

T. NIEWIEROWICZ

ZAHLENMODELL FÜR DIELEKTRISCHES TROCKEN EINER ZYLINDRISCHEN WAND

Zusammenfassung

In der Arbeit wurden die Massenauswechslungs- und Wärmegleichungen (1) von A. W. Łykow umgestaltet, wobei die Anwesenheit innerer Wärmequellen berücksichting worden ist. Danach wurden diese Gleichungen in Form von Vektoren dargestellt (9), und am Beispiel dielektrischen Trockens einer zylindrischen Wand ihre Lösung vorgetragen.

Im letzten Teil des Artikels wurden praktische Hinweise für die Auswertung des vorgeschlagenen Modells gebracht und seine Vor- und Nachteile erwogen.

Т. НЕВЕРОВИЧ

ЦИФРОВАЯ МОДЕЛЬ ДИЭЛЕКТРИЧЕСКОЙ СУШКИ ЦИЛИНДРИЧЕСКОЙ СТЕНКИ

Резюме

В статье преобразовано уравнения тепло- и массопереноса (1) предложенные А. В. Лыковым, а также учтено присутствие внутренних источников тепла. Далее эти уравнения приведены к матричной форме (9) и представлен способ их решения для случая диэлектрической сушки цилиндрической стенки.

В заключение статьи рассмотрены практические возможности использования представленной модели, а также описаны ее преимущества и недостатки.

ROZPRAWY EELEKTROTECHNICZNE, 1974 20, z. 1, ss. 171-193

621.372.22

Funkcje strat torów niejednorodnych złożonych z odcinków jednorodnych o charakterystyce typu Czebyszewa i Butterwortha

FRANCISZEK KAMIŃSKI (WARSZAWA)
Instytut Fizyki PAN, Zakład Magnetyków

Otrzymano 20.7.1972

Przedstawiono zarys postępowania przy poszukiwaniu funkcji strat toru niejednorodnego złożonego z bezstratnych odcinków jednorodnych na podstawie ustalonych wymagań. W szczególności wykazano potrzebę wprowadzenia wskaźnika technologicznej realizowalności układu celem uzyskania jednoznaczności rozwiązania. Sformułowano i rozwiązano problem optymalizacji pewnych parametrów charakterystycznych toru. Na podstawie przedstawionych zasad postępowania wyprowadzono ogólną postać funkcji strat typu Czebyszewa i Butterwortha. Wyniki pracy znajdują zastosowanie w syntezie filtrów elektromechanicznych i niektórych układów mikrofalowych.

1. WSTĘPNA CHARAKTERYSTYKA ZAGADNIENIA

Typowe wymagania dotyczące właściwości transmisyjnych oraz warunków pracy czwórnika filtrującego lub dopasowującego można ująć następująco (rys. 1):

1. $|T_{21}(j\omega)| \leqslant h$ dla $f_{-p} \leqslant f \leqslant f_p$,
2. $|T_{21}(j\omega)| \geqslant H_{-t}$ dla $f'_{-t} \leqslant f \leqslant f_{-t}$,
3. $|T_{21}(j\omega)| \geqslant H_t$ dla $f_t \leqslant f \leqslant f_t$,
4. opory pracy na wejściu i wyjściu układu — R_1, R_2,
5. kształt charakterystyki (krzywa aproksymacji) w pasmie przepustowym (np. typu Czebyszewa, typu Butterwortha).

Sens oznaczeń:

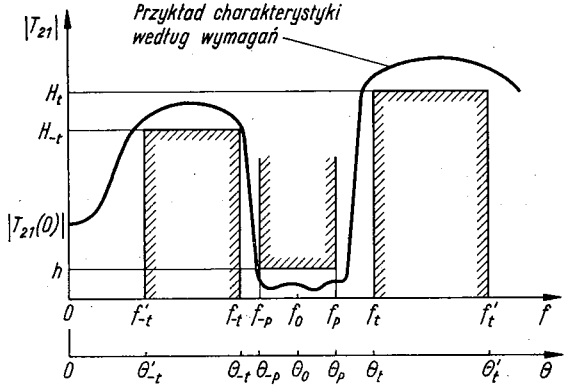
f_{-p}, f_p — częstotliwości graniczne pasma przepustowego,
f'_{-t}, f_{-t} — częstotliwości graniczne dolnego pasma tłumieniowego użytkowego,
f_t, f'_t — częstotliwości graniczne górnego pasma tłumieniowego użytkowego,
$\omega = 2\pi f$ — pulsacja,
T_{21} — funkcja filtracji układu [2], [5], [6].

Oczywiście w przypadku konkretnym niektóre z podanych punktów wymagań mogą pozostać niesprecyzowane.

W niniejszej pracy przeanalizuje się zagadnienie, jak na podstawie powyższych wymagań można wyznaczyć funkcję strat $L(s)$ toru niejednorodnego złożonego z bezstratnych odcinków jednorodnych przy obciążeniu rzeczywistym, jeżeli kształt charaktery-

styki ma być typu Czebyszewa i Butterwortha. Problematyka ta była już częściowo podejmowana w pracach [8], [9]; obecnie zostanie ona pogłębiona, a niektóre z poprzednio przedstawionych wyników — zweryfikowane.

Otóż jak wykazuje analiza, przytoczone wymagania nie są wystarczające na to, aby jednoznacznie wyznaczyć tor niejednorodny badanej klasy o pożądanych właściwościach transmisyjnych. Dla uzyskania jednoznaczności rozwiązania niezbędne są dodatkowe dane, uzyskiwane na podstawie odrębnej analizy. Obecnie omówi się w zarysie istotę tego zagadnienia.



Rys. 1. Typowe wymagania dotyczące właściwości transmisyjnych układu

W dalszych rozważaniach przyjmuje się określanie toru niejednorodnego złożonego z bezstratnych odcinków jednorodnych przy obciążeniu rzeczywistym mianem toru typu N. W tej klasie torów wyodrębnia się grupę torów typu N_o, do której należą tory niejednorodne złożone z bezstratnych odcinków jednorodnych o jednakowej długości elektrycznej przy obciążeniu rzeczywistym.

Właściwości transmisyjne toru typu N określa się przez podanie jego funkcji strat [2], [6]

$$L(p) \overset{\text{df}}{=} T_{11}(p)T_{11}(-p) = 1+T_{21}(p)T_{21}(-p), \tag{1}$$

przy czym na osi Im p obowiązuje

$$L(j\omega) = |T_{11}(j\omega)|^2 = 1+|T_{21}(j\omega)|^2, \tag{2}$$

gdzie:

$T_{11}(p)$ — funkcja transmisji toru [2], [5], [6],

$T_{21}(p)$ — funkcja filtracji toru,

$p = \delta+j\omega$ — zmienna zespolona.

W przypadku toru typu N_o można zastąpić zmienną p przez zmienną s zgodnie z relacją

$$s = \tau p = \sigma+j\Theta, \quad \Theta = 2\pi\frac{l}{\lambda} = \tau\omega, \tag{3}$$

gdzie:

τ — pewien współczynnik dodatni, zależny od parametrów fizycznych podstawowego odcinka jednorodnego w torze[1],

Θ — przesuwność falowa (długość elektryczna) odcinka elementarnego,

l — długość odcinka elementarnego,

λ — długość fali w torze.

A zatem w przypadku toru typu N_o obowiązują zależności [2], [6]

$$L(s) = T_{11}(s)T_{11}(-s) = 1+T_{21}(s)T_{21}(-s) \tag{4}$$

oraz

$$L(j\Theta) = |T_{11}(j\Theta)|^2 = 1+|T_{21}(j\Theta)|^2. \tag{5}$$

Po tych wstępnych uwagach wprowadzających można przystąpić do zreferowania istoty zagadnienia.

Niech dany będzie tor typu N_o, złożony z n odcinków jednorodnych, o funkcji strat $L(s_1)$ spełniającej omówione wymagania. Zgodnie z wynikami pracy [6] funkcja $L(s_1)-1 = T_{21}(s_1)T_{21}(-s_1)$ tego toru jest wielomianem wykładniczym o współczynnikach rzeczywistych, parzystym względem $z' = \operatorname{ch} s_1$ i nieujemnym na osi urojonej $\operatorname{Im} s_1$. A zatem

$$Q(s_1) \overset{\text{df}}{=} T_{21}(s_1)T_{21}(-s_1) = b_{2n}\operatorname{ch}^{2n}s_1+b_{2n-2}\operatorname{ch}^{2n-2}s_1+ \ldots +b_2\operatorname{ch}^2 s_1+b_o\,, \tag{6}$$

$$Q(j\Theta_1) = |T_{21}(j\Theta_1)|^2 \geqslant 0,$$

gdzie współczynniki b_{2i} są rzeczywiste.

Funkcja $Q(s_1)$ jest funkcją okresową o okresie $j\pi$ i o właściwości

$$Q\left(j\frac{\pi}{2}-s_1\right) = Q\left(j\frac{\pi}{2}+s_1\right). \tag{7}$$

Stąd wniosek, że dla uzyskania pełnej informacji o przebiegu funkcji $Q(s_1)$ na osi $\operatorname{Im} s_1 = \Theta_1$ wystarczy znać jej przebieg w przedziale $0 \leqslant \Theta_1 \leqslant \frac{\pi}{2}$.

Dla $s_1 = 0$ zachodzi równość [5]

$$Q(0) = \frac{1}{4}\left(\sqrt{\frac{R_2}{R_1}} - \sqrt{\frac{R_1}{R_2}}\right)^2, \tag{8}$$

gdzie R_1, R_2 — opory obciążenia na wejściu i wyjściu toru.

Zgodnie ze wzorem (3) między zmiennymi s_1 i p istnieje związek $s_1 = \tau_1 p$, $\Theta_1 = \tau_1\omega$, gdzie współczynnik τ_1 jest znany.

Funkcja $Q(s)$ typu (6) daje się przekształcić w wielomian algebraiczny $G(z)$ za pomocą odwzorowania

$$\operatorname{ch}^2 s = \mu z+\eta, \qquad z = x+jy = \frac{\operatorname{ch}^2 s-\eta}{\mu}, \tag{9}$$

[1] Z (3) widać, że poszczególnym torom N_0 mogą odpowiadać odmienne zmienne s, co w przypadku porównywania charakterystyk zaznacza się przez wprowadzenie odpowiednich indeksów, jak np. $s_1 = \tau_1 p_1$, $s_k = \tau_k p$ itp.

którego uwzględnienie we wzorze (6) przynosi

$$Q(s) = b_{2n}(\mu z+\eta)^n+b_{2n-2}(\mu z+\eta)^{n-1}+ \ldots +b_2(\mu z+\eta)+b_o = \\ = g_n z^n+g_{n-1}z^{n-1}+ \ldots +g_1 z+g_o = G(z), \tag{10}$$

gdzie współczynniki g_i są rzeczywiste dla μ, η rzeczywistych; ich wartości wynikają jednoznacznie z podanej równości.

Jeżeli przyjąć $\mu > 0$, η — parametr rzeczywisty, to oś urojona $\mathrm{Im}\, s = \Theta$ przejdzie w wyniku odwzorowania (9) w odcinek osi rzeczywistej $\mathrm{Re} z = x$, przy czym przedział $0 \leqslant \Theta \leqslant \frac{\pi}{2}$ zostanie odwzorowany na przedział $-\frac{\eta}{\mu} \leqslant x \leqslant \frac{1-\eta}{\mu}$. Wartości parametrów μ, η można dobrać tak, aby zadany przedział $[\Theta_{-p}, \Theta']$ na osi $\mathrm{Im} s$ został odwzorowany na przedział $[-1, 1]$ na osi $\mathrm{Re} z = x$. W tym celu należy przyjąć $\mu = \mu_o$, $\eta = \eta_o$, gdzie μ_o, η_o są pierwiastkami następującego układu równań:

$$\frac{\cos^2\Theta_{-p}-\eta}{\mu} = 1, \qquad \frac{\cos^2\Theta'-\eta}{\mu} = -1,$$

skąd

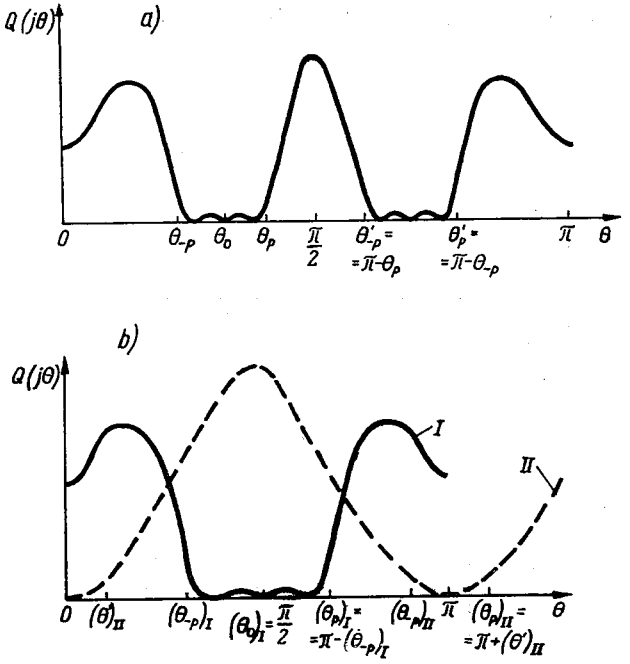
$$\mu_o = \frac{\cos^2\Theta_{-p}-\cos^2\Theta'}{2} = \frac{1}{2}\sin(\Theta'+\Theta_{-p})\sin(\Theta'-\Theta_{-p}), \\ \eta_o = \frac{\cos^2\Theta_{-p}+\cos^2\Theta'}{2} = \frac{1}{2}[1+\cos(\Theta'+\Theta_{-p})\cos(\Theta'-\Theta_{-p})]. \tag{11}$$

Z uwagi na zależność (7) wartości Θ_{-p}, Θ' we wzorze (11) dobiera się tak, aby spełnić warunek $0 \leqslant \Theta_{-p} < \Theta' \leqslant \frac{\pi}{2}$, a co za tym idzie $0 < \mu_o \leqslant \frac{1}{2}$, $0 < \eta_o < 1$.

Na szczególną uwagę zasługują przypadki $\Theta' = \frac{\pi}{2}$ oraz $\Theta_{-p} = 0$. Mianowicie dla $\Theta' = \frac{\pi}{2}$ następuje sklejenie przedziałów $[\Theta_{-p}, \Theta']$ i $[\pi-\Theta', \pi-\Theta_{-p}]$, to znaczy przedziałów, w których zachowanie się funkcji $Q(j\Theta)$ jest podobne. Jeżeli w przedziale $[\Theta_{-p}, \Theta']$ funkcja $Q(j\Theta)$ ma przebieg typu pasma przepustowego, przy czym $\Theta' = \Theta_p$, to podobny przebieg wykazuje ona także dla $\Theta'_{-p} = \pi-\Theta_p \leqslant \Theta \leqslant \Theta'_p = \pi-\Theta_{-p}$. Stąd widać, że dla $\Theta_p = \Theta' = \frac{\pi}{2}$ wystąpi równość $\Theta_p = \Theta'_{-p} = \pi-\Theta' = \frac{\pi}{2}$, a więc następuje podwojenie szerokości pasma przepustowego, które będzie rozciągało się od Θ_{-p} do $\Theta'_p = \pi-\Theta_{-p}$. Wartość $\Theta' = \frac{\pi}{2}$ traci zatem sens górnej wartości granicznej omawianego pasma, a staje się jego wartością środkową, to znaczy $\Theta' = \Theta_o = \frac{1}{2}(\Theta_{-p}+\Theta'_p)$.

Podobne zjawisko sklejenia przedziałów występuje dla $\Theta_{-p} = 0$. Wtedy w otoczeniu $\Theta = \pi$ funkcja $Q(j\Theta)$ posiada pasmo przepustowe o dwukrotnie większej szerokości w porównaniu z pasmem $0 = \Theta_{-p} \leqslant \Theta \leqslant \Theta_p < \frac{\pi}{2}$. Omówione właściwości funkcji

$Q(j\Theta)$ należy brać pod uwagę także przy szacowaniu szerokości pasma tłumieniowego. Ilustrację powyższych uwag daje rys. 2.



Rys. 2. Przykład funkcji $Q(j\Theta)$ z przebiegiem typu pasma przepustowego

a) Dwa osobne pasma przepustowe w przedziale $[0, \pi]$; b) przypadek sklejenia pasm przepustowych dla $\Theta_o = \frac{\pi}{2}$ (krzywa I) oraz dla $\Theta_{-p} = 0$ (krzywa II)

Jeżeli omówione przekształcenie zmiennych według (9), (11) zastosować względem funkcji $Q(s_1)$, wzór (6), to wtedy odpowiadająca jej funkcja $G[z(\tau_1)]$ będzie miała następujące własności:

1. $$0 \leqslant G[x(\tau_1)] \leqslant h^2 \text{ w przedziale } -1 \leqslant x(\tau_1) \leqslant 1,$$

przy czym

$$x_p(\tau_1) = \frac{\cos^2 \tau_1 \omega' - \eta_o}{\mu_o} = -1, \qquad x_{-p}(\tau_1) = \frac{\cos^2 \tau_1 \omega_{-p} - \eta_o}{\mu_o} = 1,$$

gdzie

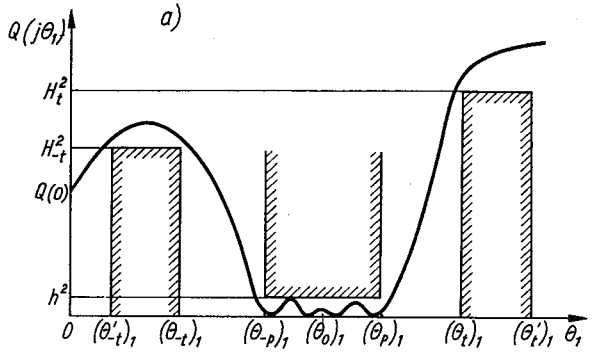
$$\omega' = \begin{cases} \omega_o = \dfrac{1}{2}(\omega_{-p} + \omega_p), & \text{jeżeli } \tau_1 \omega_o = \dfrac{\pi}{2}, \\ \omega_p, & \text{jeżeli } \tau_1 \omega_p < \dfrac{\pi}{2}; \end{cases}$$

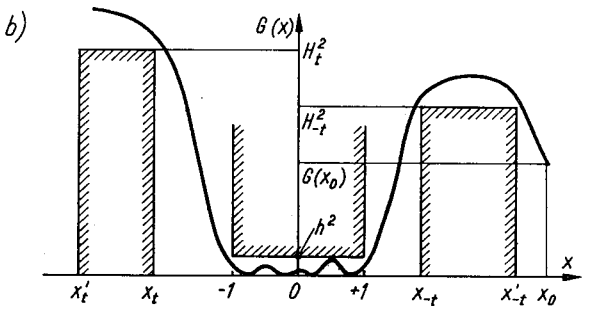
2. $$G[x(\tau_1)] \geqslant H^2_{-t} \text{ w przedziale } 1 < x_{-t}(\tau_1) \leqslant x(\tau_1) \leqslant x'_{-t}(\tau_1)$$

przy czym

$$x_{-t}(\tau_1) = \frac{\cos^2 \tau_1 \omega_{-t} - \eta_o}{\mu_o}, \quad x'_{-t}(\tau_1) = \frac{\cos^2 \tau_1 \omega'_{-t} - \eta_o}{\mu_o};$$

3. $G[x(\tau_1)] \geqslant H_{t_1}^2$ w przedziale $x'_t(\tau_1) \leqslant x(\tau_1) \leqslant x_t(\tau_1) < -1$





Rys. 3. Przykład funkcji $Q(s_1)$ i odpowiadającego jej wielomianu $G[z(\tau_1)]$ dla przypadku $\tau_1 \omega_p < \frac{\pi}{2}$

a) Wykres funkcji $Q(s_1)$ na osi Im $s_1 = \Theta_1$; b) Wykres wielomianu $G[z(\tau_1)]$ na osi Re $z(\tau_1) = x(\tau_1)$, zdeterminowany przez funkcję $Q(s_1)$

przy czym

$$x'_t(\tau_1) = \frac{\cos^2 \tau_1 \omega'_t - \eta_o}{\mu_o}, \quad x_t(\tau_1) = \frac{\cos^2 \tau_1 \omega_t - \eta_o}{\mu_o}$$

$\left(\text{punkt ten nie występuje dla } \tau_1 \omega_o = \frac{\pi}{2}\right)$;

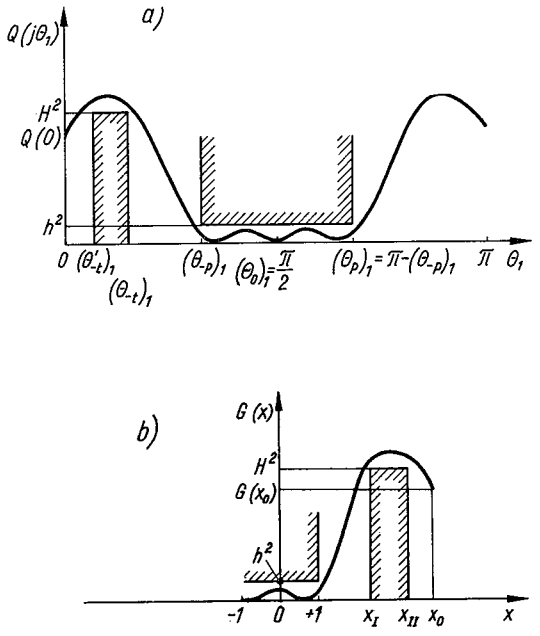
4. $$G[x_o(\tau_1)] = \frac{1}{4}\left(\sqrt{\frac{R_2}{R_1}} - \sqrt{\frac{R_1}{R_2}}\right)^2, \quad x_o(\tau_1) = \frac{1-\eta_o}{\mu_o};$$

5. $G[x(\tau_1)] \geqslant 0$ w przedziale $-\frac{\eta_o}{\mu_o} \leqslant x(\tau_1) \leqslant \frac{1-\eta_o}{\mu_o}$.

Rysunki 3 i 4 ilustrują wzajemne powiązanie przebiegów funkcji $Q(i\Theta_1)$ i wielomianu $G[x(\tau_1)]$.

W związku z przytoczonym zestawem właściwości wielomianu $G[z(\tau_1)]$, zdetermino-

wanym przez funkcję $Q(s_1)$, należy zauważyć, że właściwości objęte punktem 5 oraz wzorem (10) są podstawowe dla zapewnienia fizycznej realizowalności funkcji $G(z)$, $z = \frac{1}{\mu_o}(\mathrm{ch}^2 s - \eta_o)$, jako funkcji $Q(s) = L(s) - 1$. W dalszych rozważaniach przyjmuje się symbol $G(z)$ dla oznaczania tych wielomianów algebraicznych, które mogą być przekształcone — za pomocą podstawienia (9), (11) — w funkcję typu $Q(s)$, wzór (6).



Rys. 4. Przykład funkcji $Q(s_1)$ i odpowiadającego jej wielomianu $G[z(\tau_1)]$ dla $(\Theta_0)_1 = \frac{\pi}{2}$

a) Wykres funkcji $Q(s_1)$ na osi $\mathrm{Im}\, s_1 = \Theta_1$ dla $(\Theta_0)_1 = \frac{\pi}{2}$; b) Wykres wielomianu $G[z(\tau_1)]$ na osi $\mathrm{Re}\, z(\tau_1) = x(\tau_1)$, zdeterminowany przez funkcję $Q(s_1)$ dla $(\Theta_0)_1 = \frac{\pi}{2}$

W wyniku dotychczasowych rozważań ustalono sposób przekształcenia znanej funkcji $Q(s)$ w pewien wielomian algebraiczny $G(z)$; oczywiście można go także wykorzystać dla przejścia od wielomianu $G(z)$ do funkcji $Q(s)$. Celowość takiego postępowania wynika stąd, iż rozwiązywanie problemów aproksymacyjnych za pomocą wielomianów algebraicznych jest dziedziną dobrze zbadaną, posiadającą bogatą literaturę z gotowymi wzorami analitycznymi i algorytmami (np. [3], [4]).

Dla zrealizowania przekształcenia funkcyjnego za pomocą wzorów (9), (11) wystarczy znajomość parametrów μ_o, η_o, to znaczy wskazanie takiego przedziału położonego na odcinku $0 \leqslant \Theta \leqslant \frac{\pi}{2}$, który ma odpowiadać przedziałowi $[-1, 1]$ na osi x. Jednakże informacja ta nie jest zawarta w wymaganiach podanych na wstępie rozważań, gdyż dotyczą one właściwości transmisyjnych czwórnika w funkcji częstotliwości. Problem ten wiąże się z faktem, iż istnieje różnica między zapisem wymagań i charakterystyki toru

w funkcji częstotliwości a zapisem tychże w funkcji zmiennej $\Theta = \tau\omega$. Otóż wzajemnie jednoznaczne przyporządkowanie wartości granicznych na rys. 1 istnieje jedynie w układzie $f \sim |T_{21}|$, natomiast wartości Θ odpowiadające poszczególnym częstotliwościom granicznym mogą być różne — w zależności od przyjętego do realizacji współczynnika τ. Właśnie istnienie tej pewnej dowolności w doborze współczynnika τ rodzi niejednoznaczność syntezy toru typu N_o w oparciu o podane na wstępie wymagania.

Dla lepszego zrozumienia warto rozpatrzyć następujący przykład.

Wyznaczyć parametry odcinków jednorodnych w torze typu N_o, spełniającym punkty 1, 4 i 5 zestawu wymagań typowych, przy czym przyjmuje się, że kształt charakterystyki jest zadany za pomocą funkcji $G(z) = h^2 z^{2n}$; rozwiązania poszukuje się dla $\Theta_p < \frac{\pi}{2}$.

Na podstawie zależności (9)÷(11) otrzymuje się

$$Q(s) = h^2\left(\frac{\operatorname{ch}^2 s - \eta_o}{\mu_o}\right)^{2n}, \tag{12}$$

$$\mu_o = \frac{1}{2}\sin\tau(\omega_p+\omega_{-p})\sin\tau(\omega_p-\omega_{-p}),$$

$$\eta_o = 1-\frac{1}{2}(\sin^2\tau\omega_p+\sin^2\tau\omega_{-p}),$$

gdzie $\tau < \frac{\pi}{2\omega_p}$.

Można z łatwością przekonać się, że funkcja $Q(s)$ o postaci (12) spełnia warunki fizycznej realizacji (zob. twierdzenie o warunkach fizycznej realizacji funkcji strat toru typu N_o, podane w pracy [6]), a zatem wyznacza pewien tor typu N_o. Wyniki syntezy tego toru zależą od stopnia wielomianu $2n$ oraz parametrów h, μ_o, η_o, a także w sposób pośredni od wartości τ. Ponieważ wartość ta nie została z góry ustalona, więc wydaje się celowe obrać τ możliwie małe ($\tau \to 0$), aby w ten sposób zmniejszyć wypadkową długość geometryczną toru. A zatem

$$\underset{\tau\to 0}{Q(j\tau\omega)} = h^2\left(\frac{\operatorname{ch}^2 j\tau\omega - 1}{\frac{1}{2}\tau^2(\omega_p^2-\omega_{-p}^2)} + \frac{f_p^2+f_{-p}^2}{f_p^2-f_{-p}^2}\right)^{2n} = h^2\left(\frac{f_p^2+f_{-p}^2-2f^2}{f_p^2-f_{-p}^2}\right)^{2n}, \tag{13}$$

skąd

$$2n \geqslant \frac{\lg\frac{Q(0)}{h^2}}{\lg\frac{f_p^2+f_{-p}^2}{f_p^2-f_{-p}^2}}, \qquad \text{gdzie } Q(0) = \frac{1}{4}\left(\sqrt{\frac{R_2}{R_1}} - \sqrt{\frac{R_1}{R_2}}\right)^2. \tag{14}$$

Ze związków (13), (14) wynika, że dla $\tau \to 0$ minimalna liczba odcinków jednorodnych oraz charakterystyka toru nie zależą od konkretnej wartości τ. Na tej podstawie można wnioskować, że istnieje nieskończenie wiele torów typu N_o o jednakowej liczbie odcinków jednorodnych $2n$, spełniających identyczne wymagania odnośnie do charakterystyki przenoszenia (łącznie z punktem $f = 0$), a różniących się między sobą wartością parametru τ, a tym samym i charakterystykami w funkcji zmiennej Θ.

Z drugiej strony na podstawie dodatkowych danych można twierdzić, że gdy $\tau \to 0$, to $\mu_o \to 0$ i tym samym przynajmniej jeden współczynnik skoku impedancji falowej w omawianym torze zmierza do nieskończoności bądź zera. Ponieważ z praktyki realizacji torów niejednorodnych wiadomo, że technologicznie dopuszczalna wartość tego współczynnika jest ograniczona obustronnie, zatem ze względu na praktyczne możliwości realizacji wskazane jest ograniczenie dopuszczalnej wartości parametru μ_o od dołu. Należy to jednak uczynić tak, aby uniknąć możliwości nieporozumień, które mogą wystąpić w przypadku korzystania z zamiany zmiennej $s = ks'$ (k — dowolna liczba naturalna) we wzorze (6). Mianowicie dla odpowiednio dużych liczb k wartość parametru μ_o według (11), obliczona dla nowej zmiennej $\mathrm{Im}s'$, będzie mniejsza od każdej ustalonej wartości granicznej, aczkolwiek stopień trudności realizacyjnych układu nie uległ żadnej zmianie. Z tego widać, że należy tak sformułować ograniczenie, aby wspomniana zamiana zmiennej nie przesądzała sama przez się o przynależności toru o funkcji $Q(s')$ do grupy torów technologicznie realizowalnych.

Z przytoczonych względów przyjmuje się następującą interpretację ograniczenia w postaci zadanej minimalnie dopuszczalnej wartości parametru μ_o, oznaczonej symbolem μ_d. Otóż na podstawie wzoru (11) słuszne jest równanie

$$\mu_d = \frac{\cos^2\Theta_d \frac{f_{-p}}{f_o} - \cos^2\Theta_d \frac{f'}{f_o}}{2}, \quad f_o = \frac{1}{2}(f_{-p}+f_p), \tag{15}$$

które wyznacza wartość $0 < \Theta_d \leqslant \frac{\pi}{2}\frac{f_o}{f'} \leqslant \frac{\pi}{2}$; parametr Θ_d ma sens przesuwności falowej (długości elektrycznej) odcinka jednorodnego dla środkowej częstotliwości pasma przepustowego. Zakłada się, że układ spełnia warunek technologicznej realizowalności dla zadanej wartości μ_d, jeżeli przesuwność falowa Θ_k najkrótszego odcinka jednorodnego między dwoma skokami impedancji falowej o współczynniku skoku różnym od jedności, liczona dla środkowej częstotliwości pasma przepustowego, jest nie mniejsza od wartości Θ_d, określonej wzorem (15). Bardziej ogólnie: liczona dla częstotliwości środkowej pierwszego przedziału na osi Θ, w którym zachowanie się funkcji $Q(j\Theta)$ jest podobne do jej przebiegu w przedziale roboczym $[\Theta_{-p}, \Theta_p]$. W praktyce wartość Θ_k może być z góry określona i wtedy w oparciu o związek (15) oblicza się odpowiadającą jej wartość μ_d.

Współczynnik μ_o, jak widać z powyższych uwag, może służyć za pewien wskaźnik trudności realizacyjnych układu. Odgrywa on przy tym rolę podobną do tej, jaką spełnia względna szerokość pasma przepustowego przy ocenie trudności związanych z realizacją czwórnika o stałych skupionych. Jest to jednak wskaźnik bardziej kompleksowy, gdyż obok wpływu wielkości $(f_p - f_{-p})/f_o$ uwzględnia także wpływ parametru τ.

Wskaźnik μ_d może mieć zastosowanie zarówno przy syntezie toru niejednorodnego o elementach ściśle odpowiadających wynikom obliczeń, jak i przy syntezie tzw. toru zastępczego, którego elementy pozostają w określonym związku z elementami toru zrealizowanego. W drugim przypadku wartość μ_d jest ustalana w sposób nieco odmienny. Mianowicie celem ominięcia trudności związanych z obustronnym ograniczeniem wartości współczynnika skoku impedancji falowej stosuje się pomocnicze odcinki jednorodne

(tzn. takie, które nie wynikają bezpośrednio z założonej funkcji strat), których obecność — nie wpływając prawie na roboczą charakterystykę toru — umożliwia uzyskanie znacznie mniejszych wartości μ_o. Dlatego przy syntezie toru zastępczego należy przyjmować wartość μ_d z uwzględnieniem efektu oddziaływania pomocniczych elementów w roboczym pasmie częstotliwości.

W dalszych rozważaniach przyjmuje się, że znana jest minimalnie dopuszczalna wartość parametru $\mu_o = \mu_d$ i odpowiadająca jej wartość Θ_d oraz że tor niejednorodny należy zaprojektować z uwzględnieniem zarówno warunków typowych, wymienionych na wstępie rozdziału, jak i ograniczenia od dołu dopuszczalnych wartości Θ_k.

2. FUNKCJA STRAT TYPU CZEBYSZEWA JAKO ROZWIĄZANIE PEWNEGO PROBLEMU OPTYMALIZACJI PARAMETRÓW CHARAKTERYSTYCZNYCH UKŁADU

Praktyka syntezy torów niejednorodnych stawia następujący problem do rozwiązania:

Problem A

Wyznaczyć funkcję strat toru typu N przy obciążeniu R_1, R_2, która spełnia następujące wymagania:

1. $|T_{21}(j\omega)| \leqslant h$ dla $f_{-p} \leqslant f \leqslant f_p$,
2. $|T_{21}(j\omega)| \geqslant H_{-t}$ dla $f'_{-t} \leqslant f \leqslant f_{-t}$,
3. $|T_{21}(j\omega)| \geqslant H_t$ dla $f_t \leqslant f \leqslant f'_t$,

przy warunku dodatkowym

4. całkowita długość elektryczna toru

$$\Theta_c \overset{\text{df}}{=} \sum_{i=1}^{n} \Theta_{oi},$$

gdzie:

n — liczba odcinków jednorodnych w torze,

Θ_{oi} — przesuwność falowa (długość elektryczna) i-tego odcinka jednorodnego dla częstotliwości środkowej pasma przepustowego $f_o = \frac{1}{2}(f_{-p}+f_p)$,

ma być najmniejsza z możliwych przy założeniu, iż

$$\Theta_k \geqslant \Theta_d,$$

gdzie:

Θ_k — przesuwność falowa najkrótszego odcinka jednorodnego między dwoma skokami impedancji falowej o współczynniku skoku różnym od jedności dla częstotliwości f_o, z uwzględnieniem uwagi dotyczącej torów N_o w rozdz. 1,

$0 < \Theta_d \leqslant \frac{\pi}{2}$ — pierwiastek równania (15) dla zadanej wartości μ_d.

Poszukiwanie postaci funkcji strat toru niejednorodnego spełniającego wymagania problemu A jest realizowane w dwu etapach. W trakcie etapu I zostaje wyznaczona funkcja

strat toru typu N_o o charakterystyce Czebyszewa, która spełnia postawione wymagania. Z kolei w ramach etapu II udowadnia się, że rozwiązanie uzyskane w trakcie etapu I jest jedyne w całej klasie torów typu N.

W trakcie rozważań wykorzystuje się rezultaty rozdziału 1, w myśl których poszukuje się takiego wielomianu algebraicznego $G(z)$, który w wyniku podstawienia (9), (11) przekształca się w fizycznie realizowalną funkcję $Q(s) = L(s) - 1$, spełniającą wymagania problemu A. Poszukiwanie rozwiązania w ramach etapu I przebiega więc następująco.

W pierwszej kolejności określa się właściwości funkcji $G(z)$, wzór (10), zdeterminowane przez wymagania stawiane funkcji $Q(s)$ o postaci (6). Przyjmuje się przy tym jako zasadę, że indeksy punktów na osi $\mathrm{Re}\, z = x$ są takie same jak odpowiadających im punktów na osi częstotliwości. W ten sposób otrzymuje się, że wielomian algebraiczny $G(z)$, $z = \frac{1}{\mu_o}(\mathrm{ch}^2 s - \eta_o)$, reprezentujący funkcję $Q(s)$ o postaci (6) i o właściwościach określonych przez wymagania problemu A, powinien spełniać na osi $\mathrm{Re}\, z = x$ następujące wymagania:

1. $0 \leqslant G(x) \leqslant h^2$ w przedziale $-1 \leqslant x \leqslant 1$,

przy czym

$$x_p = \frac{\cos^2 \tau\omega_p - \eta_o}{\mu_o} = -1, \qquad x_{-p} = \frac{\cos^2 \tau\omega_{-p} - \eta_o}{\mu_o} = 1;$$

2. $G(x) \geqslant H_{-t}^2$ w przedziale $1 < x_{-t}(\tau) \leqslant x \leqslant x'_{-t}(\tau)$,

przy czym

$$x_{-t}(\tau) = \frac{\cos^2 \tau\omega_{-t} - \eta_o}{\mu_o}, \qquad x'_{-t}(\tau) = \frac{\cos^2 \tau\omega'_{-t} - \eta_o}{\mu_o};$$

3. $G(x) \geqslant H_t^2$ w przedziale $x'_t(\tau) \leqslant x \leqslant x_t(\tau) < -1$,

przy czym

$$x'_t(\tau) = \frac{\cos^2 \tau\omega'_t - \eta_o}{\mu_o}, \qquad x_t(\tau) = \frac{\cos^2 \tau\omega_t - \eta_o}{\mu_o};$$

4. $$G(x_o) = \frac{1}{4}\left(\sqrt{\frac{R_2}{R_1}} - \sqrt{\frac{R_1}{R_2}}\right)^2, \qquad x_o(\tau) = \frac{1-\eta_o}{\mu_o};$$

5. $$G(x) \geqslant 0 \text{ w przedziale } -\frac{\eta_o}{\mu_o} \leqslant x \leqslant \frac{1-\eta_o}{\mu_o}.$$

W związku z wymaganiami odnośnie do zachowania się funkcji $G(z)$ na osi $\mathrm{Re}\, z = x$ należy zwrócić uwagę na fakt, iż jedynie przedział $[x_p, x_{-p}]$ został z góry określony; natomiast pozostałe punkty charakterystyczne na osi x mogą być określone dopiero po ustaleniu współczynnika τ, który może przyjmować różne wartości pod warunkiem zachowania nierówności $\Theta_k \geqslant \Theta_d$. Współczynniki μ_o, η_o oblicza się na podstawie wzoru (11), przy czym

$$\omega' = \begin{cases} \omega_p & \text{jeżeli } \Theta_p = \tau\omega_p < \dfrac{\pi}{2}, \\ \omega_o & \text{jeżeli } \Theta_o = \tau\omega_o = \dfrac{\pi}{2}. \end{cases}$$

Należy także dodać, że powyższy wykaz właściwości funkcji $G(x)$ odpowiada w zasadzie sytuacji, gdy $\Theta_p < \frac{\pi}{2}$. Natomiast w przypadku $\Theta_o = \frac{\pi}{2}$ należy uwzględnić pewne zmiany. Mianowicie punkt $x_p = -1$, wymieniony w punkcie 1, ma odpowiadać częstotliwości f_o, a warunki 2 i 3 należy zastąpić przez jeden warunek

$$G(x) \geqslant H^2 \text{ w przedziale } 1 < x_{\mathrm{I}} \leqslant x \leqslant x_{\mathrm{II}},$$

gdzie:

H^2 — większa z liczb H^2_{-t}, H^2_t,

x_{I} — mniejsza z liczb x_{-t}, x_t,

x_{II} — większa z liczb x'_{-t}, x'_t.

Dalsze postępowanie polega na wprowadzeniu funkcji pomocniczej

$$G_1(z) \overset{\mathrm{df}}{=} G(z) - \frac{h^2}{2}, \tag{16}$$

której właściwości wynikają jednoznacznie z cech charakterystycznych funkcji $G(z)$. W ten sposób problem A w ramach etapu I został przetransformowany na równoważny mu problem poszukiwania wielomianu algebraicznego $G_1(z)$ o założonych właściwościach. Istotę tego problemu można ująć następująco.

Spośród wszystkich wielomianów algebraicznych $G_1(z)$ o współczynnikach rzeczywistych, spełniających na osi $\mathrm{Re}\, z = x$ warunki

1. $$|G_1(x)| \leqslant \frac{h^2}{2} \text{ w przedziale } -1 \leqslant x \leqslant 1,$$

2. $$G_1(x) \geqslant H^2_{-t} - \frac{h^2}{2} \text{ w przedziale } 1 < x_{-t} \leqslant x \leqslant x'_{-t},$$

3. $$G_1(x) \geqslant H^2_t - \frac{h^2}{2} \text{ w przedziale } x'_t \leqslant x \leqslant x_t < -1,$$

4. $$G_1(x_o) = \frac{1}{4}\left(\sqrt{\frac{R_2}{R_1}} - \sqrt{\frac{R_1}{R_2}}\right)^2 - \frac{h^2}{2}, \qquad x_o = \frac{1-\eta_o}{\mu_o},$$

5. $$G_1(x) \geqslant -\frac{h^2}{2} \text{ w przedziale } -\frac{\eta_o}{\mu_o} \leqslant x \leqslant \frac{1-\eta_o}{\mu_o},$$

gdzie parametry μ_o, η_o są określone przez wzór (11),

$$x = \frac{\cos^2 \tau\omega - \eta_o}{\mu_o}, \qquad 0 < \tau < \frac{\pi}{2\omega_p}, \qquad x_p = -1, \qquad x_{-p} = 1,$$

bądź warunki

1′. $$|G_1(x)| \leqslant \frac{h^2}{2} \text{ w przedziale } -1 \leqslant x \leqslant 1,$$

2′. $$G_1(x) \geqslant H^2 - \frac{h^2}{2} \text{ w przedziale } 1 < x_{\mathrm{I}} \leqslant x \leqslant x_{\mathrm{II}},$$

3'. $$G_1(x_o) = \frac{1}{4}\left(\sqrt{\frac{R_2}{R_1}} - \sqrt{\frac{R_1}{R_2}}\right)^2 - \frac{h^2}{2}, \qquad x_o = -1+2\cos^{-2}\left(\frac{\pi}{2}\frac{f_{-p}}{f_o}\right),$$

4'. $$G_1(x) \geqslant -\frac{h^2}{2} \text{ w przedziele } -1 \leqslant x \leqslant -1+2\cos^{-2}\left(\frac{\pi}{2}\frac{f_{-p}}{f_o}\right),$$

gdzie

$$x = 2\frac{\cos^2\frac{\pi}{2}\frac{f}{f_o}}{\cos^2\frac{\pi}{2}\frac{f_{-p}}{f_o}} - 1, \qquad x(\omega_o) = -1, \qquad x_{-p} = 1, \qquad f_o = \frac{1}{2}(f_{-p}+f_p),$$

wyznaczyć taki wielomian $W_n(z)$, który realizuje

$$\Theta_c = n\Theta_o = n\tau\omega_o = \underset{\Theta_k \geqslant \Theta_d}{\text{minimum}}. \tag{17}$$

Jak więc widać, należy rozwiązać problem optymalizacji pewnych parametrów, określających użytkową przydatność układu transmisyjnego. Poszukiwanie rozwiązania tego problemu przebiega według następującego schematu. Na początku określa się cechy szczególne poszukiwanej funkcji, na podstawie których można wyodrębnić taki wielomian algebraiczny $W_n'(z)$, który najlepiej spełnia postawione wymagania (to znaczy wykazuje najmniejszy moduł ekstremalnej wartości w przedziale $[-1, 1]$ na osi $\mathrm{Re}z = x$ przy jednoczesnym spełnieniu pozostałych warunków) dla założonej wartości współczynnika τ spośród wszystkich wielomianów tego samego stopnia, przy czym stopień n ma być najniższy z możliwych. W wyniku realizacji tego etapu otrzymuje się pewien zbiór wielomianów $W_n'(z)$, które — jak wykazano w rozdziale 3 — należą do grupy wielomianów Czebyszewa. Każdy element tego zbioru może służyć za podstawę do utworzenia fizycznie realizowanej funkcji $Q(s)$ według (6) z właściwą tylko dla niego zmienną $s = s(\tau)$. Po określeniu zbioru wielomianów $W_n'(z)$ przystępuje się do poszukiwania takiego wielomianu $W(z_n)$ z tego zbioru, dla którego spełniony jest warunek (17).

W tym celu przyjmuje się

$$z = \frac{\mathrm{ch}^2 s(\tau) - \eta_o(\tau)}{\mu_o(\tau)}$$

i uwzględnia się tę zamianę zmiennych we właściwym wielomianie $W_n'(z)$, a następnie za pomocą tej nowej funkcji konstruuje się funkcję

$$L[s(\tau)] = 1 + \frac{h^2}{2} + W_n'\left(\frac{\mathrm{ch}^2 s(\tau) - \eta_o(\tau)}{\mu_o(\tau)}\right), \tag{18}$$

która spełnia twierdzenie o warunkach fizycznej realizacji funkcji strat toru typu N_o, a zatem wyznacza pewien tor niejednorodny N_o. Długość odcinka jednorodnego w tym torze jest ustalona przez warunek $x_{-p} = 1$. Obliczając jej wartość dla poszczególnych wielomianów $W_n'(z)$ przy jednoczesnym wyznaczeniu wartości Θ_k można wskazać taki wielomian $W_n(z)$ (i odpowiadający mu współczynnik τ) w zbiorze wielomianów $W_n'(z)$, dla którego spełniony jest warunek (17).

Tor niejednorodny o funkcji strat

$$L(s) = 1 + \frac{h^2}{2} + W_n\left(\frac{\operatorname{ch}^2 s - \eta_o}{\mu_o}\right) \tag{19}$$

i o długości odcinka jednorodnego wyznaczonej z warunku

$$\frac{\cos^2 \Theta_{-p} - \eta_o}{\mu_o} = 1 \quad \text{dla} \quad \Theta_{-p} < \frac{\pi}{2}, \tag{20}$$

gdzie Θ_{-p} — przesuwność falowa odcinka jednorodnego dla częstotliwości f_{-p}, jest zgodnie z wynikiem rozdziału 4 poszukiwanym torem niejednorodnym, spełniającym wszystkie wymagania problemu A. W szczególności oznacza to, że tor o funkcji strat (19) i z dobraną zgodnie z (20) długością odcinków jednorodnych będzie miał najmniejszą długość całkowitą spośród wszystkich torów typu N, spełniających warunki $1 \div 3$ problemu A, jeżeli tylko są one wykonane z odcinków jednorodnych o jednakowej jednostkowej przesuwności falowej. Charakterystyka przenoszenia tego toru jest równo falista (typu Czebyszewa).

Powyższe rozważania przedstawiają ogólne zasady poszukiwania rozwiązania omawianego problemu. Z kolei omówi się w sposób bardziej szczegółowy niektóre istotne punkty przedstawionego schematu postępowania, a w szczególności cechy charakterystyczne poszukiwanego wielomianu $W_n'(z)$, a także zagadnienie optymalności toru o funkcji strat według (19), (20) w klasie torów typu N.

3. POSTAĆ ANALITYCZNA WIELOMIANÓW $W_n'(z)$

Zgodnie z tym, co zostało powiedziane w rozdziale 2, przy poszukiwaniu wielomianu $W_n(z)$ należy w pierwszej kolejności wyodrębnić taki zbiór wielomianów algebraicznych $W_n'(z)$, którego elementy mają najniższy stopień spośród wszystkich wielomianów algebraicznych typu $G_1(z)$, minimalizujących moduł ekstremalnej wartości w przedziale $[-1, 1]$ na osi x przy jednoczesnym spełnieniu warunków $2 \div 5$ (albo $2' \div 4'$) dla założonej wartości współczynnika τ. Następnie przystępuje się do poszukiwania takiego wielomianu $W_n(z)$ z tego zbioru, dla którego spełniony jest warunek (17).

Przed przystąpieniem do realizacji pierwszej części zadania trzeba ustalić maksymalną liczbę ograniczeń, jakie należy nałożyć na poszukiwany wielomian typu $G_1(z)$. Rzecz w tym, iż nie jest konieczne, aby wartość poszukiwanego wielomianu była dokładnie równa granicznej dopuszczalnej wartości według wymagań na obu końcach każdego z zadanych przedziałów na osi x. W zasadzie wystarczy, aby wielomian przyjmował tę wartość najwyżej na jednym końcu odpowiedniego przedziału, a na drugim czynił jedynie zadość postawionym wymaganiom. Oczywiście wybór tych punktów zależy od konkretnej treści zadania i nie może być z góry ustalony. Wiadomo jednak, że dla $\Theta_p < \frac{\pi}{2}$ należy przewidzieć najwyżej trzy takie punkty (oznaczane jako x_o, x_1, x_2), a dla $\Theta_o = \frac{\pi}{2}$ — najwyżej dwa punkty (oznaczane jako x_0, x_1), w których wielomian ma ewentualnie przyjmować graniczne wartości dopuszczalne. Wymienione punkty mają następujące znaczenie:

x_o odpowiada punktowi x_o według wymagań,

$$\left.\begin{aligned} x_1 &= x_t' \quad \text{lub} \quad x_1 = x_t \\ x_2 &= x_{-t} \quad \text{lub} \quad x_2 = x_{-t}' \end{aligned}\right\} \quad \text{dla} \quad \Theta_p < \frac{\pi}{2}, \tag{21}$$

$$x_1 = x_{\mathrm{I}} \quad \text{lub} \quad x_1 = x_{\mathrm{II}} \quad \text{dla} \quad \Theta_o = \frac{\pi}{2}.$$

Zgodnie z powyższymi ustaleniami poszukiwany wielomian typu $G_1(z)$ stopnia n daje się napisać w postaci

$$G_1(z) = U_{n-q-1}(z) \prod_{k=0}^{q} (z-x_k) + \sum_{i=0}^{q} G_1(x_i) \prod_{\substack{k=0 \\ k\neq i}}^{q} \frac{z-x_k}{x_i-x_k}, \tag{22}$$

gdzie:

$U_{n-q-1}(z)$ — wielomian stopnia $n-q-1$,

$$q = \begin{cases} 2 & \text{dla} \quad \Theta_p < \dfrac{\pi}{2}, \\ 1 & \text{dla} \quad \Theta_o = \dfrac{\pi}{2}. \end{cases}$$

Wyrażenie (22) można przekształcić następująco:

$$G_1(z) = \prod_{k=0}^{q} (z-x_k) \left[U_{n-q-1}(z) + \sum_{i=0}^{q} \frac{G_1(x_i)}{z-x_i} \prod_{\substack{k=0 \\ k\neq i}}^{q} (x_i-x_k)^{-1} \right]. \tag{23}$$

W tym miejscu warto zauważyć, że z uwagi na nierówność $|x| > 1$ (zob. (21)) wielomian $\prod_{k=0}^{q} (z-x_k)$ nie zeruje się w przedziale $[-1, 1]$ na osi x.

A więc poszukuje się wielomianu $W_n'(z)$ o następujących właściwościach na osi Re $z = x$:

$$\frac{h^2}{2} = \max_{-1\leqslant x\leqslant 1} |W_n'(x)| = \min \max_{-1\leqslant x\leqslant 1} |G_1(x)| = \min \max_{-1\leqslant x\leqslant 1} \left| \prod_{k=0}^{q} (x-x_k) \right| \times$$

$$\times \left| U_{n-q-1}(x) + \sum_{i=0}^{q} \frac{G_1(x_i)}{x-x_i} \prod_{\substack{k=0 \\ k\neq i}}^{q} (x_i-x_k)^{-1} \right| \tag{24}$$

dla zadanych punktów x_i $(i = 0, 1, 2)$ oraz wartości $G_1(x_0)$,

$$G_1(x_1) \geqslant H_t^2 - \frac{h^2}{2}, \qquad G_1(x_2) \geqslant H_{-t}^2 - \frac{h^2}{2} \quad \text{dla} \quad \Theta_p < \frac{\pi}{2}$$

bądź

$$G_1(x_1) \geqslant H^2 - \frac{h^2}{2} \quad \text{dla} \quad \Theta_o = \frac{\pi}{2}$$

przy warunku dodatkowym

$$W_n'(x) \geqslant -\frac{h^2}{2} \quad \text{w przedziale} \quad -\frac{\eta_0(\tau)}{\mu_0(\tau)} \leqslant x \leqslant \frac{1-\eta_0(\tau)}{\mu_0(\tau)},$$

a stopień n ma być najniższy z możliwych.

Ze sformułowania problemu w postaci (24) wynika, że wielomian $U_{n-q-1}(x)$ ma być wielomianem najlepszego przybliżenia w sensie Czebyszewa stopnia $n-q-1$ funkcji

$$\sum_{i=0}^{q} \frac{G_1(x_i)}{x-x_i} \prod_{\substack{k=0 \\ k\neq i}}^{q} (x_i-x_k)^{-1} \qquad \text{z wagą} \qquad q(x) = \left| \prod_{k=0}^{q} (x-x_k) \right|$$

w przedziale $[-1, 1]$.

Zgodnie z [3], [4] warunek konieczny i wystarczający na to, aby wielomian $U(x)$ stopnia nie większego od n był wielomianem najlepszego przybliżenia w sensie Czebyszewa funkcji $f(x)$ z wagą $q(x)$ w przedziale $[-1, 1]$ brzmi:

w przedziale $[-1, 1]$ istnieją co najmniej $n+2$ punkty $x_1^0 < x_2^0 < \ldots < x_{n+2}^0$ takie, że w nich funkcja

$$R(x) \equiv q(x)\,[U(x)-f(x)]$$

przyjmuje ekstremalne wartości równe na przemian L i $-L$, gdzie

$$L = \max_{-1\leqslant x\leqslant 1} q(x)\,|U(x)-f(x)|.$$

Korzystając z tego twierdzenia przy analizie wyrażenia (24) dochodzi się do wniosku, że wielomian $W_n'(x)$ powinien przyjmować na przemian w co najmniej $n-q+1$ punktach przedziału $[-1, 1]$ wartość ekstremalną $L = \frac{h^2}{2}$ i $-L$. Tę właściwość posiadają następujące wielomiany stopnia n [1], [3], [4]:

$$\frac{h^2}{2} P_n(x) = \frac{h^2}{2} \operatorname{ch}(n \operatorname{Arch} x), \tag{25}$$

$$\frac{h^2}{2} P_n(x, r, k) = \frac{h^2}{4} \left\{ \left[\frac{H\left(u-\frac{rK}{n}\right)}{H\left(u+\frac{rK}{n}\right)} \right]^n + \left[\frac{H\left(u+\frac{rK}{n}\right)}{H\left(u-\frac{rK}{n}\right)} \right]^n \right\}, \tag{26}$$

przy czym

$$x = \frac{\operatorname{sn}^2 u + \operatorname{sn}^2 \frac{rK}{n}}{\operatorname{sn}^2 u - \operatorname{sn}^2 \frac{rK}{n}}, \qquad r = 1, 2, \tag{27}$$

gdzie:

k — moduł funkcji eliptycznej,

K — pełna całka eliptyczna pierwszego rodzaju, odpowiadająca modułowi k,

$\operatorname{sn} u$ — sinus eliptyczny Jacobiego o module k,

$H(u) = \vartheta_1\left(\frac{u}{2K}\right)$ — funkcja theta Jacobiego.

Wielomian $P_n(x)$, wzór (25), jest wielomianem Czebyszewa pierwszego rodzaju n-tego stopnia [4]. W przedziale $[-1, 1]$ przyjmuje on wartość zerową w n różnych punktach, a wartość ekstremalną $L = 1$ i $-L$ — na przemian w $n+1$ punktach. Na uwagę zasługuje fakt, iż $P_n(-1) = (-1)^n P_n(1)$.

Wielomian $P_n(x, r, k)$, wzory (26), (27), jest uogólnionym wielomianem Czebyszewa stopnia n [1]. Dla niego charakterystyczne jest istnienie dwu przedziałów, w których jego przebieg jest równo falisty. W przedziale $[-1, 1]$ zeruje się on w n-r różnych punktach, a w pewnym przedziale $[a, b]$, $a > 1$, — w r różnych punktach. Wielomian $P_n(x, r, k)$ przyjmuje wartość ekstremalną $L = 1$ i $-L$ na przemian w $n-r+1$ punktach przedziału $[-1, 1]$ i w $r+1$ punktach przedziału $[a, b]$. Należy zaznaczyć, że $P_n(-1, r, k) = (-1)^{n-r} P_n(1, r, k)$.

Dla zorientowania się, który z powyższych wielomianów ma być zaliczony do zbioru wielomianów $W_n'(z)$, należy uwzględnić wymagania określające zachowanie się wielomianu $W_n'(z)$ poza przedziałem $[-1, 1]$ na osi $\mathrm{Re}\, z = x$. W szczególności z uwagi na warunek 5 dla $\Theta_p < \frac{\pi}{2}$ należy ograniczyć grupę badanych funkcji do wielomianów typu $P_n(x)$ stopnia parzystego oraz do wielomianów typu $P_n(x, r, k)$ (branych ze znakiem $(-1)^r$) stopnia parzystego dla $r = 2$ i stopnia nieparzystego dla $r = 1$. Nieco inaczej jest w przypadku $\Theta_o = \frac{\pi}{2}$. W tym wariancie realizacyjnym mogą być brane pod uwagę wielomiany typu $P_n(x)$ oraz typu $(-1)P(x, 1, k)$ dowolnego stopnia n. W ten sposób dochodzi się do wniosku, że poszukiwane wielomiany $W_n'(z)$ mogą mieć jedynie następującą postać analityczną:

$$W_n'(z) = \begin{cases} \frac{h^2}{2} \operatorname{ch}(n \operatorname{Arch} z), & \text{gdzie } n \text{ ma być parzyste, jeżeli } \Theta_p < \frac{\pi}{2}, \\ (-1)^r \frac{h^2}{2} P_n(z, r, k), & \text{gdzie } r = 1 \text{ dla } \Theta_o = \frac{\pi}{2}; \text{ natomiast w przy-} \\ & \text{padku } \Theta_p < \frac{\pi}{2} \ r = 1, 2, \text{ a } n\text{-}r \text{ ma być parzyste.} \end{cases} \tag{28}$$

Konkretny kształt funkcji oraz wartości poszczególnych parametrów we wzorze (28) zależą od warunków $2 \div 5$ (lub $2' \div 4'$) określających zachowanie się wielomianu $W_n'(z)$ poza przedziałem $[-1, 1]$ na osi $\mathrm{Re}\, z = x$. Z definicji wielomianu $W_n'(z)$ wynika, że stopień n wielomianów o postaci (28) ma być najniższy z możliwych dla ustalonej wartości τ i nałożonych wymagań $1 \div 5$ (bądź $1' \div 4'$). Zbiór tych wielomianów dla poszczególnych wartości współczynnika τ jest poszukiwanym zbiorem wielomianów $W_n'(z)$. W ten sposób pierwsza część zadania, sformułowanego na wstępie rozdziału 3, została zrealizowana.

Z kolei przystępuje się do wyznaczania takiego wielomianu $W_n(z)$ w zbiorze wielomianów $W_n'(z)$, dla którego spełniony jest warunek (17). Zadanie to można rozwiązać w sposób konkretny jedynie na podstawie znajomości parametrów wymienionych w problemie A.

4. DOWÓD OPTYMALNEGO CHARAKTERU TORU NIEJEDNORODNEGO O FUNKCJI STRAT WEDŁUG (19), (20) W KLASIE TORÓW TYPU N

Po określeniu postaci analitycznej wielomianów $W_n'(z)$, wzór (28), oraz zasad postępowania przy wyznaczaniu wielomianu $W_n(z)$ należy z kolei wykazać, że funkcja strat $L(s)$ skonstruowana w oparciu o ten wielomian — zgodnie z zależnością (19) z uwzględnieniem

(20) — stanowi rzeczywiście rozwiązanie problemu A. Należy zatem dać odpowiedź na następujące pytanie:

Problem B

Czy istnieje tor typu N o lepszych właściwościach niż tor o funkcji strat (19), (20), to znaczy taki, który spełnia następujące wymagania:

1. $h_2^2 \leqslant h_1^2 = h^2$ w przedziale $\omega_{-p} \leqslant \omega \leqslant \omega_p$,

$$h_i^2 = \max_{\omega_{-p} \leqslant \omega \leqslant \omega_p} Q_i(j\omega), \quad i = 1, 2,$$

2. $Q_2(j\omega) \geqslant H_{-t}^2$ w przedziale $\omega'_{-t} \leqslant \omega \leqslant \omega_{-t}$,
3. $Q_2(j\omega) \geqslant H_t^2$ w przedziale $\omega_t \leqslant \omega \leqslant \omega'_t$,
4. $Q_2(0) = Q_1(0)$,
5. $\Theta_{c2} < \Theta_{c1}$,

przy czym

6. $\Theta_d \leqslant \Theta_{k2}$.

W powyższym zestawieniu wymagań indeks 1 wyróżnia tor o funkcji strat według (19), (20), a indeks 2 — tor domniemany.

Poszukiwanie odpowiedzi przebiega następująco.

Niech dane będą dwa tory typu N (tor 1 i tor 2), które zostały wyznaczone na podstawie identycznych wymagań $1 \div 3$ problemu A. Tor 1 obejmuje n_1 odcinków jednorodnych o jednakowej długości, a jego całkowita długość elektryczna dla częstotliwości f_0 wynosi $\Theta_{c1} = n_1\Theta_{01}$. Tor 2 obejmuje n_2 odcinków jednorodnych oraz posiada $2 \leqslant r_2 \leqslant n_2+1$ skoków impedancji falowej o współczynniku skoku różnym od jedności (wśród nich współczynniki skoku na wejściu i wyjściu toru); całkowita długość elektryczna tego toru dla f_0 wynosi

$$\Theta_{c2} = \sum_{i=1}^{n_2} (\Theta_{oi})_2 .$$

Zakłada się, że $\Theta_{c2} < \Theta_{c1}$ oraz że funkcja strat toru 1 odpowiada zależności (19) z uwzględnieniem (20), a właściwości toru 2 są zgodne z punktami $1 \div 6$ problemu B.

Na podstawie zadanych wartości Θ_{c2} i Θ_d można obliczyć parametry

$$n_3 = E\left(\frac{\Theta_{c2}}{\Theta_d}\right) \quad \text{oraz} \quad \tau_3 = \frac{\Theta_{c2}}{\omega_0 n_3}, \tag{29}$$

gdzie $E(x)$ — część całkowita liczby rzeczywistej x, a następnie wartości współczynników $\mu_{03} = \mu_o(\tau_3)$, $\eta_{03} = \eta_0(\tau_3)$ (wzór (11)). Ze sposobu określenia liczby n_3 wynika, że spełniona jest nierówność $n_3 \geqslant r_2 - 1$. Fakt ten łącznie z założeniami odnośnie do właściwości toru 2 gwarantuje pozytywny wynik poszukiwań takiej funkcji strat $L_3(s_3)$, $s_3 = s(\tau_3)$, stopnia $2n_3$ o postaci

$$L_3(s_3) = 1 + \frac{h_3^2}{2} + W''_{n_3}\left(\frac{\operatorname{ch}^2 s_3 - \eta_{03}}{\mu_{03}}\right), \tag{30}$$

(gdzie $W''_{n_3}(z)$ jest wielomianem stopnia n_3 o postaci (28)), dla której odpowiadająca jej funkcja $Q_3(s_3)$ będzie spełniała na osi Im $s_3 = \Theta_3 = \tau_3\omega$ następujące wymagania:

1. $Q_3(j\tau_3\omega) \geqslant H^2_{-t}$ dla $\omega'_{-t} \leqslant \omega \leqslant \omega_{-t}$,
2. $Q_3(j\tau_3\omega) \geqslant H^2_t$ dla $\omega_t \leqslant \omega \leqslant \omega'_t$,
3. $Q_3(0) = Q_2(0)$,

przy możliwie najmniejszej wartości

4. $h_3^2 = \max\limits_{w-p \leqslant w \leqslant w_p} Q_3(j\tau_3\omega)$.

W świetle poczynionych założeń o sposobie wyznaczania funkcji strat $L_3(s_3)$ można z góry przewidzieć, iż zachodzi $h_3 > h_1 \geqslant h_2$ przy $\Theta_d \leqslant \Theta_{k3}$. W przeciwnym razie należałoby przyjąć, iż istnieje taki element w zbiorze wielomianów $W'_n(z)$ różny od $W_n(z)$, który po podstawieniu do wzoru (19) realizuje funkcję strat toru typu N_0, spełniającego wymagania problemu A łącznie z warunkiem technologicznej realizowalności przy mniejszej od Θ_{c1} całkowitej długości elektrycznej, co jest sprzeczne z określeniem wielomianu $W_n(z)$.

W dalszych rozważaniach zakłada się, że przez właściwy dobór parametrów wielomianu $W''_{n_3}(z)$ uzyskano ścisłą odpowiedniość przedziałów $[(\Theta_{-p})_2, (\Theta_p)_2]$ i $[(\Theta_{-p})_3, (\Theta_p)_3]$, a także spełnienie warunku 1 bądź 2, dotyczącego funkcji $Q_3(s_3)$, ze znakiem równości przynajmniej w jednym punkcie odpowiedniego przedziału w przypadkach, gdy $(\Theta_o)_3 = \frac{\pi}{2}$, a funkcja $W''_{n_3}(z)$ jest uogólnionym wielomianem Czebyszewa, oraz gdy $(\Theta_p)_3 < \frac{\pi}{2}$, a $W''_{n_3}(z)$ jest analogicznym wielomianem stopnia nieparzystego, lub spełnienie warunków 1 i 2 ze znakiem równości przynajmniej w jednym punkcie każdego z przedziałów, jeżeli $(\Theta_p)_3 < \frac{\pi}{2}$, a funkcja $W''_{n_3}(z)$ jest uogólnionym wielomianem Czebyszewa stopnia parzystego. Na podstawie powyższych uwag oraz znanych właściwości wielomianów Czebyszewa [1], [3], [4] dochodzi się do wniosku, że funkcja

$$F(j\omega) = Q_3(j\tau_3\omega) - Q_2(j\omega) \tag{31}$$

zeruje się co najmniej w n_3+1 różnych punktach $0 = \omega_1 < \omega_2 \ldots < \omega_{3+1} \leqslant \omega'_t$, przy czym odpowiedniki tych punktów na osi Θ_3 leżą w przedziale $\left[0, \frac{\pi}{2}\right]$. Punkty te są nawzajem niezależne.

Tor 2 obejmuje $1 \leqslant r_2 - 1 \leqslant n_3$ odcinków jednorodnych, występujących między skokami impedancji falowej o współczynniku skoku różnym od jedności (z założenia o właściwościach toru 2). Przesuwność falowa $(\Theta_{oi})_2$ i-tego odcinka jednorodnego daje się napisać w postaci

$$(\Theta_{oi})_2 = k_i(\Theta_o)_3 + y'_i, \quad k_i \text{ — liczba naturalna,}$$

przy czym z uwagi na warunek (zob. wzór (29))

$$\sum_{i=1}^{r_2-1} (\Theta_{oi})_2 = n_3(\Theta_o)_3$$

obowiązuje zależność

$$\sum_{i=1}^{r_2-1} y'_i = \sum_{i=1}^{r_2-1} k_i\left(\frac{y'_i}{k_i}\right) = \sum_{i=1}^{n_3} y_i = 0. \tag{32}$$

Na podstawie zależności (31), (32) dochodzi się do wniosku, iż istnieje układ n_3+1 niezależnych równań nieliniowych o postaci

$$\sum_{i=1}^{n_3} y_i = 0, \quad F(j\omega_k, y_1, \ldots, y_{n_3}) = 0, \quad k = 2, 3, \ldots, n_3+1,$$

który jest spełniony przez co najwyżej n_3 parametrów niezależnych y_i, co jak wiadomo jest niemożliwe. Dlatego można twierdzić, że tor 2 typu N o właściwościach według sformułowania problemu B nie istnieje, a zatem tor 1 o funkcji strat według (19), (20) jest optymalny w klasie torów typu N, co należało wykazać. W ten sposób problematyka funkcji strat typu Czebyszewa została wyczerpana.

5. FUNKCJA STRAT TORU NIEJEDNORODNEGO O CHARAKTERYSTYCE TYPU BUTTERWORTHA

Charakterystyka przenoszenia typu Butterwortha należy obok charakterystyki typu Czebyszewa do najczęściej spotykanych w praktyce syntezy torów niejednorodnych. Dlatego wydaje się celowe poświęcić jej nieco uwagi.

Charakterystyka typu Butterwortha (maksymalnie płaska) jest określana podobnie jak w teorii układów o stałych skupionych [2], [10]. Za punkt wyjścia służy funkcja

$$G(z') = h^2 z'^n, \tag{33}$$

w której uwzględnia się przekształcenie zmiennych zgodnie z wynikami rozdziału 1. Do obliczania niezbędnych parametrów służą wymagania analogiczne do przedstawionych w rozdziale 1 z tą wszakże różnicą, iż przebieg charakterystyki w pasmie przepustowym jest kontrolowany przede wszystkim dla częstotliwości f', będącej punktem zerowym funkcji $Q(j\tau\omega)$, wzór (6), oraz dla częstotliwości skrajnej pasma przepustowego użytkowego, np. dla $f_{-p} < f'$. Dlatego pożądane jest, aby parametry μ_o, η_o, wzór (11), były uzależnione od wymienionych częstotliwości. Można to uzyskać bez zmiany postaci związków (11), jeżeli we wzorze (33) wykorzystać podstawienie $2z' = z+1$, a zmienną z uzależnić od s za pomocą wzorów (9), (11). W ten sposób uzyskuje się funkcję strat typu Butterwortha w postaci

$$L(s) = 1+h^2\left(\frac{\operatorname{ch}^2 s - \cos^2\Theta'}{2\mu_o}\right)^n, \tag{34}$$

gdzie n ma być parzyste, jeżeli $\Theta' < \frac{\pi}{2}$; μ_o oblicza się ze wzoru (11) uwzględniając przy tym to, że Θ' jest przesuwnością falową podstawowego odcinka jednorodnego dla częstotliwości f'.

Jak łatwo sprawdzić, funkcja $Q(s) = L(s)-1$, gdzie $L(s)$ według (34), zeruje się w punkcie Θ' n-krotnie, jeżeli $\Theta' < \frac{\pi}{2}$, i $2n$-krotnie, jeżeli $\Theta' = \frac{\pi}{2}$.

Wszystkie rozważania dotyczące wskaźników technologicznej realizowalności toru μ_d, Θ_d (rozdział 1) znajdują zastosowanie także przy syntezie toru niejednorodnego o charakterystyce Butterwortha, wzór (34). Znajomość tych parametrów jest konieczna do wyznaczenia parametru τ (lub przesuwności falowej Θ') w przypadku, gdy zależy na zmniejsze-

niu długości całkowitej toru nawet kosztem zwiększenia liczby skoków impedancji falowej o współczynniku skoku różnym od jedności.

6. UWAGI KOŃCOWE

W niniejszej pracy przedstawiono zarys postępowania przy poszukiwaniu funkcji strat torów niejednorodnych złożonych z bezstratnych odcinków jednorodnych o charakterystyce typu Czebyszewa i Butterwortha. Główną uwagę zwrócono na specyfikę problemu, a w szczególności na konieczność określenia pojęcia wskaźnika technologicznej realizowalności układu μ_d celem uzyskania jednoznaczności rozwiązania oraz na sposób przekształcenia wielomianów algebraicznych na fizycznie realizowalne funkcje strat toru niejednorodnego, odpowiadające określonym wymaganiom.

Na podstawie problemu A, dotyczącego optymalizacji parametrów charakterystycznych toru, przeanalizowano właściwości torów o charakterystyce typu Czebyszewa. Wyprowadzono postać analityczną funkcji strat typu Czebyszewa (wzory (19), (20), (28)) oraz omówiono metodykę poszukiwania rozwiązania optymalnego.

W oparciu o przedstawione zasady postępowania wyprowadzono ogólną postać funkcji strat typu Butterwortha (wzór (34)).

Całość rozważań ma spełnić dwojaki cel. Z jednej strony służy za przykład analizy, rozumowania przy rozwiązywaniu konkretnego problemu (tzn. przy wyznaczaniu funkcji strat typu Czebyszewa i Butterwortha). A jednocześnie z drugiej strony ma aspekt bardziej ogólny, gdyż może pomóc przy poszukiwaniu postaci analitycznej funkcji strat o właściwościach odmiennych od omówionych, akcentując w szczególności konieczność spełnienia warunków fizycznej realizacji na każdym etapie przekształcania zmiennych i funkcji.

W pracy pominięto zagadnienie realizacji syntezy torów o wspomnianych charakterystykach oraz zagadnienie wzorów przybliżonych, ułatwiających proces obliczeń. Odpowiednie wskazówki na ten temat można znaleźć w literaturze, np. w [2], [8], [10].

LITERATURA CYTOWANA W TEKŚCIE

1. N. I. Achiezer, Über einige Funktionen, welche in zwei gegebenen Intervallen am wenigsten von Null abweichen, Izw. AN SSSR, Otd. Matem. i Estestw. Nauk, seria 7, nr 9, 1932, s. 1163; nr 3, 1933, s. 309; nr 4, 1933, s. 499.
2. A. L. Feldsztejn, L. R. Jawicz, Sintiez czetyriochpolusnikow i wosmipolusnikow na SWCz, wyd. II, Izd. „Swiaz", Moskwa 1971.
3. W. Golde, C. Norek, S. Paszkowski, Zarys teorii aproksymacji i jej zastosowań w elektrotechnice, PWN, Warszawa 1958.
4. W. L. Gonczarow, Tieorija interpolirowanija i pribliżenija funkcij, Gostiechizdat, Moskwa 1954.
5. F. Kamiński, Cechy charakterystyczne torów niejednorodnych złożonych z odcinków jednorodnych, Przegl. Telekom., rok 38/32, nr 8, 1966, s. 225—229.
6. F. Kamiński, Warunki fizycznej realizacji pewnej klasy torów niejednorodnych, Rozpr. Elektrotechn., t. 13, nr 4, 1967, s. 627—648.
7. F. Kamiński, Warunki fizycznej realizacji torów niejednorodnych złożonych z odcinków jednorodnych, Arch. Elektrotechn., t. 17, nr 4, 1968, s. 767—785.
8. F. Kamiński, Wybrane aspekty syntezy torów niejednorodnych złożonych z odcinków jednorodnych, o charakterystyce Czebyszewa, Rozpr. Elektrotechn., t. 15, nr 1, 1969, s. 23—51.

9. F. Kamiński, K woprosu o sintezie czebyszewskich filtrujuszczich stupienczatych linij, Prace ITE PAN, nr 29, 1968.
10. G. L. Matthaei, L. Young, E. M. T. Jones, Microwave Filters, Impedance Matching Networks and Coupling Structures, Mc Graw-Hill, New York 1964.
11. G. L. Matthaei, Short-Step Chebyshev Impedance Transformers, IEEE Trans. on Microwave Theory and Techniques, MTT-14, nr 8, 1966, s. 372—383.

F. KAMIŃSKI

LOSS FUNCTIONS OF THE STEPPED LINES WITH CHEBYSHEV AND BUTTERWORTH TRANSMISSION CHARACTERISTIC

Summary

The paper is devoted to the discussion of determination of the loss function $L(s)$, eq. (1), corresponding to the formulated conditions, for a stepped line composed of lossless homogeneous sections. Much attention is given to the specific character of the problem under study. In particular it has been proposed to make use of the technological realizability criterion μ_d for a transmission line in cascade with the object of achieving the unique solution. The transformation technique of algebraic polynomials to physical realizability loss functions of an unhomogeneous line has been considered.

The optimization problem for some characteristic parameters has been formulated and solved (problem A, § 1). In particular the analytical expression of the Chebyshev loss function has been derived (eq. (19), (20, (28)), as well as the procedure has been discussed to look for the optimum solution.

On the basis of the procedure discused above, the general analytical expression of the Butterworth loss function has been derived (eq. (34)).

The paper should help on with the investigation of the same problem but under other conditions. Results of this work are applicable to the synthesis of eletromechanical filters and appropriate microwave systems.

F. KAMIŃSKI

FONCTIONS DE L'AFFAIBLISSEMENT EFFECTIF DES LIGNES INHOMOGÈNES COMPOSÉES DE TRONÇONS DE LIGNE HOMOGÈNE À CARACTÉRISTIQUE DU TYPE DE TCHEBYSHEFF ET DE BUTTERWORTH

Résumé

On présente dans l'oeuvre un procédé de recherche d'une fonction de l'affaiblissement effectif $L(s)$, f.(1), pour la ligne inhomogène composée de tronçons de ligne homogène sans pertes, correspondant aux conditions de transmission fixées. Des aspects spécifiques du problème en question sont l'object de la grande attention. En particulier on propose de mettre à profit un critère special μ_d, nommé critère de la réalisation technologique, pour créer la condition d'une solution unique et pareillement on considère une manière de la transformation de polynômes algébriques en fonction correspondant à une certaine fonction de l'affaiblissement effectif des lignes en question.

On met à l'étude et résout un certain problème de l'optimisation de quelques paramètres de la ligne inhomogène (le problème A, § 1). En particulier la forme analytique de la fonction de l'affaiblissement effectif du type de Tchebysheff a été déterminée (f. (19), (20), (28)) et la méthode de recherche de la solution optimale a été présentée.

En tenant compte de ces résultats on obtient la forme générale de la fonction de l'affaiblissement effectif du type de Butterworth (f. (34)).

Cette étude théorique peut aider à résoudre un problème semblable mais correspondant aux conditions différentes.

Lec conclusions de l'oeuvre conviennent à la synthèse des filtres électromécaniques et de certains composants à hyperfréquence.

F. KAMIŃSKI

BETRIEBSDÄMPFUNGSFUNKTIONEN DER LEITUNGSKETTEN, MIT TSCHEBYSCHEFF- UND BUTTERWORTH-VERHALTEN

Zusammenfassung

Auf Grund gegebener Übertragungsbedingungen wird ein Verfahren zur Ermittlung der Betriebsdämpfungsfunktion $L(s)$, Gl. (1), für die Synthese von Siebschaltungen untersucht, die sich aus in Kette geschalteten verlustfreien Leitungsstücken zusammensetzen. Besondere Aufmerksamkeit ist der Spezifik dieses Problems geschenkt worden, insbesondere der Einführung der Kenngröße für technologische Realisierbarkeit μ_d, um eine Eindeutigkeit des Problems herbeizuführen. Es wird auch ein Verfahren zur Umbildung algebraischer Polynome in physikalisch realisierbare, mit gegebenen Übertragungsbedingungen übereinstimmende Betriebsdämpfungsfunktionen der Leitungsketten behandelt. Ein Optimalisierungsproblem für einige Parameter der Leitungskette (Problem A, § 1) wird formuliert und gelöst. Die analytische Form der tschebyscheffschen Betriebsdämpfungsfunktion wird abgeleitet (Gl. (19), (20), (28)) sowie eine Untersuchungsmethode einer Optimallösung dargestellt.

Auf Grund der vorliegenden Ergebnisse wurde die allgemeine Form der Butterworths-Betriebsdämpfungsfunktion (Gl. (34)) abgeleitet.

Das Gesamtergebnis dieser Arbeit soll eine Hilfe beim Ermitteln einer analytischen Form für die Betriebsdämpfungsfunktion mit abweichenden Eigenschaften, als die obengenannten sein, da diese Schwierigkeiten bei der Erfüllung physikalischer Realisierungsbedingungen aufweisen.

Die theoretischen Erwägungen dieser Arbeit finden bei der Synthese elektromechanischer Siebketten sowie gewisser Ultrahochfrequenzanlagen Anwendung.

Ф. КАМИНЬСКИ

ФУНКЦИИ РАБОЧЕГО ЗАТУХАНИЯ СТУПЕНЧАТЫХ ЛИНИЙ С ХАРАКТЕРИСТИКАМИ ТИПА ЧЕБЫШЕВА И БАТТЕРВОРТА

Резюме

Рассмотрен способ определения функции рабочего затухания $L(s)$, ф-ла (1), ступенчатой линии, соответствующей поставленным условиям. Основное внимание сосредоточено на обсуждении специфики вопроса, в частности на необходимости введения критерия технологической реализуемости линии μ_d для обеспечения однозначности решения и на метод преобразования алгебраических полиномов в физически реализуемые функции рабочего затухания неоднородной линии, соответствующие поставленным условиям.

Сформулирован и решен вопрос оптимизации некоторых параметров линии (проблема А, § 1). В связи с этим найдено аналитическое выражение функции рабочего затухания типа Чебышева (ф-лы (19), (20), (28)) и рассмотрена задача поиска оптимального решения.

На основе предложенного метода анализа выведен общий вид функции рабочего затухания типа Баттерворта (ф-ла (34)).

Теоретические выводы работы должны помочь в поисках аналитического выражения функции рабочего затухания ступенчатых линий при иных исходных условиях, так как решение этой задачи связано с известными трудностями, обусловленными необходимостью поэтапного учета условий физической реализуемости в процессе преобразования переменных и функций.

Представленные результаты находят применение в синтезе электромеханических цепочечных фильтров и некоторых устройств свч.

ROZPRAWY ELEKTROTECHNICZNE 1974, 20, z. 1, ss. 195-207

621.318.5

Niektóre zagadnienia analizy technicznej konstrukcji przekaźników

JAN ŁYSKANOWSKI (WARSZAWA)
Instytut Elektrotechniki

Otrzymano 20.2.1973

W artykule przedstawiono analizę konstrukcji przekaźników wykorzystując elementy programowania liniowego. Wyznaczono obszary prawidłowego działania oraz przydatności konstrukcyjnej przekaźników. Przedstawiona w artykule metoda stwarza warunki do obiektywnej oceny jakości i niezawodności przekaźników.

1. WSTĘP

Analiza techniczna konstrukcji przekaźnika jest jednym z etapów procesu realizacji, od którego może zależeć w znacznym stopniu jakość przekaźnika.

W artykule niniejszym podano definicje obszarów działania przekaźnika określone na podstawie własności konstrukcyjnych przekaźnika, w odróżnieniu od zdefiniowanych w pracy [1] obszarów zadziałań przekaźników na podstawie ich charakterystyki działania.

Określenia podane w niniejszym artykule stanowią punkt wyjścia do wyznaczania takich wartości parametrów pracy, które z punktu widzenia określonego kryterium można uważać za optymalne.

2. STAN ZAGADNIENIA

Analiza ogólna konstrukcji przekaźników elektromechanicznych pozwala twierdzić, że istnieje duża ilość różnych wykonań i odmian konstrukcyjnych przekaźników, produkowanych często po kilkaset lub mniej sztuk rocznie; przy tym niektóre z tych przekaźników, szczególnie starszych typów, są zaprojektowane nieekonomicznie, są mało dokładne, mają znaczne masy i wymiary. Zagadnienia te występują szczególnie jaskrawo w zakresie przekaźników trakcyjnych.

Stosowane dotychczas metody obliczeń przekaźników pozwalają przeważnie na obliczenie parametrów przekaźników o wartościach lokalnie optymalnych.

Jedna z takich metod jest przedstawiona w pracy N. E. Łysowa i A. W. Kurnosowa [7].

Ze stosunku mechanicznej pracy wykonanej przez zworę przekaźnika do jego objętości żelaza i miedzi, wyprowadzono wzory określające zależności między wymiarami elektromagnesu:

$$X_o = \frac{d}{D}, \quad Y = \frac{l_c}{D}, \tag{1}$$

w których d jest wielkością rdzenia magnetycznego, D średnicą cewki, natomiast l_c jest długością cewki.

Dla wartości $X_o = 0{,}57 \ldots 0{,}7$ oraz $Y = 1{,}5 \ldots 3$ uzyskuje się optymalne wykorzystanie materiałów czynnych przekaźnika, przy ustalonych wartościach innych jego parametrów.

W odróżnieniu od tej metody, przedstawiony poniżej sposób obliczeń pozwala na uzyskanie parametrów przekaźnika o wartościach globalnie optymalnych lub optymalnych parametrach zgodnie z założonym z góry kryterium.

W przedstawionej w artykule analizie założono, że znany jest schemat przekaźnika oraz znane są rodzaje stosowanych elementów i technologia wykonywania przekaźnika. Przy dalszym założeniu, że znane są ustalone wartości parametrów $c_l (l = 1, 2, \ldots, m)$ i możliwe realizacje zmiennych decyzyjnych $x_i (i = 1, 2, \ldots, n)$ bądź — jeżeli wartości tych wielkości podlegają rozrzutom — znane są wartości kresu dolnego i górnego ich możliwych zmian oraz ewntualnie wartości średniej (znamionowej), model zagadnienia w takim przypadku staje się modelem strategicznym. Jeżeli np. jako zmienną decyzyjną przyjęto docisk styków, a jako parametr — długość szczeliny powietrznej, to w modelu strategicznym nie traktuje się ich jako zmiennych losowych (co należałoby uczynić w modelu probabilistycznym), lecz wykorzystuje się tylko to, że wartości F i l_p są zadane poprzez zbiory wartości, jakie mogą przyjmować w okresie eksploatacji.

Przy wyprowadzaniu wzorów korzystano z danych przedstawionych w literaturze [5, 7, 9].

3. OKREŚLENIE OBSZARÓW PRAWIDŁOWEGO DZIAŁANIA I PRZYDATNOŚCI KONSTRUKCYJNEJ PRZEKAŹNIKA

Własności konstrukcyjne przekaźnika zależą przede wszystkim od parametrów elementów zastosowanych w przekaźniku, od stanów pracy przekaźnika oraz od wyboru punktu pracy.

W zależności od stosunku zachodzącego pomiędzy wymaganiami stawianymi przekaźnikowi a stopniem realizacji tych wymagań można mówić o określonym stanie pracy przekaźnika.

Po doprowadzeniu do przekaźnika wielkości zasilającej (prądu lub napięcia) o określonej wartości ma on zmienić swój stan pracy, tzn. ze stanu początkowego ma przejść do stanu zadziałania. Można więc powiedzieć, że przekaźnik realizuje na wyjściu funkcję logiczną o wartościach 0 i 1. Zakładamy przy tym, że przekaźnik zmienia swój stan w sposób bezzwłoczny.

Zmiana stanu pracy przekaźnika następuje, gdy wartość wielkości zasilającej przekroczy wartość rozruchową przekaźnika. Wielkości zasilające przekaźnik mogą mieć charakter dyskretny lub ciągły.

Przekaźnik ma spełnić szereg wymagań podanych we właściwych przepisach i normach.

Funkcję, którą ma realizować przekaźnik zgodnie ze stawianymi mu w przepisach wymaganiami stanowi na ogół funkcja logiczna n argumentów x_i, którą określono przez

$$\varphi_w(x_i) \quad \underset{i=1,2,\ldots,n}{x_i} \in (0, 1). \tag{2}$$

Natomiast funkcję, którą rzeczywiście realizuje przekaźnik oznaczono przez:

$$\varphi_r(x_i),$$

przy czym parametry opisujące wielkości zasilające wejściowe i wyjściowe określono za pomocą ciągów

$$\{b_{1i}\}\{b_{2i}\} \qquad (i = 1, 2, \ldots, n). \tag{3}$$

Założono, że:

a) przekaźnik zabezpieczeniowy znajduje się w stanie prawidłowym pracy, jeżeli jest spełniona tożsamość

$$\varphi_w(x_i) = \varphi_r(x_i), \tag{4}$$

gdy

$$b_{2i\min} \leqslant b_{2i} \leqslant b_{2i\max}, \tag{5}$$

przy

$$b_{1i\min} \leqslant b_{1i} \leqslant b_{1i\max}, \tag{6}$$

przy czym $b_{...\min}$ i $b_{...\max}$ — dopuszczalne granice parametrów b_{1i} oraz b_{2i}, określone w odnośnych wymaganiach;

b) przekaźnik zabezpieczeniowy znajduje się w stanie nieprawidłowym pracy, jeżeli

$$\varphi_w(x_i) \neq \varphi_r(x_i), \tag{7}$$

lub jeżeli istnieje taka, spełniająca (6) kombinacja wartości b_{1i}, przy której jest spełniona przynajmniej jedna z relacji

$$b_{2i} > b_{2i\max},$$

$$b_{2i} < b_{2i\min}.$$

Z przytoczonych określeń (4) ... (7) wynika, że w dowolnej chwili t istnieją dwa jedynie możliwe i wzajemnie się wykluczające stany przekaźnika: prawidłowy stan pracy i nieprawidłowy stan pracy.

Stan przekaźnika może ulec zmianie nie tylko przez wielkość zasilającą wejściową, lecz również za pośrednictwem zmiennych opisujących wielkości fizyczne przekaźnika. Zmienne te tworzą ciągi

$$\{c_{1i}\} \qquad (i = 1, 2, \ldots, r) \tag{8}$$

parametrów wewnętrznych charakteryzujących elementy stosowane w konstrukcji przekaźnika, oraz ciągi

$$\{c_{2i}\} \qquad (i = 1, 2, \ldots, q) \tag{9}$$

parametrów zewnętrznych, które opisują własności źródeł zasilających, warunki otoczenia, przy czym $r+q = n$.

Przestrzeń tych parametrów można interpretować geometrycznie jako n-wymiarową przestrzeń euklidesową, przy czym współrzędne każdego punktu tej przestrzeni są równe odpowiednim wartościom parametrów pracy i określają położenie wektora charakteryzującego konstrukcję przekaźnika.

Każdemu egzemplarzowi przekaźnika w określonej chwili t można przypisać jednoznacznie jeden z punktów przestrzeni parametrów przekaźnika.

Punkt w przestrzeni o współrzędnych odpowiadających ustalonym parametrom pracy

$$c_{11} = \alpha_1, \quad c_{12} = \alpha_2, \quad \ldots, c_{ir} = \alpha_r \ldots c_{21} = \alpha_{r+1}, \quad c_{22} = \alpha_{r+2}, \ldots, c_{2q} = \alpha_n$$

nazywa się punktem pracy przekaźnika, jeżeli jest spełniona inkluzja

$$\{w(\alpha_1, \alpha_2, \ldots, \alpha_n)\} \in M, \tag{10}$$

przy czym M — przestrzeń parametrów.

Punkt ten nazywa się znamionowym punktem pracy, jeżeli znamionowe wartości wynoszą

$$\alpha_{1zn}, \alpha_{2zn}, \ldots, \alpha_{nzn}. \tag{11}$$

Zbiór punktów w określony za pomocą relacji

$$A_{pr} = \{w{:}F(w) = 1,\}, \quad A_{pr} \subset M \tag{12}$$

nazwano obszarem prawidłowego działania A_{pr} przekaźnika zabezpieczeniowego, przy czym $F(w) = 1$ jest zdaniem określającym stan prawidłowego działania.

Relację tę można wysłowić: obszarem prawidłowego działania A_{pr} przekaźnika zabezpieczeniowego nazwiemy podzbiór punktowy przestrzeni parametrów M, zawierający jako elementy jedynie te wszystkie punkty pracy w, którym odpowiada stan prawidłowego działania danego przekaźnika.

Zbiór punktów w określony za pomocą relacji

$$A_p = \{w{:}F_p(w) = 1\}, \quad A_p \subset M \tag{13}$$

nazwano obszarem przydatności konstrukcyjnej przekaźnika zabezpieczeniowego, przy czym $F_p(w) = 1$ jest zdaniem określającym stan przydatności konstrukcyjnej wyznaczonej np. na podstawie przyjętego w założeniach kryterium przydatności.

Relację tę można wysłowić: obszarem przydatności konstrukcyjnej A_p przekaźnika zabezpieczeniowego nazwiemy podzbiór punktowy przestrzeni parametrów M, zawierający jako elementy jedynie te wszystkie punkty pracy w, którym odpowiada stan przydatności konstrukcyjnej przekaźnika.

4. ZAPIS ANALITYCZNY OBSZARÓW PRAWIDŁOWEGO DZIAŁANIA I PRZYDATNOŚCI KONSTRUKCYJNEJ PRZEKAŹNIKA

Założenia konstrukcyjne na przekaźniki w ich części dotyczącej elektrycznych własności, formułują zazwyczaj trzy zasadnicze grupy wymagań:

a) przekaźnik (lub jego elementy) powinien przyjmować jeden z dwóch możliwych stanów stabilnych przy odpowiednich wielkościach zasilających wejściowych,

b) parametry wielkości zasilających na wyjściu, w założonych warunkach pracy powinny być zawarte w dopuszczalnych przedziałach,

c) wartości parametrów charakteryzujących czynniki wymuszające oddziaływujące na elementy układu, powinny być ograniczone do takich poziomów, by intensywność uszkodzeń zupełnych i prędkość zmian starzeniowych nie przekroczyły dopuszczalnych wartości.

Dla elektromagnetycznych przekaźników, typ a) wymagań sprowadzi się do żądania co najmniej dwóch warunków: przyciągania zwory przekaźnika i zwolnienia tej zwory.

W przypadku przekaźników na elementach półprzewodnikowych warunkami będą zatkania tranzystora i nasycenia tranzystora. Warunki te noszą nazwę układu więzów funkcjonalnych lub warunków działania przekaźnika. Ich liczba d zależy od rodzaju przekaźnika i na ogół waha się od 2 do 4.

Warunki typu b) noszą nazwę układu więzów realizacyjnych lub warunków zasilania. Ich liczba s zależy od wymagań stawianych przekaźnikowi. Jeżeli nieistotne są własności dynamiczne przekaźnika, występują co najwyżej dwa warunki typu b) wyrażające wymagania na amplitudę sygnałów wejściowych dla stanów ustalonych przekaźnika.

Wymagania typu c) mają charakter prostych ograniczeń mocy wydzielanej w elementach przekaźnika lub ograniczeń napięciowych, prądowych termicznych. Ograniczenia tego typu wynikają z danych katalogowych lub wymagań normalizacyjnych. Często są związane z wytrzymałością elementów lub z intensywnością występowania uszkodzeń zupełnych. Ich liczba z nie podlega żadnych prawidłowościom i silnie zależy od własności przekaźnika i założeń konstrukcyjnych.

Omówione wyżej wymagania tworzą układ nierówności określających prawidłowe działanie przekaźnika

$$\begin{gathered} g_k(X, C) \leqslant 0, \\ k = 1, 2, \ldots, d+s+z \end{gathered} \tag{14}$$

determinujący zbiór rozwiązań dopuszczalnych

$$A = \{X : G_k(X, C) \leqslant 0\}, \tag{15}$$

stanowiący analityczny zapis związków pomiędzy parametrami pracy, których spełnienie jest konieczne dla poprawnej realizacji wszystkich wymagań stawianych temu przekaźnikowi. We wzorach (14) i (15):

$X = (x_1, x_2, \ldots, x_n)$ jest wektorem zmiennych decyzyjnych,

$C = (c_1, c_2, \ldots, c_m)$ jest wektorem parametrów.

Ostatnia cecha, w stosunku do której wymagania nie zostały sformułowane w postaci nierówności wchodzącej w skład układów więzów, stanowi czynnik rządzący wyborem optymalnego rozwiązania spośród rozwiązań wchodzących w skład zbioru rozwiązań dopuszczalnych.

Cechę tę uzależnia się od zmiennych oraz parametrów i przyjmuje jako funkcję celu $f(X, P)$. Żąda się przy tym, aby wartość funkcji celu była największa lub najmniejsza na zbiorze rozwiązań dopuszczalnych.

Funkcje $g_k(X, P)$ i $f(X, P)$ stanowią model matematyczny układu.

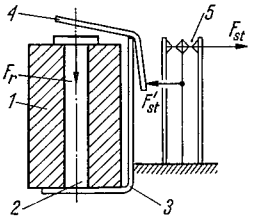
Z powodu nieuniknionych uproszczeń dokonywanych przy wyznaczaniu poszczególnych relacji, model ten tylko z pewnym przybliżeniem opisuje rzeczywisty obszar A.

Jakie wielkości, od których zależą funkcje wchodzące w skład układu więzów i zbioru funkcji celu, należy traktować jako zmienne decyzyjne, a jakie jako parametry, jest decyzją wstępną konstruktora.

Na przykład, przy projektowaniu przekaźników typu przedstawionego na rys. 1 istotnym zagadnieniem jest obliczenie podstawowych wymiarów magnetowodu i cewki wzbudzającej. Dlatego funkcją celu może być zależność wyrażająca stosunek pracy mechanicznej

wykonywanej przez przekaźnik do sumy objętości materiałów czynnych. Zależność ta określana za pomocą współczynnika geometrycznego jest funkcją pracy mechanicznej wykonywanej przez przekaźnik, objętości stali magnetowodu przekaźnika i objętości uzwojenia cewki.

W celu uproszczenia analizy w przedstawionym dalej przykładzie wyznaczenia obszarów prawidłowego działania i przydatności konstrukcyjnej przekaźnika przyjęto szereg



Rys. 1. Szkic konstrukcji przekaźnika pośredniczącego prądu stałego
1 — cewka, 2 — rdzeń magnetyczny, 3 — jarzmo, 4 — zwora, 5 — zestyki

uproszczeń, które pozwoliły na zobrazowanie wyznaczonych obszarów w przestrzeni dwuwymiarowej.

5. PRZYKŁAD WYZNACZENIA OBSZARÓW PRAWIDŁOWEGO DZIAŁANIA I PRZYDATNOŚCI KONSTRUKCYJNEJ PRZEKAŹNIKA POŚREDNICZĄCEGO (PRĄDU STAŁEGO)

Obszar prawidłowego działania i przydatności konstrukcyjnej przekaźnika pośredniczącego prądu stałego wyznaczono dla konstrukcji tegoż przekaźnika pokazanej na rys. 1. W oparciu o wzory podane w literaturze, wprowadzono szereg warunków działania i warunków ograniczających.

5.1. Warunek rozruchu przekaźnika

Rozruch przekaźnika następuje, gdy na zworę działa siła F_r:

$$F_r = k \cdot n \cdot F_{st}, \tag{16}$$

przy czym:

k — współczynnik mechanicznego przełożenia siły F_{st},

F_{st} — siła z jaką dociskane są styczki zestyku rozwiernego w przekaźniku niepobudzonym,

n — liczba zestyków.

Dla uzyskania siły rozruchowej F_r konieczna jest określona wartość strumienia magnetycznego ϕ przenikającego przez szczelinę powietrzną.

Według W. Kosteckiego [5] wartość siły rozruchowej rozwijanej przez elektromagnes w osi jego szczeliny „roboczej" można wyznaczyć ze wzoru

$$F_r = \frac{1}{2}\left(\frac{\phi_\delta}{\Lambda_\delta}\right)^2 \left|\frac{d\Lambda_\delta}{d\delta}\right|, \tag{17}$$