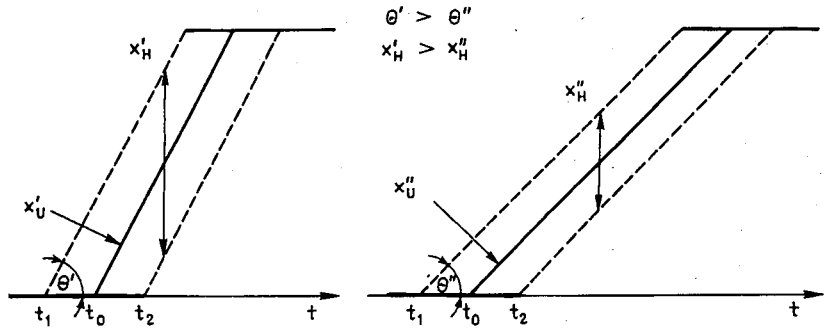
$$H(t) \leqslant H_{max} = max|t_2 - t_1| \tag{15}$$

$$x_H(t) = H(t) \cdot \frac{\mathrm{d}x_U(t)}{\mathrm{d}t} \tag{16}$$

Dla przypadku sygnału użytecznego z rys. 4: $x_H(t) = H(t) \cdot \mathrm{tg}\,\theta$. Składowa x_H zmienia amplitudę sygnału użytecznego x_U wobec jego „niestałości" w każdej chwili t. Jest to błąd wynikający z przyjęcia w każdej chwili t_0 zamiast amplitudy B amplitudę A lub C, którą sygnał w nieobecności zakłóceń horyzontalnych osiągnąłby w chwili t_2 lub powinien osiągnąć w chwili t_1 odpowiednio.

Występowanie zakłóceń horyzontalnych uwidacznia się na zboczach obserwowanego sygnału. Im stromość zbocza jest większa, tym wpływ ten jest silniejszy (składowa amplitudowa x_H staje się większa, zgodnie z zależnością (16)). Ilustruje to rys. 5. Na płaskich odcinkach sygnału użytecznego, zakłócenia horyzontalne nie zniekształcają jego amplitudy. Może się wówczas ujawnić jedynie wpływ zakłóceń wertykalnych.



Rys. 5. Wpływ zakłóceń horyzontalnych na sygnały o różnej dynamice przy tej samej niestałości położenia sygnału użytecznego

Uśrednianie sygnału pomiarowego stosowane w celu ograniczania zakłóceń wertykalnych, zastosowane w tym samym celu, gdy zakłócenia są horyzontalne, nie daje zbieżności do sygnału użytecznego. Rezultatem uśredniania jest sygnał zniekształcony, o wygładzonym kształcie. Jednakże sygnał ten, w wyniku odpowiedniego przetworzenia umożliwia odtworzenie niezakłóconeogo sygnału użytecznego [5,7].

Uśredniona wartość sygnału z fluktuacją położenia wynosi:

$$\bar{x}(t) = \int_{-\infty}^{\infty} x_U(t - \tau) \cdot f_H(\tau)\,\mathrm{d}\tau \tag{17}$$

czyli jest splotem niezakłóconego sygnału użytecznego x_U i funkcji gęstości prawdopodobieństwa zakłóceń horyzontalnych f_H:

$$\bar{x}(t) = x_U(t) * f_H(t). \tag{18}$$

W dziedzinie częstotliwości odpowiada temu zależność:

$$\bar{X}(f) = X_U(f) \cdot \mathscr{F}\{f_H(t)\}. \tag{19}$$

Uśrednianie sygnału z fluktuacją położenia jest więc równoważne filtracji niezakłóconego sygnału użytecznego. Charakterystyka filtru jest równa transformacie Fouriera funkcji gęstości prawdopodobieństwa zakłóceń horyzontalnych. Jeśli znana jest funkcja f_H, lub można ją estymować, wówczas mając uśrednioną wartość sygnału pomiarowego $\bar{x}$ można określić sygnał użyteczny x_U, uwolniony od zakłóceń horyzontalnych. Wyznacza się go na podstawie zależności (18) stosując algorytm rozplatania lub z zależności (19) stosując odwrotną transformatę Fouriera.

W przypadku gaussowskich zakłóceń horyzontalnych zachodzą związki:

$$f_H(t) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \exp\left(\frac{t^2}{2 \cdot \sigma^2}\right) \tag{20}$$

$$\mathscr{F}\{f_H(t)\} = \exp(-2\pi^2 f^2 \sigma^2) \tag{21}$$

$$x_U(t) = \mathscr{F}^{-1}\left\{\frac{\bar{X}(f)}{\mathscr{F}\{f_H(t)\}}\right\} = \mathscr{F}^{-1}\{\bar{X}(f) \cdot \exp(2\pi^2 f^2 \sigma^2)\}. \tag{22}$$

4. OGRANICZANIE ZAKŁÓCEŃ Z PRZESTRZENI DWUWYMIAROWEJ

Gdy zakłócenia są z przestrzeni dwuwymiarowej (zakłócenia wertykalne i horyzontalne występują jednocześnie) wówczas w sygnale pomiarowym x, oprócz sygnału użytecznego x_U występują dodatkowo składowe amplitudowe x_W i x_H (por. wzór 3a, 3b, 16). Sygnał zakłóceń (z) wynosi wówczas:

$$z(t) = x_W(t) + H(t)\frac{\mathrm{d}x_U(t)}{\mathrm{d}t}. \tag{23}$$

Algorytm ograniczania sygnału (z) jest dwuetapowy. W pierwszym etapie jest ograniczany wpływ zakłóceń horyzontalnych. W tym celu realizowane są operacje:

- uśrednianie sygnału pomiarowego,
- estymacja funkcji gęstości prawdopodobieństwa f_H,
- wyznaczenie sygnału x_0, uwolnionego od zakłóceń horyzontalnych, z następującej zależności (filtracja homomorficzna względem splotu):

$$\bar{x}(t) = x_0(t) * f_H(t). \tag{24}$$

W drugim etapie jest ograniczany wpływ zakłóceń wertykalnych. W tym celu sygnał x_0 należy poddać przetworzeniu według jednej z przedstawionych uprzednio metod (uśrednianie, metody korelacyjne, filtracja).

Aby na podstawie sygnału $\bar{x}$ wyodrębnić z zależności (24) sygnał x_0 uwolniony od zakłóceń horyzontalnych, musi być znana funkcja gęstości prawdopodobieństwa tych zakłóceń f_H. W przypadku braku danych o właściwościach statystycznych zakłóceń, f_H należy estymować.

Sygnał zakłóceń $z(t)$ jest sumą dwóch składników losowych (por. zal. 23). Odpowiadająca mu funkcja gęstości prawdopodobieństwa $f(z)$ jest więc splotem funkcji gęstości prawdopodobieństwa obu składników (x_W i H – wielkości losowe):

$$f(z) = f(x_W) * f(H \cdot \dot{x}_U) \tag{25}$$

Przekształcając funkcję $f(H \cdot \dot{x}_U)$ [2], można wyznaczyć $f_H(t)$:

$$f_H(t) = f(H) = \dot{x}_U(t) \cdot f(H \cdot \dot{x}_U). \tag{26}$$

Algorytm estymacji $f_H(t)$ składa się z następujących kroków:

a_1) estymacja $f(z)$ – na podstawie histogramu łącznych zakłóceń amplitudy sygnału odniesienia.

a_2) estymacja $f(x_W)$ – na podstawie histogramu zakłóceń amplitudy sygnału odniesienia, w zakresach, w których nie ujawnia się wpływ zakłóceń horyzontalnych (tj. gdy sygnał ma stałą wartość lub na płaskich odcinkach sygnału zmiennego),

a_3) wyznaczenie funcki $f(H \cdot \dot{x}_U)$ – z zależności (25) (rozplot),

a_4) wyznaczenie funkcji $f_H(t)$ – z zależności (26).

5. UWAGI KOŃCOWE

Zagadnienia przedstawione w artykule dotyczą metod wpływania na jakość wyników pomiarów w komputerowych urządzeniach pomiarowych nazwanych urządzeniami pomiarowymi trzeciej generacji. Operacje na sygnale pomiarowym w tych układach są wykonywane w osi amplitudy oraz w osi czasu i dlatego niedokładność pomiaru należy traktować jako wielkość dwuwymiarową o składowej wertykalnej i horyzontalnej. Każda ze współrzędnych składa się z części deterministycznej oraz z części losowej. Części te stanowią odpowiednio błąd systematyczny oraz przypadkowy (nazywany zakłóceniami).

W pracy przedstawiono rodzaje operacji jakie mogą być wykonywane na sygnale pomiarowym, w celu ograniczenia zakłóceń wertykalnych i horyzontalnych. Jeśli zakłócenia występujące w sygnale pomiarowym wpływają addytywnie na amplitudę sygnału użytecznego, to mogą być one ograniczone w rezultacie zastosowania:

- uśredniania sygnału pomiarowego według odpowiedniego algorytmu,
- metody korelacyjnej,
- filtracji liniowej.

Jeśli natomiast wpływ zakłóceń na amplitudę sygnału użytecznego jest multiplikatywny lub splotowy (oddziaływanie zakłóceń horyzontalnych), wówczas mogą być one ograniczone w wyniku filtracji nieliniowej (homomorficznej).

W UP3G operacje związane z kształtowaniem ich właściwości metrologicznych oraz z otrzymywaniem wyniku pomiaru są realizowane głównie metodami programowymi. Aby rozszerzyć możliwości w zakresie rozwiązywania wyżej wymienionych zagadnień, układy te wyposaża się w procesory sygnałowe.

BIBLIOGRAFIA

1. K. G. Beauchamp: *Przetwarzanie sygnałów metodami analogowymi i cyfrowymi,* WNT, Warszawa 1978
2. J. S. Bendat, A. G. Piersol: *Random Data. Analysis and Measurement Procedures.* John Wiley and Sons, New York 1986
3. W. Borodziewicz, K. Jaszczak: *Cyfrowe przetwarzanie sygnałów.* WNT, Warszawa 1987
4. L. E. Franks: *Teoria sygnałów,* PWN, Warszawa 1975
5. Y. Guo, W. Su: *Automatic time – domain measurement system.* IEEE Trans. Instrum. Meas. 1989, nr 2
6. A. Kareem, S. K. Balakrishnan: *Impact of digital signal processing on measurement and test instruments.* IEEE Trans. Instrum. Meas., 1988, nr 4
7. T. Miki, H. Yamaguchi, Y. Nagaki: *An accurate wide – band automatic waveform analyzer.* IEEE Trans. Instrum. Meas. 1977, nr 4
8. W. Nowakowski: *Tendencje rozwojowe w dziedzinie testowania i kontroli jakości wyrobów (CAT–CAQ). Komputerowe systemy pomiarowe, wirtualne przyrządy pomiarowe.* PAK 1991, nr 1
9. W. Nowakowski: *Systemy pomiarowe z komputerami osobistymi.* PAK, 1991, nr 2
10. A. V. Oppenheim, R. W. Schafer: *Cyfrowe przetwarzanie sygnałów.* WKŁ, Warszawa, 1979
11. A. V. Oppenheim: *Sygnały cyfrowe. Przetwarzanie i zastosowania.* WNT, Warszawa 1982
12. E. Ozimek: *Podstawy teoretyczne analizy widmowej sygnałów.* PWN, Warszawa–Poznań 1985
13. S. M. Riad, N. S. Nahman: *Application of the homomorphic deconvolution for the separation of TDR signals occuring in overlapping time windows.* IEEE Trans. Instrum. Meas. 1976, nr 4

A. DOMAŃSKA

PROCESSES REDUCING INFLUENCE OF DISTORTIONS IN COMPUTER MEASUREMENT INSTRUMENTS

Summary

This paper presents a concept of the vector form of inaccuracy in the measurement instruments of the third generation. It describes the methods, which can be used to reduce distortions of the utility signal. Three cases of the influence of distortions are distinguished: vertical, horizontal and joint vertical – horizontal. Three types of relations, which can occur in time domain between utility signal and distorsions are taken into account: addition, multiplication and convolution.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1992, 38, Z. 2–3, ss. 405–417

621.376.233

Błędy pomiarowe detektorów fazoczułych spowodowane napięciami i prądami niezrównoważenia

ANATOL JAWOREK

Instytut Maszyn Przepływowych, Polska Akademia Nauk, Gdańsk

Otrzymano 1991.10.24

Autoryzowano do druku 1991.12.6

Dokonano teoretycznej oceny błędu pomiaru składowych rzeczywistej i urojonej immitancji dwuelementowych, liniowych dwójników biernych za pomocą detektorów fazoczułych, spowodowanego przez wejściowe napięcia i prądy niezrównoważenia wzmacniaczy operacyjnych tworzących poszczególne stopnie toru pomiarowego. Porównano pod tym względem trzy typy detektorów fazoczułych. Z wyprowadzonych zależności wynika, że pomiar za pomocą detektora fazoczułego z próbkowaniem jest najmniej wrażliwy na wejściowe napięcia i prądy niezrównoważenia wzmacniaczy operacyjnych. Błąd pomiarowy w detektorze fazoczułym z całkowaniem w stałym przedziale jest nie większy niż w detektorze fazoczułym z całkowaniem w zmiennym przedziale.

1. WSTĘP

Detektor fazoczuły jest układem, na którego wyjściu otrzymuje się sygnał stałonapięciowy proporcjonalny do amplitudy detekowanego sygnału zmiennonapięciowego i będący funkcją jego przesunięcia fazowego względem sygnału odniesienia. Detektory fazoczułe synchroniczne są stosowane do detekcji sygnałów na tle szumów lub do detekcji składowych rzeczywistej i urojonej immitancji zespolonej dwuelementowych, liniowych dwójników biernych. Potrzeba rozróżnienia składowych immitancji występuje w przypadku diagnostyki dwójników biernych w sieciach elektrycznych [1,2] lub w pomiarach sygnałów generowanych przez parametryczne przetworniki pomiarowe [3], posiadające jednocześnie składową rzeczywistą i urojoną immitancji (np. w przetwornikach indukcyjnych, albo w przetwornikach pojemnościowych o dużej upływności lub rezystancyjnych o niepomijalnej susceptancji).

Detektory fazoczułe współpracujące z przetwornikami pomiarowymi umożliwiają pomiar zmian wartości wielkości mierzonej w czasie nie przekraczającym kilku okresów sygnału zasilającego przetwornik pomiarowy, co pozwala na dostatecznie wierne odtworzenie jej przebiegu czasowego. W niektórych specjalnych rozwiązaniach czas pomiaru może być skrócony do jednego lub nawet do pół okresu sygnału pobudzającego.

Wyróżnić można trzy podstawowe typy integracyjnych detektorów fazoczułych do bezpośredniego pomiaru składowych immitancji, w których pomiar polega na całkowaniu sygnału odpowiedzi dwójnika na pobudzenie harmoniczne w ściśle określonym dla danej składowej przedziale kąta fazowego:

1. Detektor fazoczuły z całkowaniem sygnału odpowiedzi w stałym przedziale, wyznaczonym przez fazę sygnału odniesienia i wynoszącym połowę okresu sygnału odniesienia [4–6].

2. Detektor fazoczuły z całkowaniem sygnału odpowiedzi w zmiennym przedziale, zależnym od fazy sygnału odniesienia (pobudzenia dwójnika) i fazy odpowiadającej wartości maksymalnej sygnału odpowiedzi badanego dwójnika [7,8].

3. Detektor fazoczuły z próbkowaniem sygnału odpowiedzi (sample–and–hold), w którym momenty próbkowania wyznaczane są przez sygnał odniesienia [9–11].

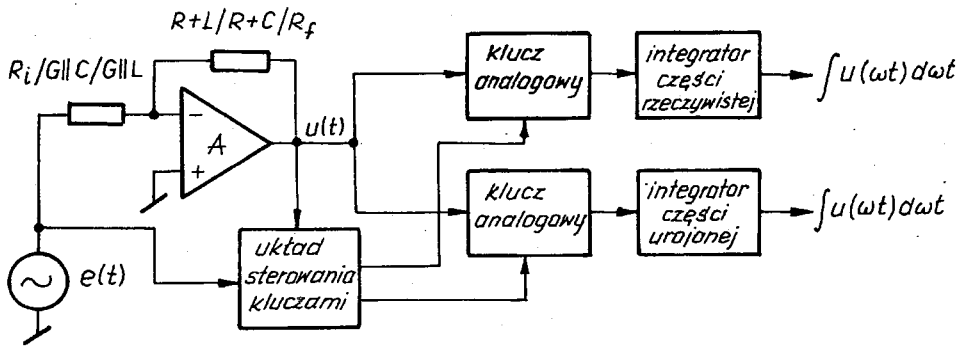
Pomimo licznych zastosowań detekcji fazoczułej w aparaturze naukowej i przyrządach pomiarowych [12–17], niewiele jest opracowań analizujących problem dokładności przetwarzania sygnałów za ich pomocą [2,9,18–21]. Zagadnienie błędów przetwarzania spowodowanych składowymi harmonicznymi zawartymi w mierzonym sygnale i niedokładnym wyznaczeniem przedziałów całkowania omówione zostało w [18,19]. Problem dokładności pomiaru składowych immitancji w układzie wzmacniacza rzeczywistego tzn. o skończonym wzmocnieniu i ograniczonym paśmie rozważony został w [20,21]. Wpływ napięć i prądów niezrównoważenia wzmacniaczy operacyjnych w torze pomiarowym na wynik pomiaru składowych immitancji dla poszczególnych typów detektorów fazoczułych nie był dotychczas w literaturze rozważany.

W pracy przeprowadzono teoretyczne oszacowania błędów pomiaru składowych immitancji dwuelementowych dwójników biernych spowodowanych napięciami i prądami niezrównoważenia wzmacniaczy operacyjnych występujących w torze pomiarowym detektorów fazoczułych i uczestniczących w przetwarzaniu sygnału pomiarowego.

2. UKŁADY POMIAROWE DWÓJNIKÓW BIERNYCH

Niezależnie od wielu możliwych praktycznych realizacji detekcji fazoczułej z przeznaczeniem do pomiaru składowych immitancji dwuelementowych dwójników biernych, podstawowe operacje wykonywane na sygnale odniesienia i odpowiedzi pozostają niezmienione. Schemat blokowy układu zawierającego detektor fazoczuły i mierzącego składowe immitancji dwójników biernych przedstawiono na rys. 1.

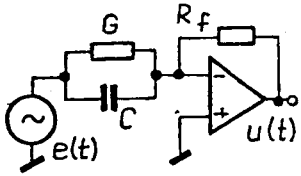
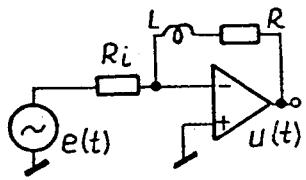
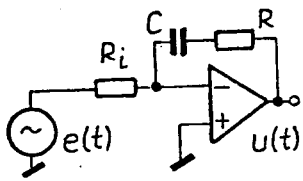
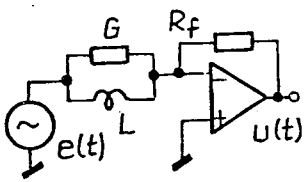
Pomiar składowych immitancji dwójników biernych za pomocą synchronicznych detektorów fazoczułych polega na całkowaniu sygnału odpowiedzi dwójnika na pobudzenie harmoniczne w przedziałach kąta fazowego zależnych od mierzonej składowej immitancji. Wartość całki jest proporcjonalna do mierzonej składowej immitancji. Mierzony dwójnik załączony jest w obwodzie wzmacniacza operacyjnego w sposób przedstawiony w tabeli 1. Dwójnik równoległy G||C lub G||L załącza się na wejściu wzmacniacza operacyjnego a w pętli sprzężenia zwrotnego znajduje się



Rys. 1. Schemat blokowy układu do pomiaru składowych immitancji dwójników za pomocą detekcji fazoczułej

Tabela 1

Podstawowe układy pomiarowe immitancji dwójników

Nr Ozn.	Układ pomiarowy	Odpowiedź czasowa / Transmitancja	Moduł / Przesun. fazowe
1 G||C		$u(t) = -ER_f\|Y\|\sin(\omega t + \varphi)$ $\frac{\hat{u}}{\hat{e}} = -R_f(G + j\omega C)$	$\|Y\| = \sqrt{G^2 + \omega^2C^2}$ $\varphi = \operatorname{atn}\frac{\omega C}{G}$
2 R+L		$u(t) = \frac{-E}{R_i}\|Z\|\sin(\omega t + \varphi)$ $\frac{\hat{u}}{\hat{e}} = \frac{-1}{R_i}(R + j\omega L)$	$\|Z\| = \sqrt{R^2 + \omega^2L^2}$ $\varphi = \operatorname{atn}\frac{\omega L}{R}$
3 R+C		$u(t) = \frac{-E}{R_i}\|Z\|\sin(\omega t - \varphi)$ $\frac{\hat{u}}{\hat{e}} = \frac{-1}{R_i}(R - j/(\omega C))$	$\|Z\| = \sqrt{R^2 + 1/(\omega C)^2}$ $\varphi = \operatorname{atn}\frac{1}{R\omega C}$
4 G||L		$u(t) = -ER_f\|Y\|\sin(\omega t - \varphi)$ $\frac{\hat{u}}{\hat{e}} = -R_f(G - j/(\omega L))$	$\|Y\| = \sqrt{G^2 + 1/(\omega L)^2}$ $\varphi = \operatorname{atn}\frac{1}{G\omega L}$

rezystancja wzorcowa R_f. Dwójnik szeregowy R+L lub R+C załącza się w pętli sprzężenia zwrotnego wzmacniacza operacyjnego, a na wejściu wzmacniacza umieszcza się rezystancję wzorcową R_i. Spośród szesnastu możliwych konfiguracji pomiarowych dwójników biernych [19], ze źródłem pobudzającym układ pomiarowy załączonym do wejścia inwersyjnego lub nieinwersyjnego wzmacniacza operacyjnego, tylko cztery podane w tabeli 1, zapewniają proporcjonalność sygnału wyjściowego z detektorów fazoczułych do mierzonych składowych immitancji [18,19].

Oprócz metod opisanych w pracy istnieje również sposób polegający na mierzeniu napięcia i prądu na elemencie badanym [1], co umożliwia pomiar parametrów dwójników znajdujących się w dowolnym miejscu sieci, oddzielonym od źródeł sygnału wieloma innymi elementami czynnymi i biernymi. Sygnał napięciowy jest wówczas sygnałem odniesienia. Możliwe jest również zastąpienie rezystancji wzorcowych R_f lub R_i przez pojemność wzorcową [22]. Chociaż prezentowane w tym artykule zależności odnoszą się do układów pomiarowych i błędów przetwarzania detektora jednopołówkowego to jednak mogą być one rozszerzone z niewielkimi modyfikacjami na przypadek detektorów dwupołówkowych.

W rozważaniach założono, że operacyjny wzmacniacz pomiarowy posiada nieskończone wzmocnienie napięciowe i nieograniczone pasmo przenoszonych częstotliwości. Przyjęto, że w torze pomiarowym obejmującym generator, wzmacniacz pomiarowy, klucze analogowe i integratory (rys. 1), tylko prądy i napięcia niezrównoważenia wzmacniacza pomiarowego i integratora wnoszą wkład do błędu pomiarowego, a pozostałe bloki sterujące nie mają wpływu na błąd. Przy pobudzeniu każdego z rozważanych układów sygnałem harmonicznym o postaci:

$$e(t) = E\sin(\omega t) \tag{1}$$

przy czym: E — amplituda, ω — pulsacja generatora zasilającego, na wyjściu wzmacniacza operacyjnego otrzymuje się napięcia zestawione w tabeli 1. $|Y|$, $|Z|$ jest modułem admitancji i impedancji dwójnika, φ — kątem przesunięcia fazowego

Tabela 2

Przedziały całkowania w każdym okresie sygnału odniesienia dla trzech metod detekcji fazoczułej

metoda	dwójnik	składowa immitancji	
		rzeczywista	urojona
stały przedział całkowania	wszystkie	$0;\ \pi$	$-\frac{\pi}{2};\frac{\pi}{2}$
zmienny przedział całkowania	G‖C, R+L	$0;\ \varphi_m = \frac{\pi}{2} - \varphi$	$\varphi_m = \frac{\pi}{2} - \varphi;\frac{\pi}{2}$
	R+C, G‖L	$\varphi_m = \frac{\pi}{2} + \varphi;\ \pi$	$\frac{\pi}{2};\ \varphi_m = \frac{\pi}{2} + \varphi$
próbkowanie (sample and-hold)	wszystkie	$\delta\left(\frac{\pi}{2}\right)$	$\delta(0)$

sygnału odpowiedzi względem sygnału odniesienia. Sygnały te są całkowane w odpowiednich przedziałach właściwych danej metodzie detekcji fazoczułej i podane zostały w tabeli 2.

Przedziały całkowania podane w tabeli 2 odnoszą się do detekcji jednopołówkowej. W przypadku detekcji dwupołówkowej całkowanie odbywa się w pierwszym półokresie sygnału odniesienia w takich samych przedziałach jak w tabeli 2, a w drugim półokresie sygnał odpowiedzi jest odwracany w fazie i dodatkowo całkowany w przedziałach przesuniętych o π względem tych podanych w tabeli 2.

Tabela 3

Napięcia wyjściowe detektorów fazoczułych mierzących składowe immitancji dwójników

metoda	składowa rzeczywista		składowa urojona	
	G‖C, R+L	R+C, G‖L	G‖C, R+L	R+C, G‖L
stały przedział całkowania	$-2EM\cos(\varphi)$	$-2EM\cos(\varphi)$	$-2EM\sin(\varphi)$	$+2EM\sin(\varphi)$
zmienny przedział całkowania	$-EM\cos(\varphi)$	$-EM\cos(\varphi)$	$-EM\sin(\varphi)$	$-EM\sin(\varphi)$
próbkowanie (sample-and-hold)	$-EM\cos(\varphi)$	$-EM\cos(\varphi)$	$-EM\sin(\varphi)$	$+EM\sin(\varphi)$

Wyniki całkowania sygnałów z tabeli 1 w przedziałach z tabeli 2 zestawiono w tabeli 3. W wyrażeniach tabeli 3, dla uproszczenia rozważań założono, że znormalizowany względem $\omega=2\pi/T$ moduł funkcji przejścia układu całkującego równy jest $T/2\pi\tau$. Moduł wzmocnienia dla prądu zmiennego w układzie z zamkniętą pętlą sprzężenia zwrotnego M, oraz wzmocnienie stałoprądowe K dla dwójników równoległych G‖C i G‖L zdefiniowano wyrażeniami:

$$\left.\begin{aligned} M &= R_f|Y| \\ K &= R_f G \end{aligned}\right\} \tag{2}$$

a dla dwójników szeregowych R+C i R+L wyrażeniami:

$$\left.\begin{aligned} M &= \frac{|Z|}{R_i} \\ K &= \frac{R}{R_i} \end{aligned}\right\} \tag{3}$$

Dokładniejsze omówienie zasady pomiaru składowych immitancji dwójników biernych za pomocą detektorów fazoczułych można znaleźć w pracach [18 i 19].

3. BŁĘDY SPOWODOWANE NAPIĘCIAMI I PRĄDAMI NIEZRÓWNOWAŻENIA WZMACNIACZY OPERACYJNYCH W TORZE POMIAROWYM

Wejściowe prądy i napięcia niezrównoważenia wzmacniaczy operacyjnych stosowanych w torze pomiarowym dodają się do wyniku pomiaru stając się przyczyną błędu. W ogólnym przypadku, przyrost napięcia na wyjściu integratorów w przedziale całkowania od kąta fazowego φ_d (początek przedziału całkowania) do kąta fazowego φ_g (koniec przedziału całkowania — oba kąty zależne od rozpatrywanej metody detekcji i mierzonej składowej immitancji, por. tabela 2), spowodowany wejściowymi napięciami i prądami niezrównoważenia wzmacniaczy operacyjnych uczestniczących w przetwarzaniu sygnału pomiarowego można zapisać w postaci:

$$\Delta U_s = \frac{T}{2\pi\tau}\int_{\varphi_d}^{\varphi_g}(R_f I_{sp} + (1+K)\,U_{sp})\,\mathrm{d}\omega t + \frac{T}{2\pi\tau}\int_{0}^{2\pi}(R_c I_{sc} + U_{sc})\,\mathrm{d}\omega t + U_{sc} \qquad (4)$$

przy czym: $\tau = R_c C_c$ — stała czasowa układu całkującego, U_{sp}, I_{sp} — wejściowe napięcie i prąd niezrównoważenia wzmacniacza pomiarowego, U_{sc}, I_{sc} — wejściowe napięcie i prąd niezrównoważenia wzmacniacza tworzącego integrator, $T = 2\pi/\omega$ — okres sygnału pobudzającego badany dwójnik. Rezystancja R_f jest rezystancją wzorcową w pętli sprzężenia zwrotnego wzmacniacza operacyjnego, a rezystancja R_i jest rezystancją wzorcową dołączoną do wejścia wzmacniacza pomiarowego, w zależności od konfiguracji układu pomiarowego.

Pierwsza całka w (4) jest składnikiem błędu pochodzącym od napięcia i prądu niezrównoważenia wzmacniacza pomiarowego i obliczana jest w przedziale $[\varphi_d, \varphi_g]$, czyli w przedziale kąta fazowego, w którym odbywa się pomiar składowej immitancji za pomocą danej metody detekcji fazoczułej (tabela 2). Składnik $R_f I_{sp}$ w pierwszej całce jest napięciem na wyjściu wzmacniacza pomiarowego wywołanym przez wejściowy prąd niezrównoważenia I_{sp}. Drugi składnik jest proporcjonalny do wejściowego napięcia niezrównoważenia tego wzmacniacza [23].

Druga całka obliczana jest w przedziale granicznym $[0, 2\pi]$ ponieważ prądy i napięcia niezrównoważenia wzmacniacza tworzącego integrator całkowane są w całym okresie pomiarowym (zakłada się, że integrator zostaje wyzerowany po przesłaniu wyniku całkowania do pamięci i oczekuje na następny pomiar, który niezależnie od metody kończy się po całym okresie). Pierwszy składnik w drugiej całce $R_c I_{sc}$ jest napięciem na wyjściu integratora spowodowanym przez wejściowy prąd niezrównoważenia wzmacniacza. Do obu całek dodawane jest napięcie niezrównoważenia U_{sc} wzmacniacza operacyjnego tworzącego integrator, występujące po zakończeniu całkowania [23]. Założono przy tym, że wejściowe prądy spoczynkowe wzmacniaczy operacyjnych są skompensowane.

Względny błąd pomiaru składowej immitancji jest stosunkiem przyrostu napięcia ΔU_s (4) do wartości napięcia U_c uzyskanego dla mierzonej składowej immitancji na wyjściu integratora w przypadku, gdy napięcia i prądy niezrównoważenia równe są zero:

$$\delta_s = \left|\frac{\Delta U_s}{U_c}\right|. \tag{5}$$

Zakładając stałość napięć i prądów niezrównoważenia w całym okresie integracji $[0, 2\pi]$ oraz liniowość układu całkującego, wyrażenie (4) można sprowadzić do postaci:

$$\Delta U_s = \frac{T(\varphi_g - \varphi_d)}{2\pi\tau}(R_f I_{sp} + (1 + K) U_{sp}) + \frac{T}{\tau}(R_c I_{sc} + U_{sc}) + U_{sc}. \tag{6}$$

Wartość liczbową błędu pomiaru spowodowanego napięciami i prądami niezrównoważenia można wyznaczyć tylko dla konkretnego przypadku ze względu na jego zależność od mierzonej immitancji oraz konieczność znajomości wartości napięć i prądów niezrównoważenia dla zastosowanego typu wzmacniacza operacyjnego. Dlatego poniżej podane zostaną tylko wyrażenia ogólne opisujące błąd względny δ_s dla trzech układów detektorów fazoczułych i obu mierzonych składowych immitancji, a następnie porównane zostaną wrażliwości tych układów na prądy i napięcia niezrównoważenia.

3.1. DETEKTOR FAZOCZUŁY Z CAŁKOWANIEM W ZMIENNYM PRZEDZIALE

Przy pomiarze składowej rzeczywistej i urojonej immitancji dwójników biernych G||C, R+L i G||L błędy pomiaru spowodowane wejściowymi napięciami i prądami niezrównoważenia obliczone z (5) wynoszą dla składowej rzeczywistej:

$$\delta_{sRe} = \frac{\Delta\varphi(R_f I_{sp} + (1 + K) U_{sp}) + 2\pi(R_c I_{sc} + (1 + \tau/T) U_{sc})}{EM\cos(\varphi)} \tag{7}$$

i dla składowej urojonej:

$$\delta_{sIm} = \frac{\Delta\varphi(R_f I_{sp} + (1 + K) U_{sp}) + 2\pi(R_c I_{sc} + (1 + \tau/T) U_{sc})}{EM\sin(\varphi)} \tag{8}$$

Błędy pomiaru (7) i (8) są proporcjonalne do długości przedziału całkowania $\Delta\varphi$ i zależą od kąta przesunięcia fazowego φ.

W przypadku dwójnika szeregowego R+C (tabela 1) należy uwzględnić to, że tworzy on ze wzmacniaczem operacyjnym układ całkujący wejściowe napięcia i prądy niezrównoważenia i w skończonym czasie musi osiągnąć stan nasycenia. Jest to poważna wada utrudniająca zastosowanie tej konfiguracji do pracy ciągłej w układach z przetwornikami pomiarowymi. Mierzony dwójnik R+C musi być okresowo zwierany w celu eliminacji ładunku zgromadzonego na pojemności C. W takim przypadku przyrost napięcia na wyjściu całego układu pomiarowego wyraża się zależnością:

$$\Delta U_s = \frac{T}{2\pi\tau}\int_{\varphi_d}^{\varphi_g}\left(\frac{T}{2\pi C(R+R_i)}\int_0^{2\pi}(R_i I_{sp} + U_{sp})\,d\omega t\right)d\omega t + \frac{T}{2\pi\tau}\int_0^{2\pi}(R_c I_{sc} + U_{sc})\,d\omega t + U_{sc}, \tag{9}$$

przy czym całka wewnętrzna w pierwszym składniku opisuje napięcie pojawiające się w każdym okresie T na wyjściu wzmacniacza pomiarowego na skutek prądu i napięcia niezrównoważenia wzmacniacza.

Z zależności (9) wynika błąd względny pomiaru składowej rzeczywistej dwójnika R+C:

$$\delta_{sRe} = \frac{\frac{T\Delta\varphi}{C(R+R_i)}(R_i I_{sp} + U_{sp}) + 2\pi(R_c I_{sc} + (1 + \tau/T) U_{sc})}{EM\cos(\varphi)} \tag{10}$$

i błąd pomiaru składowej urojonej:

$$\delta_{sIm} = \frac{\frac{T\Delta\varphi}{C(R+R_i)}(R_i I_{sp} + U_{sp}) + 2\pi(R_c I_{sc} + (1 + \tau/T) U_{sc})}{EM\sin(\varphi)}. \tag{11}$$

3.2. DETEKTOR FAZOCZUŁY Z CAŁKOWANIEM W STAŁYM PRZEDZIALE

W tej metodzie detekcji fazoczułej, w której przedziały całkowania są stałe, niezależnie od kąta przesunięcia fazowego φ, względne błędy pomiaru składowej rzeczywistej dla poszczególnych dwójników wynoszą dla dwójników równoległych G‖C i G‖L oraz szeregowego R+L:

$$\delta_{sRe} = \frac{\pi(R_f I_{sp} + (1 + K) U_{sp}) + 2\pi(R_c I_{sc} + (1 + \tau/T) U_{sc})}{2EM\cos(\varphi)}, \tag{12}$$

a dla dwójnika szeregowego R+C:

$$\delta_{sRe} = \frac{\frac{T\pi}{C(R+R_i)}(R_i I_{sp} + U_{sp}) + 2\pi(R_c I_{sc} + (1 + \tau/T) U_{sc})}{2EM\cos(\varphi)}. \tag{13}$$

Błąd względny pomiaru składowej urojonej dla dwójnika równoległego G‖C, G‖L i szeregowego R+L wynosi:

$$\delta_{sIm} = \frac{\pi(R_f I_{sp} + (1 + K) U_{sp}) + 2\pi(R_c I_{sc} + (1 + \tau/T) U_{sc})}{2EM\sin(\varphi)} \tag{14}$$

i dla dwónika szeregowego R+C:

$$\delta_{sIm} = \frac{\frac{-T\pi}{C(R+R_i)}(R_i I_{sp} + U_{sp}) + 2\pi(R_c I_{sc} + (1 + \tau/T) U_{sc})}{2EM\sin(\varphi)}, \tag{15}$$

przy czym w odniesieniu do dwójnika R+C stosują się zastrzeżenia podane w rozdz. 3.1.

We wszystkich układach z całkowaniem błąd spowodowany napięciami i prądami niezrównoważenia jest proporcjonalny do długości przedziału całkowania. Z tego powodu błąd pomiaru za pomocą detektora z całkowaniem w zmiennym przedziale może być mniejszy, gdyż przedział całkowania jest krótszy.

3.3. DETEKTOR FAZOCZUŁY Z PRÓBKOWANIEM

W układzie detektora fazoczułego z próbkowaniem napięcie i prąd niezrównoważenia wzmacniacza pomiarowego nie są całkowane lecz dodają się do wartości chwilowej sygnału odpowiedzi w momencie próbkowania. Błąd pomiaru jest więc niezależny od wartości kąta fazowego φ. Dlatego ta metoda detekcji fazoczułej jest najmniej wrażliwa na niezrównoważenie wzmacniacza pomiarowego. Czułość na niezrównoważenie wzmacniacza spełniającego funkcję integratora pozostaje bez zmian, gdyż prąd wejściowy wzmacniacza operacyjnego może ładować kondensator stanowiący element pamiętający, w czasie między kolejnymi próbkowaniami, czyli w okresie od 0 do 2π.

Błędy względne pomiaru składowej rzeczywistej dla wszystkich konfiguracji dwójników wyrażają się następującą zależnością:

$$\delta_{sRe} = \frac{R_f I_{sp} + (1 + K) U_{sp} + \dfrac{TI_{sc}}{C_p} + U_{sc}}{EM\cos(\varphi)}. \tag{16}$$

Błędy względne pomiaru składowej urojonej immitancji wynoszą:

$$\delta_{sIm} = \frac{R_f I_{sp} + (1 + K) U_{sp} + \dfrac{TI_{sc}}{C_p} + U_{sc}}{EM\sin(\varphi)} \tag{17}$$

W odniesieniu do dwójnika R+C obowiązują uwagi podane w rozdz. 3.1.

4. PORÓWNANIE BŁĘDÓW METOD

Wrażliwość wszystkich trzech metod pomiarowych na wejściowe napięcia i prądy niezrównoważenia wymaga szacunkowego porównania. Rozważono przypadki najmniej korzystne, gdy następuje sumowanie się napięć niezrównoważenia.

Błędy pomiaru składowej rzeczywistej i urojonej immitancji w układzie pomiarowym dwójnika G||C, R+L i G||L za pomocą detekcji fazoczułej z całkowaniem w zmiennym przedziale (oznaczanej w poniższych wzorach przez IV) odniesione do błędów pomiaru za pomocą detekcji fazoczułej z całkowaniem w stałym przedziale (oznaczanej przez *IC*) są jednakowe dla obu składowych i zawierają się w granicach:

$$1 - \frac{1}{2 + \dfrac{R_f I_{sp} + (1+K) U_{sp}}{2(R_c I_{sc} + (1+\tau/T) U_{sc})}} \leqslant \frac{\delta_{sRe}(IC)}{\delta_{sRe}(\mathrm{IV})} \leqslant$$

$$\leqslant \frac{1}{2}\left[1 + \frac{R_f I_{sp} + (1+K) U_{sp}}{2(R_c I_{sc} + (1+\tau/T) U_{sc})}\right] \tag{18}$$

przy zmianach długości przedziału całkowania $\Delta\varphi$ w metodzie detekcji fazoczułej z całkowaniem w zmiennym przedziale od 0 do $\pi/2$ lewa strona nierówności ($\Delta\varphi = \pi/2$) jest zawsze mniejsza od 1 lecz większa od 1/2. Prawa strona nierówności (dla $\Delta\varphi = 0$) jest zawsze większa od 1/2 i może być większa od 1. Stosunek błędów spowodowanych przez napięcia i prądy niezrównoważenia jest większy od 1 dla przedziału całkowania $\Delta\varphi$ w metodzie detekcji fazoczułej z całkowaniem w zmiennym przedziale spełniającego warunek:

$$0 < \Delta\varphi < \frac{\pi}{2}\left[1 - \frac{2(R_c I_{sc} + (1+\tau/T) U_{sc})}{R_f I_{sp} + (1+K) U_{sp}}\right]. \tag{19}$$

W układzie pomiarowym dwójnika szeregowego R+C, stosunek błędów zawiera się w granicach:

$$1 - \frac{1}{2 + \dfrac{\dfrac{T}{C(R+R_i)}(R_i I_{sp} + U_{sp})}{2(R_c I_{sc} + (1+\tau/T) U_{sc})}} \leqslant \frac{\delta_{sRe}(IC)}{\delta_{sRe}(\mathrm{IV})} \leqslant$$

$$\leqslant \frac{1}{2}\left[1 + \frac{\dfrac{T}{C(R+R_i)}(R_i I_{sp} + U_{sp})}{2(R_c I_{sc} + (1+\tau/T) U_{sc})}\right]. \tag{20}$$

Z zależności (18)–(20) wynika, że pomiar metodą detekcji fazoczułej z całkowaniem w zmiennym przedziale jest bardziej wrażliwy na napięcia i prądy niezrównoważenia od pomiaru za pomocą metody detekcji fazoczułej z całkowaniem w stałym przedziale. Wrażliwość ta jest funkcją kąta fazowego $\Delta\varphi$ przedziału całkowania, czyli funkcją stosunku składowej rzeczywistej i urojonej immitancji badanego dwójnika.

Porównanie obu metod całkowych z metodą detekcji fazoczułej z próbkowaniem nie prowadzi do jednoznacznych wyników, ze względu na dużą ilość parametrów, których nie można zredukować. Stosunek względnych błędów pomiaru składowej rzeczywistej za pomocą metody detekcji fazoczułej z całkowaniem w zmiennym przedziale i detekcji fazoczułej z próbkowaniem (oznaczanej przez *SH*) wynosi:

$$\frac{\delta_{sRe}(SH)}{\delta_{sRe}(\mathrm{IV})} = \frac{R_f I_{sp} + (1+K) U_{sp} + \dfrac{TI_{sc}}{C_p} + U_{sc}}{\Delta\varphi (R_f I_{sp} + (1+K) U_{sp}) + 2\pi (R_c I_{sc} + (1+\tau/T) U_{sc})}. \tag{21}$$

Pomiar za pomocą detektora fazoczułego z próbkowaniem jest obarczony zawsze mniejszym błędem niż pomiar detektorem fazoczułym z całkowaniem w zmiennym przedziale przy tych samych napięciach i prądach niezrównoważonego wzmacniacza pomiarowego i integratora jeśli spełniona jest nierówność:

$$\frac{2\pi(R_c I_{sc} + (1 + \tau/T) U_{sc})}{R_f I_{sp} + (1 + K) U_{sp} + \frac{TI_{sc}}{C_p} + U_{sc}} > 1. \tag{22}$$

W celu porównania błędów pomiarowych w metodach detekcji fazoczułej z całkowaniem w stałym przedziale i z próbkowaniem należy przeprowadzić analogiczne rozumowanie. Stosunek błędów w przypadku pomiaru składowej rzeczywistej i urojonej jest jednakowy i wynosi:

$$\frac{\delta_{sRe}(IC)}{\delta_{sRe}(SH)} = \frac{\pi(R_f I_{sp} + (1+K) U_{sp}) + 2\pi(R_c I_{sc} + (1+\tau/T) U_{sc})}{2\left(R_f I_{sp} + (1+K) U_{sp} + \frac{TI_{sc}}{C_p} + U_{sc}\right)}. \tag{23}$$

Przeprowadzone rozważania odnoszą się do przypadku, gdy wyjście wzmacniacza pomiarowego jest połączone galwanicznie z wejściem integratora. Jeśli możliwe jest odseparowanie tych dwóch układów za pomocą pojemności (która jednak może wnosić dodatkowe przesunięcia fazowe) lub jeśli wzmacniacz pomiarowy jest idealny, to można pominąć błąd wynikający z niezrównoważenia wzmacniacza pomiarowego. Wyrażenia (4) i (6) określające wartość bezwzględną błędu pomiaru w metodach fazoczułych z całkowaniem upraszczają się do postaci:

$$\Delta U_s = \frac{T}{2\pi\tau}\int_0^{2\pi}(R_c I_{sc} + U_{sc})\,d\omega t + U_{sc} = \frac{T}{\tau}(R_c I_{sc} + U_{sc}) + U_{sc}, \tag{24}$$

a dla metody detekcji fazoczułej z próbkowaniem mają postać:

$$\Delta U_s = \frac{TI_{sc}}{C_p} + U_{sc}. \tag{25}$$

Wyrażenia na wartość błędu pomiarowego mają więc identyczną postać dla obu metod detekcji fazoczułej z całkowaniem. Sytuacja zmienia się jednak, gdy integrator jest zerowany w czasie między kolejnymi przedziałami całkowania. Wówczas zmieniają się granice całkowania w drugiej całce w (4) lub (24) i błąd bezwzględny wynosi:

$$\Delta U_s = \frac{T}{2\pi\tau}\int_{\varphi_d}^{\varphi_g}(R_c I_{sc} + U_{sc})\,d\omega t + U_{sc} = \frac{T(\varphi_g - \varphi_d)}{2\pi\tau}(R_c I_{sc} + U_{sc}) + U_{sc}. \tag{26}$$

Ze względu na występujący w (26) składnik proporcjonalny do przedziału całkowania $(\varphi_g - \varphi_d)$, wrażliwość detekcji fazoczułej z całkowaniem w zmiennym przedziale na napięcie i prąd niezrównoważenia wzmacniacza tworzącego integrator

jest w tym przypadku mniejsza od detekcji fazoczułej z całkowaniem w stałym przedziale, co wynika ze stosunków błędów:

$$\left|\frac{\delta_{sRe}(\mathrm{IV})}{\delta_{sRe}(IC)}\right| = 2\frac{\Delta\varphi}{\pi} \leqslant 1. \tag{27}$$

W tym przypadku dla $0 \leqslant \Delta\varphi \leqslant \pi/2$ detekcja fazoczuła z całkowaniem w zmiennym przedziale jest mniej wrażliwa na napięcia i prądy niezrównoważenia integratora od metody detekcji fazoczułej z całkowaniem w stałym przedziale. Wynika to z krótszego przedziału całkowania w przypadku pierwszej z metod.

PODSUMOWANIE

Dokonano teoretycznego oszacowania niedokładności pomiaru składowych immitancji dwójników biernych za pomocą detektorów fazoczułych, spowodowanej przez wejściowe napięcia i prądy niezrównoważenia wzmacniaczy operacyjnych tworzących poszczególne stopnie toru pomiarowego. Porównano pod tym względem trzy typy detektorów fazoczułych.

Z podanych zależności wynika, że wejściowe napięcia i prądy niezrównoważenia w detektorze fazoczułym z całkowaniem w stałym przedziale wnoszą mniejszy błąd do wyniku pomiaru, niż w detektorze fazoczułym z całkowaniem w zmiennym przedziale. Wrażliwość detektora fazoczułego z całkowaniem w zmiennym przedziale może być w pewnych warunkach mniejsza niż detektora fazoczułego z całkowaniem w stałym przedziale. Detektor fazoczuły z próbkowaniem jest najmniej wrażliwy na wejściowe prądy i napięcia niezrównoważenia.

BIBLIOGRAFIA

1. J. H o j a: *Metoda pomiarowa parametrów elementów RLC wykorzystująca sygnały proporcjonalne do prądu i napięcia na elemencie mierzonym.* Rozprawy Elektrotechniczne, 1984, 30, Nr 3, 643–58
2. J. H o j a: *Optymalizacja parametrów detektora fazoczułego przy niekorzystnym stosunku składowych ortogonalnych sygnału pomiarowego.* XVII Międzyuczelniana Konferencja Metrologów, Poznań, Wrzesień 1984
3. A. J a w o r e k: *Wyznaczanie składowych immitancji parametrycznych przetworników pomiarowych za pomocą detekcji fazoczułej.* VII Krajowa Konferencja Metrologów, Wrocław 1986
4. J. E. S i g d e l l: *A principle for capacitance measurement suitable for linear evaluation of capacitance transducers.* IEEE Trans. Instrum. Meas. 1972, 21, N 1, 60–4
5. R. V. J o n e s, J. C. S. R i c h a r d s: *The design and some applications of sensitive capacitance micrometers.* J. Phys. E: Sci. Instrum. 1973, 6, No 7, 589–600
6. H. F i s h e r: *Lock-in amplifiers obtain measurements even if there is more noise than signal.* Laser Focus. Nov. 1977, 82–8
7. A. J a w o r e k: *Pomiary składowych impedancji dwójników.* PAK, 1987, 33, Nr 2, 38–9
8. A. J a w o r e k: *Sposób i układ do pomiaru składowych impedancji lub admitancji dwójników.* Patent P–149187

9. L. Q. Orsini: *Conversion of immitance parameters to dc voltage.* IEEE Trans. Instrum. Meas. 1973, 21, N 3, 196—8
10. A. Jaworek: *Sposób szybkiego pomiaru składowych immitancji dwójników metodą próbkowania odpowiedzi.* PAK, 1987, 33, Nr 5, 102—3
11. A. Jaworek: *Sposób i układ do pomiaru składowych rzeczywistej i urojonej immitancji dwójników.* Patent P—149188
12. G. B. Clayton: *Ecperiments with operational amplifiers. Pt. 11: An op-amp used as a phase sensitive detector.* Wireless World, July 1973, 355—6
13. C. T. Morse: *A computer controlled apparatus for measuring ac properties of materials over the frequency range 10^{-5} to 10^{5} Hz.* J. Phys. E: Sci. Instrum 1974, 7, 657—62
14. J. Hoja: *Bezpośrednia metoda pomiaru elementów składowych dwójników wieloelementowych o znanej strukturze.* Praca doktorska. Politechnika Gdańska, Gdańsk 1979
15. K. Maeda: *An Automatic Precision 1 MHz Digital LRC Meter.* Hewlett-Packard J., March 1974, 2—9
16. M. von Ortenberg, W. Staguhn, F. Böbel, S. Takeyama, T. Sakakibara, N. Miura: *Phase-sensitive detection of magnetoresistivity by the four-probe technique in pulsed high magnetic fields.* J. Phys. E: Sci. Instrum. 1989, 22, Nr 6, 359—62
17. S. M. Huang, A. L. Stott, R. G. Green, M S. Beck: *Electronic transducers for industrial measurement of low value capacitances.* J. Phys. E: Sci. Instrum. 1988, 21, Nr 3, 242—50
18. A. Jaworek: *Pomiar immitancji dwójników biernych za pomocą detektorów fazoczułych.* Pomiary, Automatyka, Kontrola 1990, 36, Nr 9, 180—3
19. A. Jaworek: *Detektory fazoczułe w pomiarach składowych immitancji dwójników.* Kwart. Elektron. Telekom. 1990, 36, z. 3—4, s. 189—219
20. A. Jaworek: *Dynamiczne właściwości fazoczułych metod detekcji składowych immitancji dwójników biernych.* Pomiary, Automatyka, Kontrola 1990, 36, Nr 2—3, 42—44
21. A. Jaworek: *Wpływ parametrów wzmacniacza pomiarowego na pomiar składowych immitancji dwójników za pomocą detektorów fazoczułych.* Kwart. Elektron. Telekom. 1990, 36, zesz. 3—4
22. J. C. R. Richard s: *Some aspects of transducer immittance measurement.* J. Phys. E: Sci. Instrum. 1982, 15, 1251—56
23. M. Nadachowski, Z. Kulka: *Analogowe układy scalone.* Warszawa: WKŁ, 1980

A. JAWOREK

INPUT OFFSET CURRENT AND VOLTAGE EFFECT ON MEASURING ERROR OF PHASE-SENSITIVE DETECTORS

Summary

The measuring error caused by offset currents and voltages of operational amplifiers transforming the signal in phase-sensitive detectors measuring the real and imaginary parts of a complex immitance of a linear and passive two-terminal has been evaluated. Three types of phase-sensitive detectors have been compared. From the relations derived it follows that the phase-sensitive detector of sample-and-hold type is the least sensitive to the offset currents and voltages. The measuring error in phase-sensitive detector with constant interval of integration is equal to or less than in phase-sensitive detector with adaptive interval of integration.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1992, 38, z. 2–3, ss. 419–429

(621.384.3)

Modulacja fotoemisji

JÓZEF WOJAS

Profesor emerytowany Politechniki Warszawskiej

Otrzymano 1991. 06. 20

Autoryzowano do druku 1991. 12.

W artykule przedstawiono zagadnienie pomiaru przesuwania się prądowo-napięciowych charakterystyk fotoemisji zewnętrzenej pod wpływem dotatkowego podświetlania podczerwienią próbek półprzewodników o dużej przerwie energetycznej.

1. TEORETYCZNE WPROWADZENIE

W półprzewodnikach o dużej przerwie energetycznej (np. w arsenku galu) występuje następujące zjawisko: Kierując na próbkę, oprócz skoncentrowanej wiązki promieniowania ultrafioletowego wzbudzającego fotoemisję, wiązkę promieniowania podczerwonego (o energii w przedziale 1,1 – 1,4 eV) obserwuje się przesuwanie charakterystyk prądowo-napięciowych.

Próbki typu *n* (z dużą koncentracją elektronów) przesuwają swoje charakterystyki prądowo-napięciowe w prawo i wykazują około 10% wzrost prądu nasycenia.

Próbki typu *p* dają niewielkie przesunięcie charakterystyk w lewo; ich prąd nasycenia prawie nie wzrasta.

Kwanty promieniowania podczerwonego (o wyżej wspomnianej energii) wybijają elektrony ze stanów powierzchniowych, a prawie nie mogą przerzucać elektronów z pasma walencyjnego do pasma przewodnictwa, gdyż ich energia jest mniejsza od szerokości przerwy. W przerwie energetycznej nie ma swobodnych elektronów; są one związane z donorami lub stanami powierzchniowymi. Podświetlanie uwalnia je i wtedy mogą przechodzić do pasma walencyjnego z którego są wyemitowywane pod wpływem ultrafioletu. Przy podświetlaniu próbek typu *n* występują dwa efekty:

1) Dopływ „uwolnionych" z powierzchni elektronów do warstwy przypowierzchniowej, co powoduje zmniejszenie zagięcia pasm, czyli obniżenie bariery potencjału (na skutek pojawienia się dodatniego ładunku na samej powierzchni), albo inaczej mówiąc inwersji z typu *n* na *p* .

2) Niewielki wzrost prądu nasycenia (10 do 20%). Bierze się to stąd, że jeśli zagięcie pasm na typ *p* maleje przy powierzchni, to w górnej części pasma walencyjnego

pojawia się więcej elektronów, które tu „*pospadały*", a mniej dziur i emisja jest nieco większa, bo, jak wiemy, główna fotoemisja pochodzi z pasma walencyjnego. Inaczej mówiąc, ładunek dodatni na samej powierzchni polepsza warunki wychodzenia fotoelektronów, bo obniża barierę powierzchniową; powstała warstwa podwójna wytwarza pole nadające elektronom dużą energię skierowaną na zewnątrz ułatwiając przez to ich wychodzenie.

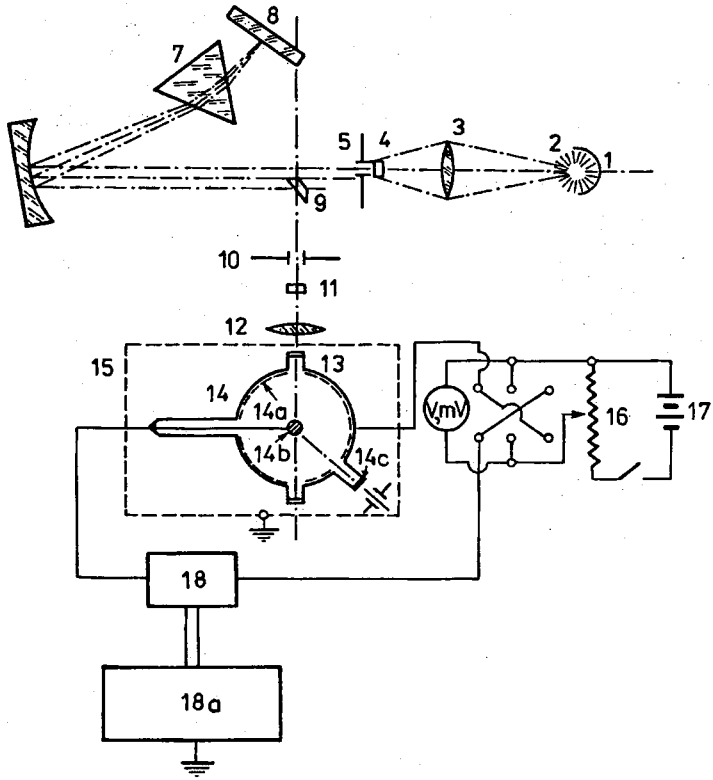
Podświetlane próbki typu *p* (konkretnie arsenku galu) wykazują niewielkie zagięcie do dołu (od 0,25 do 0,3 eV). A więc powierzchnia posiada zarówno ładunek dziurowy jak elektronowy. Elektrony bowiem zawsze tkwią w stanach powierzchniowych. Ale na omawianej powierzchni przeważa ładunek dziurowy, gdyż powierzchnia ta już wcześniej oddała swoje elektrony do dolnych poziomów warstwy przypowierzchniowej przez co spowodowała niewielkie zagięcie do dołu. Przy dodatkowym podświetleniu tej powierzchni podczerwienią, trochę elektronów z pułapek powierzchni uwalnia się, ale niewiele, gdyż przeważają tu dziury. Natomiast uwalniają się elektrony w warstwie pzypowierzchniowej do której podczerwień łatwo wnika. Te elektrony zmieniają strukturę warstwy przypowierzchniowej powodując zagięcie pasm do dołu. Część tych elektronów rekombinuje, a część wychodzi na powierzchnię i daje tam niewielki ładunek ujemny.

Z powodu „ucieczki" pewnej ilości elektronów z warstwy przypowierzchniowej, przy samej powierzchni zagięcie na typ *n* zmniejsza się a naładowana ujemnie sama powierzchnia podwyższa trochę barierę potencjału do góry; praca wyjścia przy tym trochę rośnie. Dlatego omawiane charakterystyki przesuwają się nieco w lewo. Innymi słowy, zagięcie w wyniku inwersji na typ *n* w próbce typu *p* zmniejsza się, gdy elektrony z warstwy przypowierzchniowej uciekają na samą powierzchnię — bo w warstwie przypowierzchniowej zwiększyła się ilość dziur.

Jeśli teraz podczerwień w warstwie przypowierzchniowej będzie generowała pary: elektron — dziura, to elektrony mogą być schwytane przez zjonizowane donory na powierzchni, a dziury (dodatkowe) zostaną w warstwie przypowierzchniowej. To ostatnie powoduje spłaszczenie zagięcia przy powierzchni, czyli „powrót" do typu *p* (przy powierzchni). Należy pamiętać, że efekty te spowodowane podczerwienią są w próbkach typu *p* znacznie mniejsze niż w próbkach typu *n*.

2. ZJAWISKO MODULACJI FOTOEMISJI PRZEZ OŚWIETLANIE PODCZERWIENIĄ

Zmiany właściwości przypowierzchniowej warstwy półprzewodnika (głównie rodzaju i gęstości ładunku przypowierzchniowego) mają wpływ na prąd fotoemisji. Efektywne modulowanie tych właściwości może następować przez efekt Schottky'ego lub przez pobudzanie promieniowaniem o takiej energii która nie wywołuje fotoefektu zewnętrznego a powoduje przemieszczanie się nośników w przypowierzchniowych warstwach półprzewodników. Ten drugi efekt jest wyraźny w półprzewodnikach o dużej przerwie energetycznej i był badany przez Szubę [1] w monokryształach CdS, a znacznie później i w innym ujęciu przez Łagowskiego i współpracowników [2] także w CdS. Doświadczenie realizuje się przez jednoczesne pobudzanie próbki półprzewod-



Rys. 1. Układ do pomiaru charakterystyk prądowo-napięciowych.
Układ pomiarowy do zdejmowania charakterystyk fotoemisji. 1. reflektor z aluminium, 2. źródło ultrafioletu, 3. i 12. suprasilowe soczewki skupiające, 4. przesłona, 5. wejściowa szczelina monochromatora z dużą siłą światła np. 3MP–3, 6. aluminiowe zwierciadło paraboliczne, 7. pryzmat kwarcowy, 8. i 9. aluminiowe zwierciadła płaskie, 10. szczelina wyjściowa monochromatora, 11. przesłona, 13. okienko suprasilowe, 14. próżniowy kondensator kulisty, 14a. kolektor, 14b. emiter, 14c. wejście podczerwieni połączenie z pompą jonowo-sorpcyjną, 15. ekran od pól el-magnetycznych, 16. potencjometr-helipot jako dzielnik napięcia, 17. źródło napięcia stałego, 18. sonda elektrometru o wysokiej oporności (rzędu $10^{10}-10^{12})\Omega$, 18a. elektrometr wibracyjny, np. firmy „Ekco" o czułości 5.10^{-15}A/dz

nika promieniowaniem nadfioletowym oraz źródłem podczerwieni. Energie dodatkowego oświetlenia winny być nie mniejsze od szerokości przerwy energetycznej półprzewodnika (w CdS – 2,5 eV, w GaAs – 1,4 eV).

Autor zaobserwował po raz pierwszy to zjawisko w orientowanych kryształach GaAs typu „*n*" i „*p*". Wykonano pomiary prądów fotoemisji z dodatkowym oświetlaniem (w układzie podanym na rys. 1) w próbkach GaAs typu „*n*" o orientacji (100) przy wielu wartościach energii fotonów. Te same pomiary wykonano w próbkach (111) typu „*n*" przy kilku energiach pobudzającego promieniowania. Badano także przy kilku hν próbki typu „*p*" głównie o orientacji (111).

3. WYNIKI BADAŃ DOŚWIADCZALNYCH

3a) OMÓWIENIE WYNIKÓW POMIARÓW Z PRÓBKAMI „*n*"

Na rys. 2, 3a) i b) przedstawiono niektóre z otrzymanych przebiegów dla próbek „*n*" o orientacjach (100) i (111B). Część wyników autora dotyczących modulacji fotoemisji podczerwienią została opublikowana w 1969 roku [3].

Otrzymane krzywe sugerują następujące spostrzeżenia:

a) wyraźne przesuwanie się prądowo-napięciowych charakterystyk fotoemisji w prawo a stąd rozszerzanie niskoenergetycznych obszarów widm fotoelektronów,

b) wzrost napięcia kontaktowego kolektor-emiter,

c) niewielki i występujący nie na wszystkich charakterystykach wzrost prądu nasycenia przy pobudzaniu podczerwienią.

Ponadto sprawdzano czy występuje zmiana spektralnego rozkładu wydajności przy dodatkowym oświetlaniu i stwierdzono, że przebieg krzywej $Y = f(h\nu)$ oraz wartość fotoelektrycznego progu nie zmieniają się.

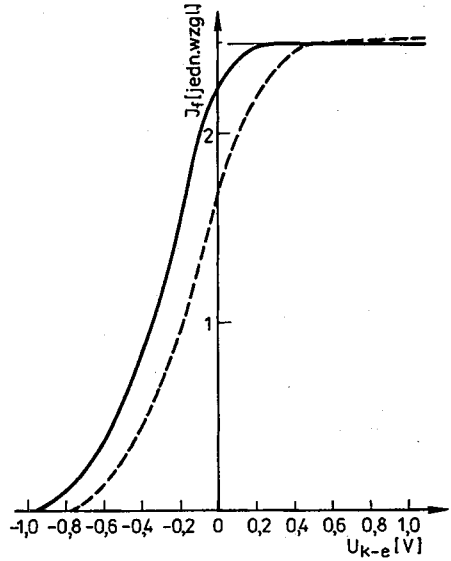
We wspomnianej pracy autora [3] dotyczącej badania wpływu „podświetlania" na fotoemisję z powierzchni (111B) monokryształów GaAs typu *n* i *p* zaproponowano wyjaśnienie obserwowanych efektów przez zjawisko zmiany zagięcia pasm przy powierzchni półprzewodnika.

Dinan, Galbraith i Fischer [4] badając wpływ podświetlania podczerwienią powierzchni (110) monokryształów GaAs zauważyli występowanie efektów a), b) i c) w próbkach typu *n*, natomiast nie stwierdzili zmian fotoemisyjnych właściwości próbek typu *p*. Potwierdzają oni wstępną interpretację Wojasa opartą na przypowierzchniowym modelu pasmowym półprzewodnika (na zjawisku „spłaszczenia" zagięcia pasm).

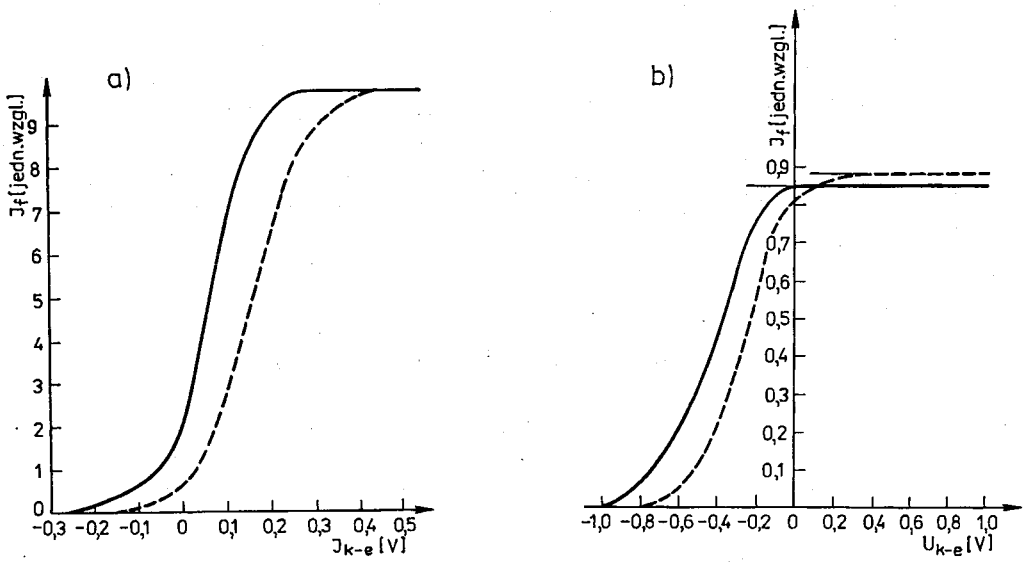
Rys. 4 przedstawia ten przypowierzchniowy model kryształu GaAs typu *n* w kierunku (100) i jego zmianę po oświetleniu podczerwienią.

Wyjaśnienie efektów a) i b) jest, zdaniem autora, następujące: Brak przesunięcia fotoelektrycznego progu $h\nu_t$ świadczy o tym, że pobudzanie podczerwienią nie wpływa na odległość energetyczną: poziom próżni — szczyt pasma walencyjnego. Położenie szczytu pasma walencyjnego w obszarze przypowierzchniowym względem E_p nie ulega zmianie, lecz zmienia się ono względem poziomu Fermiego. Wzrost kontaktowej różnicy potencjałów przy stałej pracy wyjścia z kolektora oznacza zbliżenie się poziomu Fermiego do poziomu próżni. Jest ono spowodowane wzrostem koncentracji elektronów w pasmie przewodnictwa przerzucanych tam przez promieniowanie podczerwone. Efekt ten jest równoważny zmniejszeniu zagięcia energetycznych pasm do góry (rys. 4). Inwersja pasm w obszarze przypowierzchniowym maleje na skutek pojawienia się ujemnego ładunku elektronów, co nazywamy spłaszczeniem zagięcia.

Obecnie autor potwierdził fakt, że „podnoszenie" się poziomu Fermiego przy dodatkowym oświetleniu jest większe dla wyższych energii pobudzającego nadfioletu (Tabela 1). Efekt ten nie jest łatwy do wyjaśnienia. Być może część uwalnianych w procesie fotoemisji elektronów jest w czasie ich wędrówki w krysztale pułapkowana w defektach, z których przechodzi do pasma przewodnictwa, zwiększając ujemny

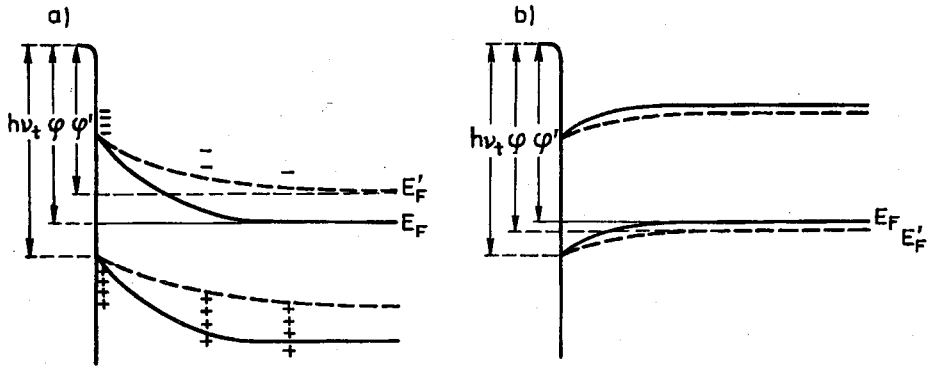


Rys. 2. Przesunięcia charakterystyk fotoemisji pod wpływem dodatkowego oświetlania podczerwienią próbki GaAs typu *n* o orientacji (100) (energia wzbudzania fotoemisji $h\nu = 4.89$ eV)



Rys. 3. Przesunięcia charakterystyk fotoemisji pod wpływem dodatkowego oświetlania dwu próbek GaAs typu n o orientacji (111): a) próbka nr. 1 (energia fotonów $h\nu = 4.89$ eV), b) próbka nr. 2 (energia fotonów $h\nu = 5.80$ eV). Krzywe przerywane oznaczają charakterystyki przy dodatkowym pobudzaniu podczerwienią

ładunek elektronów swobodnych. Elektronów tych byłoby tym więcej, im większe są energie pobudzającego nadfioletu.



Rys. 4. Ilustracja zmian w przypowierzchniowych modelach energetycznych GaAs przy pobudzaniu energią fotonów generujących pary nośników: a) kryształ typu *n*, b) kryształ typu *p*

Efekt c) wzrostu prądu nasycenia przy dotatkowym oświetleniu jest, jak się wydaje, również związany ze zmianą zagięcia pasm. Spłaszczenie zagięcia może spowodować penetrację światła w większym obszarze pasma walencyjnego niż poprzednio (zwiększenie „pola emisji", rys. 3b). W związku z tym wzrasta ilość emitowanych elektronów o średnich i niskich energiach. Zależnie od wysokości zagięcia pasm oraz innych czynników wzrost prądu nasycenia może być bądź znaczny bądź niewielki jak w przypadku danych autora. Natomiast w obserwacjach amerykańskich badaczy [4] dochodził on do 20%. Doświadczalnym potwierdzeniem powyższego wyjaśnienia byłoby występowanie większych zmian I_s przy stymulowaniu emisji fotonami o mniejszych energiach. Potwierdzenie to uzyskano dla próbek o orientacji (100).

Zjawisko przesuwania się charakterystyk przy jednoczesnym zachowaniu wartości progu fotoemisji $h\nu_t$ świadczy, zdaniem autora, że w badanej fotoemisji dominują elektrony z pasm walencyjnych.

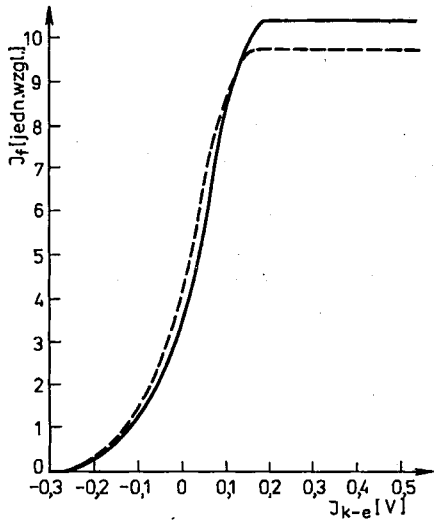
Tabela 1

Wzrost poziomu Fermiego przy zwiększaniu energii UV

	wzrost kontaktowej różnicy potencjałów ΔU_k [V]	
$h\nu$ [eV]	(100)	(111B)
4,66	0,1	
4,89	0,17	0,14
5,20	0,25	0,25
5,38	0,30	
5,48	0,33	0,35

3.1. WYNIKI DOŚWIADCZEŃ Z PRÓBKAMI TYPU „p"

Wykonano kilkanaście charakterystyk prądowo-napięciowych z powierzchni monokryształów (111B) typu *p* dla kilku energii nadfioletu. Przy pobudzaniu podczerwienią, obserwowany przez autora w kryształach GaAs typu „*p*" (rys. 5) efekt przesuwania się charakterystyk jest o wiele mniejszy i ma charakter przeciwny niż w próbkach typu „*n*".



Rys. 5. Porównanie charakterystyk fotoemisji próbek GaAs typu *p* przy użyciu i po zgaszeniu dodatkowego oświetlenia podczerwienią: a) orientacja próbki (111), hν=4.89 eV

Występuje też zmniejszanie się prądów nasycenia przy zastosowaniu dodatkowego oświetlenia. Efekt spłaszczenia zagięcia pasm jest tutaj zapewne znikomy, gdyż – jak już wykazywano – samo przypowierzchniowe zagięcie w próbkach typu „*p*" jest bardzo małe względem próbek typu „*n*". Przesuwanie krzywych w przeciwną stronę (w lewo) niż w doświadczeniach z próbkami typu „n" świadczy jednak o pewnym obniżaniu się poziomu Fermiego względem poziomu próżni.

4. ANALIZA WYNIKÓW

Z otrzymanych wyników widać, że najwyraźniej występuje tu efekt zmniejszania się pracy wyjścia φ fotoelektronów z próbek typu „n" przy jednoczesnym pobudzaniu ich podczerwienią i nadfioletem. Efekt ten mogą wywoływać następujące przyczyny:

a) zapełnianie się stanów powierzchniowych wzbudzonymi podczerwienią elektronami walencyjnymi,

b) uwalnianie par nośników w obszarach absorpcji podczerwieni i – na skutek różnych ruchliwości – zmiana rozkładu ładunków w warstwach przypowierzchniowych

(efekt Dembera [5] [6]. Zdaniem autora, w badanych próbkach decydującym jest drugi efekt.

Jeśli pobudzające promieniowanie ma energię wystarczającą do generacji par „elektron — dziura", to w obszarze absorpcji znacznie wzrasta ilość tych nośników, przy czym w stanie stabilizacji $\Delta n_e = \Delta n_d$ na skutek ich różnych ruchliwości i współczynników dyfuzji. W pewnym obszarze, przy powierzchni półprzewodnika wytwarza się nadmiar ładunku dodatniego względem objętości co jest przyczyną napięcia Dembera, które dla półprzewodnika jednorodnego wyraża się wzorem:

$$U_D = \mathrm{kT}\,(\mu_e - \mu_d)\frac{\Delta n_d}{\sigma_0} \tag{1}$$

gdzie Δn_d — średni makroskopowy przyrost koncentracji dziur,

σ_0 — przewodność próbki bez oświetlenia.

Przy rozpatrywaniu pobudzania nadfioletem i podczerwienią półprzewodnika ważnym jest mieć na uwadze, że przyrost gęstości nośników występuje w obszarze o grubości około 10^6 nm, przypowierzchniowe zagięcia pasm w próbkach GaAs typu n są poniżej 100 nm, zewnętrzne fotoelektrony zaś pochodzą z głębokości $\leqslant$ 10 nm.

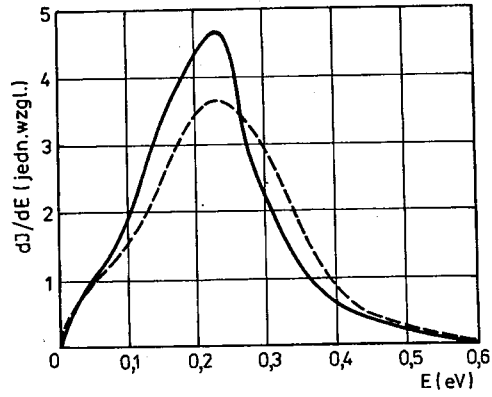
Rozpatrzmy zjawiska zachodzące w tej ostatniej warstwie w którą penetruje zarówno promieniowanie nadfioletowe jak i podczerwień. Generowane są tutaj pary nośników przez co występują przyrosty Δn_e i Δn_p koncentracji elektronów i dziur, przy czym ruchliwości tych drugich są przeszło 10 razy mniejsze, można więc traktować dziury jako nieruchome.

Opuszczanie cienkiej, kilkudziesięcioangstremowej warstwy[1] przez nadmiarowe elektrony może odbywać się bądź na skutek obsadzania stanów powierzchniowych, bądź przez dyfuzję tych nośników do wnętrza półprzewodnika, albo przez oba procesy. Fakty fizyczne o ktorych będzie mowa dalej wskazują, że pierwszy proces nie występuje, lub współdziała w znikomym stopniu. Dyfuzja elektronów z warstw najbliższych powierzchni w głąb kryształu jest także utrudniona ze względu na ekranujące działanie dużej ilości tych nośników wytworzonych we wspomnianym obszarze o grubości ca 10^6 nm. Im głębiej, tym łatwiejsza ucieczka do objętości półprzewodnika, zatem wytwarza się gradient potencjału w warstwie przypowierzchniowej mający charakter „spłaszczenia" zagięcia pasm (rys. 4). Najbliżej powierzchni — na skutek utrudnionej ucieczki elektronów — można przyjąć, że gęstości elektronów i dziur generowanych podczerwienią są zbliżone, co z kolei uzasadnia stałość wartości $h\nu_t$ i powinowactwa elektronowego χ_S. Wspomniane spłaszczenie zagięcia pasm musi spowodować przesunięcie poziomu Fermiego do góry, (obserwowany na charakterystykach wzrost napięcia kontaktowego) bez zmiany fotoelektrycznego progu na powierzchni.

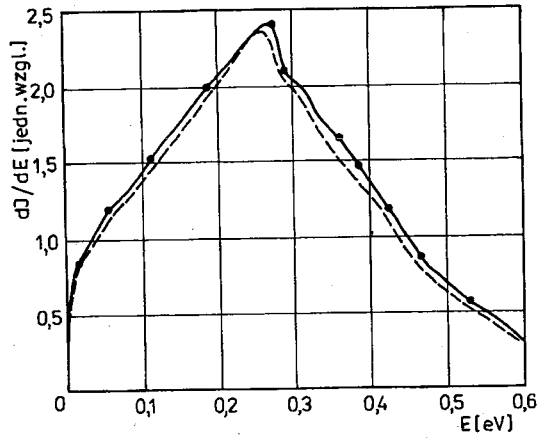
Gdyby dominującą rolę w tym zjawisku grała zmiana ładunku napowierzchniowego wówczas fotoelektryczne progi przy powierzchni i wydajności wyraźnie

[1] Przypowierzchniowe zagięcie w tych warstwach jest mniejsze i odwrotne dla próbek GaAs typu p w porównaniu z próbkami typu n.

rosłyby lub malały zależnie od zapełniania stanów powierzchniowych elektronami czy dziurami. Jednakże stałość $h\nu_t$ i pokrywanie się charakterystyk wydajności fotoelektrycznej przy oraz bez dodatkowego oświetlania świadczą o braku elektrycznych zmian stanu powierzchni. Potwierdza to też brak wyraźnych zmian kształtu otrzymywanych rozkładów energetycznych elektronów po oświetlaniu podczerwienią (rys. 6 i 7). Zmiany te powinny wystąpić w przypadku gdy rośnie emisja ponadwalencyjna pochodząca ze stanów powierzchniowych [7].



Rys. 6. Porównanie energetycznych rozkładów elektronów z próbki GaAs typu n emitowanych przy energii fotonów 4,89 eV bez i po użyciu dodatkowego oświetlania (krzywa przerywana — z podświetlaniem podczerwienią)



Rys. 7. Porównanie energetycznych rozkładów fotoelektronów emitowanych przy energii fotonów 5,28 eV z powierzchni GaAs (111) typu p pobudzonych oraz niepobudzonych podczerwienią (krzywa przerywana — z podświetlaniem podczerwienią).

Zdaniem autora wymiana elektronów między warstwą przypowierzchniową a powierzchnią w GaAs typu *n* jest znikoma z powodu umiejscowionego na powierzchni znacznego ładunku, o czym świadczą duże przypowierzchniowe zagięcia pasm.

Mały wzrost prądu nasycenia po dodatkowym oświetlaniu próbek *n* jest spowodowany raczej efektem Schottky'ego (przeważnie odcinki prądów nasycenia słabo wznoszą się ku górze) niż zmianą wydajności próbek.

Wzrost prądu nasycenia w kryształach typu *n* cytowani autorzy tłumaczą „obniżeniem bariery fotoelektronów". Ponieważ w swej pracy [4] nie sprecyzowali bliżej pojęcia bariery nasuwa się przypuszczenie, że chodzi o wartości $h\nu_t$ lub φ. Ta druga istotnie ulega zmniejszeniu. Jednakże wydajność fotoelektryczna zależy od progu $h\nu_t$. Zmniejszenie się odległości $(E_V - E_F)$ nie może powodować wzrostu I_s chyba że mamy do czynienia ze znaczną emisją ponadwalencyjną [8].

Przy oświetlaniu podczerwienią próbek z przewagą przewodnictwa dziurowego mechanizm zjawiska jest ten sam (spłaszczanie zagięcia pasm), choć generowane przyrosty Δn_e i Δn_d mogą być mniejsze ze względu na „opróżnienie" wierzchołków pasm walencyjnych z elektronów. Efekt przesuwania się charakterystyk w kryształach GaAS typu p obserwowany przez autora (rys. 5a) i b) jest mniej wyraźny[2]. Przesuwanie się charakterystyk w lewo świadczy o wzroście energetycznej odległości $(E_V - E_F)$, a zatem podniesieniu się poziomu Fermiego przy powierzchni (rys. 4) [9] [10].

5. PODSUMOWANIE I WNIOSKI

Efekty równoczesnego podświetlania próbek można ująć w następujące punkty:

1) Stosując dodatkowe „podświetlanie" podczerwienią badanych próbek GaAs — uzyskane wyniki różniły się wyraźnie w przypadku próbek typu *n* i typu *p*.

2) Dla płaszczyzn (100) i (111) próbek typu *n* GaAs zaobserwowano pod wpływem podczerwieni znaczne przesuwanie się prądowo-napięciowych charakterystyk fotoemisji w prawo, a dla tych samych orientacji próbek typu *p* — mniejsze przesunięcie w lewo.

3) Fotoprądy nasycenia próbek typu *n* niewiele wzrastały pod wpływem podczerwieni. W próbkach typu *p* fotoprądy nasycenia zmalały.

4) Położenia maksimów energetycznych rozkładów fotoelektronów oraz wartości długofalowych progów fotoemisji z powierzchni (100) i (111) pobudzanych oraz niepobudzanych podczerwienią, nie wykazywały mierzalnych zmian.

5) Położenia maksimów energetycznych rozkładów fotoelektronów oraz wartości długofalowych progów fotoemisji zewnętrznej z powierzchni (111) typu *p* pobudzanych oraz niepobudzanych podczerwienią również nie wykazują mierzalnych zmian.

[2] Przypowierzchniowe zagięcie pasm w tych próbkach są też mniejsze i odwrotne w stosunku do próbek typu n.

6. PRAKTYCZNE WYKORZYSTANIE ZJAWISKA MODULACJI FOTOEMISJI

1) Z badań fotoemisji modulowanej podczerwienią można uzyskać zarówno podstawowe informacje jak i praktyczne wskazówki, na przykład:

– można określić typ przewodnictwa nieznanych próbek, wtedy gdy zachodzi konieczność poznania typu przewodnictwa bez użycia kontaktów,

– przez dodatkowe pobudzanie podczerwienią można zwiększyć fotoczułość emiterów do pożądanych wartości,

– można modulować wysokość przypowierzchniowego zagięcia pasm.

2) Niektóre z powyższych stwierdzeń pozwalają wnioskować, że fotoemisja z GaAs ma charakter objętościowy. Za objętościowym charakterem fotoemisji z GaAs (na przykładzie próbek typu *p*) może przemawiać spadek prądu nasycenia przy podświetlaniu. Gdyby bowiem znaczny udział miała tu fotoemisja ze stanów powierzchniowych, to przy wzroście obsadzania tych stanów przez elektrony (co ma miejsce przy podświetlaniu światłem podczerwonym) fotoprąd nasycenia powinien wzrastać.

3) Zjawisko modulacji fotoprądu można zastosować do pracy niektórych przyrządów półprzewodnikowych.

BIBLIOGRAFIA

1. Yu. A. s z u b a; Zh. E.T.F., 1956 Tom XXVI, 1129
2. J. Ł a g o w s k i, Ch. L. B a l e s t r a, H. C. G a t o n; Surf. Sci. 1971 27, 547
3. J. W o j a s; *Influence of Additional Illumination on the Photoemission from n– and p– Type Gallium Arsenide Single Crystals.* Phys. Stat. Sol., 1969, 35, 903
4. J. H. D i n a n, L. K. G a l b r a i t h, T. E. F i s c h e r; *Electronic properties of clean cleaved {110} GaAs surfaces.* Surface Scienće 1971, t. 26, 587
5. T. M o s s, L. P i n c h e r e l e, A. M. W o o d w a r d; Proc. Phys. Soc. 1953, B66, 743
6. S. S i k o r s k i, J. Ś w i d e r s k i; *Fotoelektryczne kryteria oceny materiałów półprzewodnikowych.* Prace ITE, Warszawa 1968
7. J. W o j a s; *Badania fotoelektrycznych i powierzchniowych własności monokryształów GaAs o termicznie czyszczonych powierzchniach.* Wyższa Szkoła Pedagogiczna, Olsztyn 1976
8. B. S e r o c z y ń s k a – W o j a s; *O pewnej metodzie wykrywania nadprogowej fotoemisji z półprzewodników III–V.* Politechnika Warszawska – Prace Instytutu Fizyki, Warszawa, 1980, z. 22
9. O. V. S n i t k o; *Problemy fiziki poverchnosti poluprowodnikow.* Naukowa Dumka, Kiev 1984
10. B. I. B e d n i i, I. A. K a r p o w i c z, A. N. S a v i n o w; Poverkhnost 1988, 4, 83

J. WOJAS

MODULATION OF PHOTOEMISSION

S u m m a r y

In this paper is presented problem of measurement the shifting of current – voltage characteristics external photoemissions under the influence of additional infrared illumination samples semiconductors at big energy gap.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1992, 38, z. 2–3, ss. 431–442

(537.533.3)

Modelowanie temperatury translacyjnej w laserowym wzmacniaczu CO_2 typu TEA z prejonizacją UV

ZDZISŁAW KURZYŃSKI, ANDRZEJ KALBARCZYK

Instytut Fizyki Plazmy i Laserowej Mikrosyntezy, Warszawa

WIESŁAW WOLIŃSKI

Instytut Mikroelektroniki i Optoelektroniki, Politechnika Warszawska

Otrzymano 1991. 09. 04

Autoryzowano do druku 1991. 12.

Pomiary współczynnika wzmocnienia słabego sygnału na wielu liniach, we wzmacniaczu CO_2 typu TEA z prejonizacją UV, wykorzystano, zarówno do wyznaczenia temperatury translacyjnej ośrodka czynnego, jak również do modyfikacji tzw. modelu pięcio-temperaturowego. Zaproponowano uwzględnienie dodatkowych mechanizmów powodujących wzrost temperatury przez dodanie do równania na temperaturę translacyjną członu ξjE (ξ – stała, j – gęstość prądu, E – natężenie pola elektrycznego). Dobrą zgodność wyników eksperymentalnych i obliczeń numerycznych uzyskano dla $\xi = 0{,}3$.

1. WSTĘP

Ośrodkiem czynnym w laserach CO_2, typu TEA, z prejonizacją UV jest mieszanina gazów $CO_2 + N_2 + He$, o ciśnieniu p (koncentracji N) zbliżonym do atmosferycznego i temperaturze T. Proces pompowania odbywa się w trakcie impulsowego, samoistnego wyładowania jarzeniowego (SWJ), poprzez wytworzenie słabozjonizowanej plazmy, o stopniu jonizacji $N_e/N \cong 10^{-6} \div 10^{-5}$ (N_e – koncentracja elektronów w wyładowaniu). W plazmie tej zachodzi cały szereg zderzeń między cząstkami naładowanymi a molekułami (atomami), między samymi molekułami (atomami) oraz między cząstkami naładowanymi. Zderzające się molekuły (atomy) oraz jony mogą znajdować się w różnych stanach wzbudzonych lub w stanie podstawowym. Wyróżnia się pobudzanie rotacyjne, oscylacyjne oraz elektronowe. Energia z tych stanów może, między innymi relaksować i zwiększać energię translacyjnego ruchu molekuł – temperaturę translacyjną T. Tak więc, w zależności od intensywności poszczególnych procesow, mamy do czynienia z różnymi temperaturami ośrodka czynnego lasera (wzmacniacza). Artykuł jest próbą analizy,

poprzez konfrontowanie wyników obliczeń z wynikami eksperymentalnymi, procesów odpowiedzialnych za wzrost temperatury translacyjnej.

Do analizy wykorzystano tzw. kompleksowy model wzmacniacza CO_2, typu TEA, z prejonizacją UV [1], koncentrując się na równaniu opisującym temperaturę translacyjną ośrodka czynnego wzmacniacza.

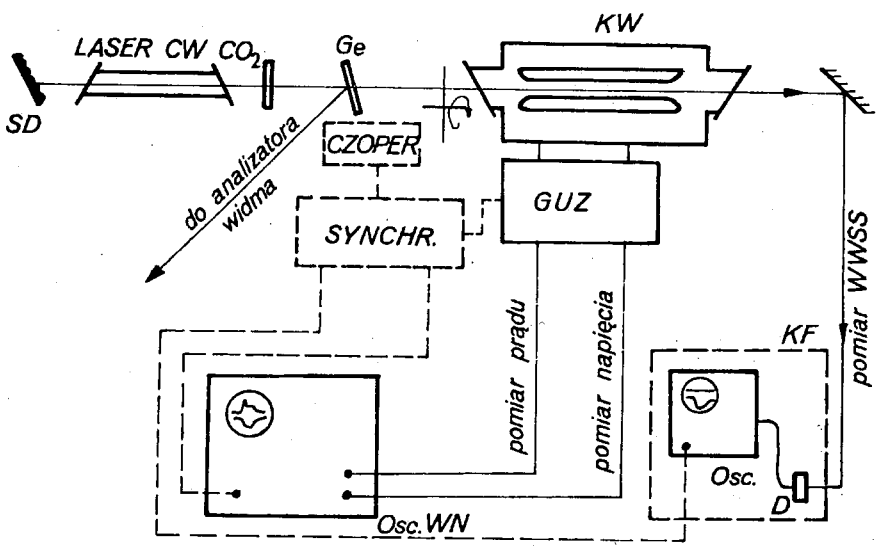
W rozdziale drugim opisano układ eksperymentalny, wykorzystywany do wyznaczania podstawowych charakterystyk. W rozdziale trzecim przedstawiono zarys metody wyznaczania temperatury, a w czwartym — analizę czynników mających wpływ na temperaturę i propozycję modyfikacji równania na temperaturę translacyjną ośrodka. Rozdział piąty zawiera krótką dyskusję otrzymanych wyników.

Przedstawione podejście umożliwia uzyskanie narzędzia, które może być szczególnie wygodne, gdy nie zależy nam na analizie poszczególnych procesów, lecz chcemy przewidywać parametry np. projektowanego lasera, czy wzmacniacza.

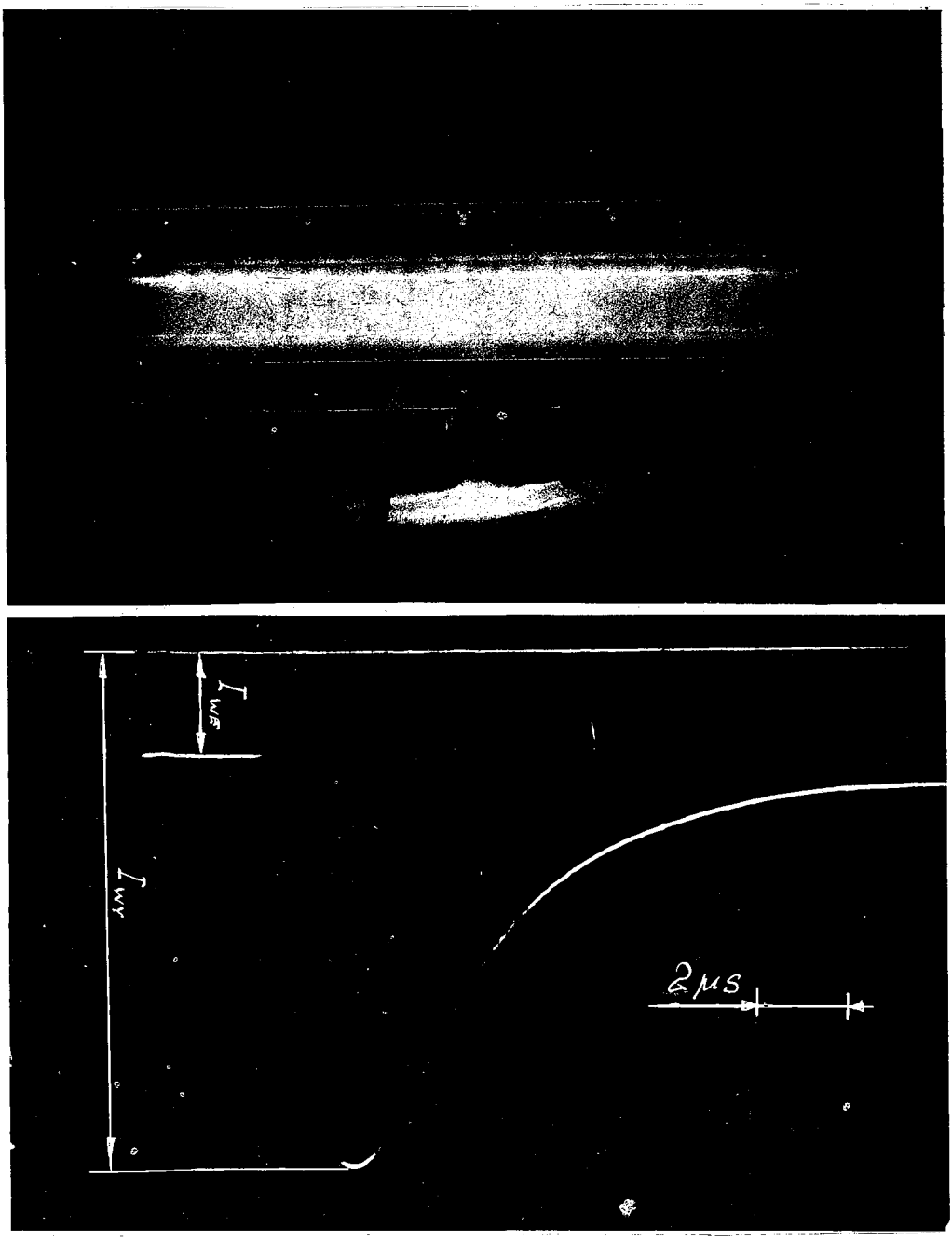
2. UKŁAD EKSPERYMENTALNY

Obiektem badań był moduł wzmacniacza CO_2, typu TEA, z prejonizacją UV, a w szczególności jego obszar wyładowania głównego. Równolegle prowadzone były pomiary parametrów elektrycznych i własności optycznych ośrodka. Pod pojęciem właściwości optycznych ośrodka rozumie się tu jego stan napompowania oraz jednorodność optyczną. Miarą tych dwu wielkości jest wartość współczynnika wzmocnienia słabego sygnału (WWSS) i jego rozkład poprzeczny. Dodatkowo rejestrowano także świecenie wyładowania głównego.

Układ do pomiaru WWSS zbudowano w oparciu o dane zaczerpnięte z literatury [2, 3, 4, 5]. Jego schemat przedstawiony jest na rys. 1. Stanowi on rozbudowaną wersję



Rys. 1. Schemat układu pomiarowego



Rys. 2. Wzmocnienie słabego sygnału w pojedynczym module wzmacniacza

układu pomiarowego, opisanego przez Wolińskiego i innych [5] i umożliwia pomiar współczynnika wzmocnienia słabego sygnału dla różnych, dowolnie wybranych, przejść oscylacyjno-rotacyjnych.

Pomiar współczynnika wzmocnienia słabego sygnału prowadzono dla jednej, wybranej linii (np.: P(18) przejścia $00°-10°0$), chcąc określić takie parametry jak: ewolucja w czasie WWSS, czas narastania wzmocnienia, czy maksymalne wzmocnienie w funkcji parametrów wzmacniacza. Wykorzystanie do sondowania wielu linii, pozwalało określić inne parametry ośrodka aktywnego, a mianowicie: jego temperaturę T oraz obsadzenia poziomów – górnego N_{001} i dolnego N_{100}. Ograniczono się do pomiarów WWSS dla gałęzi P przejścia $00°1-10°0$ (linie P(10) ÷ P(38). Ustawienie wybranej linii emisyjnej lasera kontrolowano za pomocą analizatora widma firmy Optical Engineering Inc. Santa Rosa California, umieszczając w torze głównym wiązki sondującej światłodzielącą, płaską płytkę germanową (Ge), która wydzielała część wiązki głównej i kierowała ją do analizatora widma.

Laser cw CO_2 pracował w modzie podstawowym TEM_{00}, zaś jego moc nie przekraczała pojedyńczych watów i można ją było regulować poprzez zmianę ciśnienia mieszaniny roboczej lub poprzez zmiany prądu wyładowania. Ciągła wiązka lasera była zamieniana na szereg prostokątnych impulsów przy pomocy czopera. Urządzenie to, oprócz wymienionej funkcji, pozwalało synchronizować jeden z wyciętych impulsów prostokątnych z impulsowym wyładowaniem zachodzącym w badanej komorze wzmacniacza (KW) oraz z układem rejestrującym (pomiarowym).

Impuls sondujący oraz wzmocniony były rejesterowane za pomocą detektorów (D) i oscyloskopu (Osc), umieszczonych w klatce Faraday'a (KF). Użyte do pomiaru WWSS detektory (prod. Spółki VIGO) były niechłodzonymi fotorezystorami CdHgTe, działającymi w oparciu o efekt fotoprzewodnictwa. Na jednym przebiegu oscyloskopowym rejestrowano zarówno impuls sondujący, jak i wzmocniony (rys. 2).

Fotografowanie świecenia wyładowania głównego prowadzono w celu określenia jego jednorodności przestrzennej. Umożliwiało to właściwą interpretację pomiaru WWSS. Rejestracja przebiegów prądowo-napięciowych pozwalała, co prawda, na wychwycenie wyładowań łukowych, nie pozwalała jednak stwierdzić występowania bardziej subtelnych niejednorodności, np.: wyładowań włóknistych.

3. WYZNACZANIE TEMPERATURY TRANSLACYJNEJ OŚRODKA CZYNNEGO WZMACNIACZA CO_2 TYPU TEA

Jedną z metod, najszerzej stosowaną, wyznaczania temperatury translacyjnej ośrodka aktywnego T we wzmacniaczach CO_2 jest pomiar, na wielu liniach, WWSS. Wykorzystuje się przy tym różne przejścia oscylacyjno-rotacyjne [6, 7, 8,]. W celu opisania metody ograniczymy się do przejścia 10P. Wzór na WWSS można zapisać w postaci [1, 9]:

$$g_0(J) = \frac{2hcB}{kT}\sigma_0(J,T)\left\{N_{001}(2J-1)\exp\left[-\frac{hcB}{kT}J(J-1)\right]-\right.$$

$$\frac{2J-1}{2J+1} N_{100} (2J+1) \exp\left[-\frac{hcB}{kT} J(J+1)\right]\Big\}, \quad (1)$$

gdzie: h – stała Plancka, c – prędkość światła, B – stała rotacyjna, k – stała Boltzmanna, σ_0 (J, T) – przekrój czynny na emisję wymuszoną, J – rotacyjna liczba kwantowa. Pomiar WWSS dla co najmniej trzech różnych J (różnych przejść oscylacyjno-rotacyjnych) pozwala uzyskać układ 3-ch równań typu (1), z niewiadomymi: T, N_{001} i N_{100}. Rozwiązując numerycznie ten układ równań można wyznaczyć poszukiwane wielkości, w tym również temperaturę T.

Dokładność metody pomiarowej rośnie ze wzrostem ilości punktów pomiarowych WWSS, a więc do sondowania ośrodka, należy stosować możliwie dużą ilość linii oscylacyjno-rotacyjnych. Jednakże, jak wynika z prac [6, 10, 11, 12, 13], niektóre linie posiadają większe wzmocnienia w wyniku udziału w nich pasm sekwencyjnych i gorących. W celu uniknięcia wpływu powyższych efektów na wyznaczanie interesujących nas wielkości, przyjęto odpowiednią procedurę postępowania opisaną w [1]. W wyniku zastosowanej procedury, z szesnastu linii, ograniczono się do następujących siedmiu: P(14), P(18), (P22), P(24), P(30), P(32), P(36).

Pomiary przeprowadzono dla mieszanin aktywnych o składzie CO_2:N_2:He = = 1:1:4, 1:1:6 i 1:1:8. Uzyskane w pracy wyniki eksperymentalne są zgodne z danymi literaturowymi [2, 6, 14, 15, 16]. Temperatura ośrodka osiąga (400 ∓ 10)K i brak wyraźnej zależności od składu mieszaniny [1].

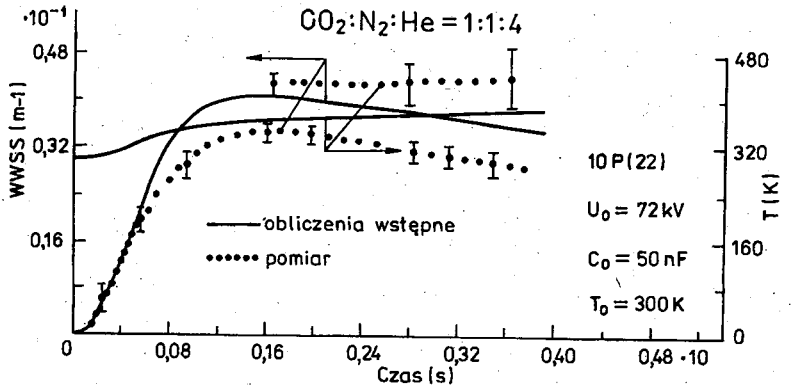
4. MODELOWANIE TEMPERATURY TRANSLACYJNEJ OŚRODKA CZYNNEGO WZMACNIACZA CO_2 TYPU TEA

Równanie opisujące temperaturę translacyjną w modelu pięciotemperaturowym ma postać [1, 9]:

$$\frac{d\mathscr{E}}{dt} = C_v (dT/dt) = \frac{\mathscr{E}_1 - \mathscr{E}_1(T)}{\tau_{10}(T)} + \frac{\mathscr{E}_2 - \mathscr{E}_2(T)}{\tau_{20}(T)} + (1 - h\nu_1/h\nu_3 - h\nu_2/h\nu_3) \frac{\mathscr{E}_3 - \mathscr{E}_3(T, T_1, T_2)}{\tau_3(T, T_1, T_2)}, \quad (2)$$

gdzie: $\mathscr{E}$ – gęstość energii translacyjnej molekuł ośrodka, C_v – ciepło właściwe mieszaniny gazów, $\mathscr{E}_i$ – gęstość energii w modach oscylacyjnych molekuły CO_2 (C_i = 1 – mod podłużny symetryczny, i = 2 – mod deformacyjny poprzeczny, i = 3 – mod podłużny asymetryczny), T_i – efektywne temperatury oscylacyjne (i = 1,2,3 odpowiednio), ν_i – częstości podstawowe modów (i = 1,2,3 odpowiednio), τ_i – czasy relaksacji z poszczególnych modów do translacji (relaksacja V – T).

Wykorzystując równanie (2) w modelu kompleksowym [1], wyznaczono czasowe zmiany WWSS dla linii 10P(22). Obliczenia wykonano dla następujących warunków: skład mieszaniny aktywnej CO_2:N_2:He = 1:1:4, warunki początkowe $T(0) = T_i(0) = T_0 = 300$ K (i = 1, . . . , 4). Wyniki przedstawiono na rys. 3. Na rysunku



Rys. 3. Zależności czasowe współczynnika wzmocnienia słabego sygnału WWSS i temperatury translacyjnej T, obliczonych z modelu pięciotemperaturowego dla $T_0 = 300$ K oraz zależność czasowa zmierzonego WWSS

zaznaczono także rezultaty odpowiednich pomiarów WWSS oraz, dodatkowo, obliczoną z modelu pięciotemperaturowego, czasową ewolucję temperatury translacyjnej. Widać, że obliczony WWSS, dla czasów $t \geqslant 0{,}5\mu s$, jest większy od wartości zmierzonej, a różnica w maksimum WWSS wynosi około 30%. Widać dalej, że wyliczana wartość T, w maksimum wzmocnienia, dla analizowanego wzmacniacza, osiąga około 350 K. Wyznaczona eksperymentalnie, dla warunków zastosowanych w obliczeniach, temperatura translacyjna przekracza nieznacznie wartość 400 K [1]. Podobne rezultaty uzyskiwano na innych układach o zbliżonych parametrach [2, 6,14, 15, 16]. Takie duże rozbieżności między wynikami eksperymentalnymi, a odpowiadającymi im wynikami obliczeń, wymagają bardziej wnikliwego rozważenia procesów mogących wpływać na analizowane wielkości.

Wpływ temperatury T na wyliczane watości WWSS jest bardzo złożony. Wynika on z zależności od temperatury czasów relaksacji, czy przekroju czynnego na emisję wymuszoną . Od temperatury T zależy także obsadzenie poziomów rotacyjnych danego poziomu oscylacyjnego.

Przyczyn opisanych powyżej rozbieżności można upatrywać w niewłaściwym dobraniu lub wyliczeniu takich wielkości, występujących w równaniach modelu 5—T, jak: stałe szybkości pobudzania, koncentracja elektronów, czasy relaksacji i inne, jak również, w nieuwzględnieniu w modelu 5—T wszystkich, mających wpływ na WWSS zjawisk, np.: przy modelowaniu zmian temperatury translacyjnej T (równanie (2)).

Stałe szybkości pobudzania wyznaczano na podstawie rozwiązania równania Boltzmanna [1]. Dla ustalonego składu i ciśnienia mieszaniny gazowej są one funkcją natężenia zredukowanego pola elektrycznego E/N, którego zmiany czasowe określano z rozwiązania pełnego modelu SWJ. Użyte w równaniu Boltzmanna przekroje czynne są typowe i powszechnie używane do modelowania [17]. Wyznaczane, na ich podstawie, współczynniki szybkości pobudzania porównywano z wartościami uzyskiwanymi przez innych autorów dla innych, niż stosowane w pracy, mieszanin; różnice nie przekraczały 5% [1, 18]. Tak więc dane te przyjęto za wiarygodne. Koncentracja

elektronów N_e wyliczana jest praktycznie z eksperymentalnie zmierzonego prądu $I(N_e = 1/Se \cdot I/W)$, zaś prędkość dryfu W jest najdokładniej wyznaczoną, z równania Boltzmanna, wielkością.

Czasy relaksacji, z kolei, są wzięte z [1] (patrz literatura tam cytowana), stosowane w innych modelach [11, 19], dają wyniki poprawne i dlatego przyjęto je za wiarygodne. Są one jednak funkcjami temperatury translacyjnej T. Fakt ten potwierdza wcześniejsze sugestie o konieczności, w celu uzyskania dokładniejszych wyników z modelu 5–T, właściwego modelowania zmian temperatury ośrodka aktywnego T. Tym bardziej, że parametryczne miany warunków początkowych (T_0) [1] wskazują na wyraźną zależność otrzymanych wartości WWSS od zmian T.

W trakcie wyprowadzania równań modelu pięciotemperaturowego dokonano szeregu założeń upraszczających. Dotyczyły one przede wszystkim pominięcia niektórych procesów zderzeniowych (patrz np.: [9, 20]). Do pominiętych, a mogących mieć bezpośredni wpływ na wzrost temperatury translacyjnej, do momentu uzyskania maksimum wzmocnienia, można zaliczyć:

– grzanie ośrodka przez elektrony samoistnego wyładowania jarzeniowego,

– relaksację typu V–T modu $\mathscr{E}_4$ (gęstość energii drgań oscylacyjnych molekuł azotu), opisywaną czasem relaksacji τ_{40}, relaksację typu V–T modu $\mathscr{E}_3$ z czasem relaksacji τ_{30} i inne procesy relaksacyjne typu V–V opisywane np.: w [9].

Uwzględnienie, wymienionych powyżej procesów relaksacyjnych napotyka na trudności, wynikające z braku danych odnośnie odpowiednich czasów relaksacji. Brakuje również dokładniejszych informacji na temat wpływu tych czasów na zmianę temperatury translacyjnej T.

Istnienie grzania ośrodka przez elektrony samoistnego wyładowania jarzeniowego sugeruje, natomiast, praca teoretyczna [21]. Rozważmy jej wyniki w świetle warunków występujących w badanym wzmacniaczu. Z uzyskanych z rozwiązania modelu SWJ, dla CO_2:N_2:He = 1:1:4, wynika, że podstawowa część energii wprowadzana do wyładowania odbywa się dla $E/N \approx (2{,}5 \div 4{,}3)10^{-16}$ V cm^2 [1]. Dla tego zakresu natężeń zredukowanych pól elektrycznych w [21] podaje się, że dla mieszaniny CO_2:N_2:He = 1:2:3 około 10%, natomiast dla mieszaniny 1:0,25:3 aż ok. 35% mocy wyładowania jest rozpraszane w zderzeniach sprężystych elektronów z molekułami (atomami). Jest to istotny argument za uwzględnieniem efektu grzania ośrodka aktywnego przez elektrony SWJ. Za jego uwzględnieniem przemawiają również niektóre prace dotyczące modelowania wzmacniaczy za pomocą niesamoistnego wyładowania jarzeniowego [11, 20].

Zmianę energii średniego elektronu, na skutek zderzeń sprężystych elektronów z molekułami (atomami) ośrodka czynnego, jeśli średnia energia elektronów jest dużo większa od średniej energii molekuł (atomów), można zapisać [22]:

$$\frac{d\varepsilon}{dt} = \varepsilon \nu_m \delta . \tag{3}$$

Uwzględniając, że $\varepsilon = \beta E$, $W/v \approx \sqrt{\delta}$ oraz $\nu_m = NQ_m v$, po prostych przekształceniach, uzyskuje się:

$$\frac{d\varepsilon}{dt} = \beta N\sqrt{\sigma Q_m}\, WE \tag{4a}$$

gdzie: ε — energia średniego elektronu, ν_m, Q_m — częstość zderzeń i przekrój czynny na transport pędu w zderzeniu elektronu z molekułą (atomem), σ — średnia stała energii, jakiej doznaje elektron przy każdym zderzeniu z molekułą (atomem); w zderzeniu sprężystym $\sigma = 2m/M$, gdzie M — masa molekuły (atomu), β — stała, E — natężenie pola elektrycznego, W — prędkość dryfu elektronu, v — prędkość elektronu. Ponieważ w jednostce objętości ośrodka aktywnego znajduje się N_e elektronów, wobec tego zmiana energii ruchu translacyjnego molekuł (atomów) ośrodka, w wyniku zderzeń sprężystych elektronów z tymi molekułami (atomami), będzie:

$$\left(\frac{d\mathscr{E}}{dt}\right)_{e-T} = \frac{d\varepsilon}{dt} N_e = \beta N\sqrt{\sigma Q_m}\, WN_eE. \tag{5}$$

Stąd ostatecznie można uzyskać:

$$\left(\frac{d\mathscr{E}}{dt}\right)_{e-T} = \varepsilon jE. \tag{6}$$

gdzie:

$$\xi = \frac{\beta N\sqrt{\delta Q_m}}{e} \tag{7}$$

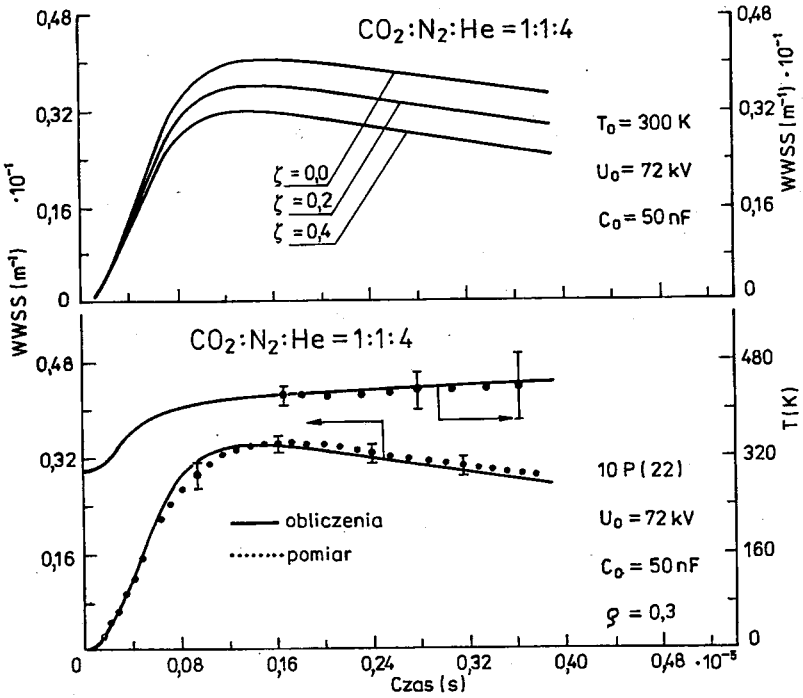
i jest ono w przybliżeniu stałe [1]. Dla rozważanych warunków oszacowana wartość $\xi \approx 0{,}7$. Jest to wartość zawyżona i w sposób prosty wynika z przyjętych uproszczeń, daje jednak ocenę wielkości ξ i stwarza możliwość poszukiwania dokładniejszej jej wartości. Mając dalej świadomość, że powyższe zależności uzyskano na podstawie teorii elementarnej oraz zakładając dominującą rolę zderzeń sprężystych, wartość parametru ξ należy dobrać, konfrontując wyniki obliczeń i pomiarów. Ponadto, w całym zakresie natężeń zredukowanych pól elektrycznych, występujących w trakcie wyładowania, ξ może się zmieniać w czasie, tym bardziej, że w całych dotychczasowych rozważaniach nie uwzględnione było pobudzanie poziomów elektronowych. Równanie (2) można więc zapisać:

$$\frac{d\mathscr{E}}{dt} = C_v(dT/dt) = \frac{\mathscr{E}_1 - \mathscr{E}_1(T)}{\tau_{10}(T)} + \frac{\mathscr{E}_2 - \mathscr{E}_2(T)}{\tau_{20}(T)}$$

$$(1 - h\nu/h\nu_3 - h\nu_2/h\nu_3)\frac{\mathscr{E}_3 - \mathscr{E}_3(T, T_1, T_2)}{\tau(T, T_1, T_2)} + \xi jE. \tag{8}$$

Takie podejście powoduje, że w parametrze ξ, oprócz efektu grzania ośrodka czynnego przez elektrony wyładowania (zderzenia elektronów z przekazaniem pędu i pobudzanie poziomów elektronowych), może również tkwić udział pominiętych, bądź niedokładnie opisanych procesów relaksacyjnych.

Wprowadzając więc do równań modelu 5 — T w miejsce równania (2) równanie (8), przeprowadzono obliczenia, parametryzując wielkość ξ w zakresie $0 \div 0{,}5$ co 0,1,



Rys. 4. Zmiany w czasie współczynnika wzmocnienia słabego sygnału WWSS i temperatury translacyjnej T dla $\zeta = 0{,}3$, $T_0 = 300$ K i CO_2:N_2:H_e = 1:1:4. Porównanie obliczeń z pomiarami

przyjmując jako warunki początkowe $T(0) = T_i(0) = T_0 = 300$ K. Dobre dopasowanie do wyników eksperymentalnych uzyskano dla $\xi = 0{,}3$. Porównanie wyników eksperymentalnych i obliczeniowych, dla tych warunków, przedstawia rys.4. Analogiczne wyniki otrzymano dla pozostałych analizowanych mieszanin gazowych CO_2:N_2:He = 1:1:6 i 1:1:8.

Na podstawie wszystkich przeprowadzonych obliczeń i ich porównań z pomiarami ustalono, że parametr ξ powinien wynosić około 0,3 czyli, że około 30% energii całkowitej tracone jest na grzanie ośrodka czynnego. Dla takiej wartości parametru ξ uzyskuje się najlepszą zgodność wielkości obliczonych z modelu 5—T (WWSS i T) z pomiarami, dla wszystkich analizowanych składów mieszanin gazowych.

5. DYSKUSJA I WNIOSKI

Rozpatrzmy powyżej uzyskany wynik, z punktu widznia bilansu energii. Z przeprowadzonej w [1] analizy wynika, że ponad 70% energii wprowadzanej do wyładowania, transportowana jest w drgania oscylacyjne molekuł i bilansuje się z dokładnością do pojedyńczych procentów z energią traconą na bezpośrednie grzanie molekuł (atomów) ośrodka. Osobnym zagadnieniem jest możliwość określenia

udziału poszczególnych procesów w energii kanału bezpośredniego. Biorąc pod uwagę, że energia tracona na jonizację jest, dla zakresów natężeń zredukowanego pola elektrycznego występującego w rozważanym wzmacniaczu, do pominięcia [21], można przyjąć, że dominujące straty energii wprowadzanej do wyładowania realizowane są przez sprężyste rozpraszanie elektronów na molekułach (atomach) ośrodka czynnego i przez zderzenia z molekułami (atomami), prowadzące do ich wzbudzeń elektronowych. Rozdział energii na wymienione zjawiska jest złożony i wyznaczenie go wymagałoby wprowadzenia istotnych zmian w modelu kompleksowym. Na podstawie cytowanej już pracy teoretycznej [21], dla nieco innych składów mieszanin gazowych, w rozważanym zakresie natężeń zredukowanych pól elektrycznych, uzyskano, że na zderzenia sprężyste tracone jest około dwukrotnie więcej energii niż na zderzenia, prowadząc do wzbudzeń poziomów elektronowych. Zakładając, że ten stosunek niewiele zmienia się ze składem mieszaniny gazowej, dla analizowanego wzmacniacza należałoby przyjąć, że około 10% energii całkowitej tracone jest na pobudzanie elektronowe, a kilkanaście do kilkudziesięciu na zderzenia sprężyste. W pracy [23], zamieszczona jest informacja, podana na podstawie niedostępnego w kraju raportu Air Force Weapons Laboratory (AFWL-74-106), że energia powodująca grzanie wyładowania, pochodzi głównie z pobudzonych poziomów elektronowych. Nie podano jednak, jaki mechanizm jest odpowiedzialny za taki transfer energii.

Na podstawie przeprowadzonych rozważań widać, że wprowadzenie do równania (2) poprawki, przez dodanie członu ξjE, daje dobrą zgodność wyników obliczeń z pomiarami dla WWSS i T. Parametr ξ, mimo że wprowadzony na podstawie rozważenia sprężystych rozproszeń elektronów na molekułach (atomach), zawiera udział wszystkich innych, nie uwzględnionych procesów podwyższających temperaturę T. Określenie, który z pominiętych procesów jest dominujący, przy stosowanej metodzie, jest niemożliwe. Rozwiązanie takiego problemu wymagałoby poszerzenia modelu o zaniedbane procesy.

Przedstawione podejście umożliwia uzyskanie narzędzia, które może być szczególnie wygodne, gdy nie zależy nam na analizie poszczególnych procesów, lecz chcemy przewidywać parametry, np. projektowanego lasera, czy wzmacniacza.

BIBLIOGRAFIA

1. Z. Kurzyński, A. Kalbarczyk: *Laserowy wzmacniacz* CO_2 *typu* TEA z prejonizacją UV, Rozprawa doktorska PW., 1989
2. E. A. Balik, B.K. Garside, J. Reid, T. Tricker: *Reduction of the pumping efficiency in* CO_2 *laser at high discharge energy*. J. Appl. Phys. 1975 **46**, pp. 1322 ÷ 1331
3. A. M. Robinson: *Effect of Capacitance on Gain in a Transversely Pulsed* CO_2 *Discharge*. Can. J. Phys. 1972 **50**, pp. 2138 ÷ 2148
4. A. M. Robinson: *Gain Distribution in a* CO_2 TEA *Laser*. J. Phys. 1989 **50**, pp. 2471 ÷ 2473
5. W. Woliński, Z. Olbromski, W. Korbowicz: *Pomiar współczynnika wzmocnienia w laserze molekularnym* CO_2–TEA. EKON'76, A – 39
6. L. A. Weaver, L. H. Taylor, L. J. Denes: *Rotational temperature determinations in molecular gas lasers*. J. Appl. Phys. 1975 **49**, pp. 3951 ÷ 3964

7. S. P. Wagin, Ju. A. Jacobi: *Izmerenie temperatury gazovych sred soderzaszczich uglekislyj gaz, metodom lasernogo zondirovanija.* Inž. Fiz. Zurn. 1978, t. **35**, pp. 5 ÷ 10
8. P. V. Avizonis, D. R. Dean, R. Grotbeck: *Determination of vibrational and ranslational temperatures in gas – dynamic lasers,* Appl. Phys. Lett., 1973, **23**, pp. 375 ÷ 378
9. K. Smith, R. M. Thomson: *Cislennoe modelirovanie gazovych lazerov.* Perevod s angliskogo, izd. „Mir", Moskwa, 1981
10. W. T. Leland, M. G. Kircher, M. G. Nutter, G. T. Schapper: *Rotational temperature in a high-pressure pulsed* CO_2 *laser.* J. Appl. Phys. 1975, **46**, pp. 2174 ÷ 2176
11. J. C. Comly JR., W. T. Leland, C. J. Elliott, R. M. Hunter II, M. J. Kircher: *Discharge and Kinetics Modeling in Electron–Beam Controlled* CO_2 *Laser Amplifiers.* IEEE J. Quant. Electr. 1981, **QE–17**, pp. 1786 ÷ 1798
12. P. Lavigne, J. L. Lachambre, G. Otis: *Gain measurements on the sequence bands in a* TEA–CO_2 *amplifier.* J. Appl. Phys. 1978, **49**, pp. 3714 ÷ 3718
13. J. Reid, K. Siemsen: *Laser power and gain measurements on the sequence bands of* CO_2. J. Appl. Phys. 1977, **48**, pp. 2712 ÷ 2117
14. D. K. Garside, J. Reid, E. A. Ballik: TE CO_2 *laser Pulse Decay and Molecular Relaxation Rates.* IEEE J. Quant. Electr. 1975, **QE–11**, pp. 538 ÷ 589
15. C. Dang, J. Reid, B. K. Garside: *Gain limitations in* TE CO_2 *Laser Amplifiers.* IEEE J. Quant. Electron. 1980, **QE–16**, pp. 1097 ÷ 1102
16. G. Girrard, J. C. Farcy, M. Michon, P. Vincent: *Experiments and Theory on Energy Extraction in an Atmospheric–Pressure* CO_2 *Amplifier with 60ns ÷ 1μs Pulses.* IEEE J. Quant. Electron. 1972, **QE–11**, pp. 241 ÷ 247
17. L. J. Kieffer: *A Compilation of Electron Collision Cross Section Date for Modeling Gas Discharge Lasers.* University of Colorado, Boulder, Colorado, 1973
18. A. Kalbarczyk, Z. Kurzyński i inni: *Analysis of energy storage and extraction in* TEA UV–*preaionized* CO_2 *laser amplifiers. Part I. Numerical determination of effective vibrational excitation rate coefficients and transport coefficients.* J. Techn. Phys. 1987, **28**, pp. 251 ÷ 262
19. R. L. Carlson, J. P. Carpenter, D. E. Casperson, R. B. Gibson, R. P. Godwin, R. F. Haglund Jr., J. A. Haneon, E. L. Jolles, T. F. Stratton: *Helios: A 15 TW Carbon dioxide Laser–fusion focility.* J. Quant. Electr. 1981, **QE–17**, pp. 1662 ÷ 1678
20. E. W. Mc Daniel, W. L. Nighan: *Gas Lasers in Applied Atomic Collision Physics.* Volume 3, Academic Press 1982, A Subsidiary of Harcourt Brace Jovanovich-Publishers
21. J. J. Lowke, A. V. Phelps, B. W. Irwin: *Predicted Electron Transport Coefficients and Operating Characteristics of* CO_2–N_2–He *Laser Mixtures.* J. Appl. Phys. 1973, **44**, pp. 4664 ÷ 4671
22. Ju. P. Rajzer: *Fizyka gazovogo razjada.* Izd. „Nauka", Moskwa, 1987, glava 4, 7, 10
23. B. E. Cherrington: *Gaseous electronics and gas lasers.* International series in natural philosophy 1979, vol. 94

Z. KURZYŃSKI, A. KALBARCZYK, W. WOLIŃSKI

MODELLING OF TRANSLATION TEMPERATURE IN PREJONIZED TEA CO_2 LASER AMPLIFIER

Summary

The measurements of the small-signal gain coefficient in the many lines were exploit in the TEA CO_2 laser amplifier PREJONIZED TEA for the calculations of temperatures as well for the modification of the five-temperature modell. It was suggested the allowance of the additional mechanisms caused of the rise of translational temperature by adding ξjE (ξ – constant, j – density of the current, E – intensity of electric field). The good agreement of the experimental and numerical results was obtained.

KWARTALNIK ELEKRONIKI I TELEKOMUNIKACJI, 1991. 38, z. 2—3, ss. 443—454

(621.3.088.3)

Przestrzeń probabilistyczna sumowania błędów instrumentalnych

Ogólne zasady doboru wyrażeń na błędy graniczne i obliczania poziomów ufności

ROMUALD RAKOWSKI

Wyższa Oficerska Szkoła Radiotechniczna, Jelenia Góra

Otrzymano 1991. 09. 13

Autoryzowano do druku 1991. 11. 20

Przedstawiono przestrzeń probabilistyczną odpowiadającą statystycznemu składaniu błędów cząstkowych pomiarów pośrednich i układów przetworników. Określono, bądź założono — na podstawie badań eksperymentalnych — graniczne rozkłady błędów dla kilku podstawowych zbiorów zdarzeń. Podano ogólnie procedurę obliczania poziomów ufności i dobierania wyrażeń na błędy graniczne.

1. WSTĘP

Zagadnienia obliczania błędów pomiarowych i układów przetworników metodami probabilistycznymi należą do istotnych w teorii pomiarów. Aby je przejrzyście przedstawić należy zacząć od skonstruowania właściwego modelu probabilistycznego, co obejmuje zazwyczaj odpowiedni dobór przestrzeni zdarzeń elementarnych, a następnie określenia miar na ciele zdarzeń tej przestrzeni-prawdopodobieństw nie przekroczenia błędów granicznych wybranych sposobów składania niepewności cząstkowych. Należy więc określić przestrzeń zdarzeń elementarnych i wyodrębnić w niej podzbiory odpowiadające spotykanym w technice rodzajem pomiarów — zbiorom wyników doświadczeń. Występuje tu duża dowolność i konieczność kierowania się informacjami o danym zagadnieniu, tym razem informacjami o pomiarach i własnościach aparatury pomiarowej, które nie należą do teorii prawdopodobieństwa.

Niniejsze opracowanie ogranicza się do określenia przestrzeni probabilistycznej i poziomów ufności sumowania błędów narzędziowych, instrumentalnych przy założeniu, że inne błędy mogą być pominięte. Dotyczy ono obliczania błędów wypadkowych, granicznych, pomiarów pośrednich i układów przetworników, a więc

zagadnień ważnych dla praktyki i teorii pomiarów, którymi często musi zajmować się technik, inżynier i pracownik naukowy.

Prezentację zagadnienia wskazanym jest zacząć od przedstawienia modelu błędu instrumentalnego Δx, a następnie przejść do opisu przestrzeni zdarzeń elementarnych i jej podzbiorów występujących w praktyce pomiarowej. Dopiero w oparciu o te ustalenia można poszukiwać metod obliczania wyrażeń na błędy graniczne i poziomy ufności. Przyjęta procedura postępowania powinna uchronić od nieporozumień i stworzyć klarowny obraz analizowanych zależności.

2. MODEL BŁĘDU PODSTAWOWEGO, NARZĘDZIOWEGO

Błąd narzędzia pomiarowego, instrumentalny, zawarty w przedziale ustalonym klasą jest zdarzeniem w przestrzeni zdarzeń elementarnych, określonym przez następujące współrzędne:

- typ przyrządu, producenta — v
- wartość mierzona (punkt na podzielni) — x
- czs eksploatacji, mierzony od momentu wzorcowania — t_e
- egzemplarz, numer egzemplarza — e.

Model błedu podstawowego narzędziowego można zatem zapisać w postaci [2, 3, 5]

$$\Delta x = \Delta x\,(v, x, t_e, e). \tag{1}$$

Zależność od typu przyrządu v rozumie się w sensie zależności błędu podstawowego od producenta, rodzaju konstrukcji, populacji przyrządów jednakowo oznakowanych np. woltomierzy V-640. Typy przyrządów różnią się między sobą także wartościami parametrów rozkładów błędów, co stanowi kryterium losowego doboru typu.

3. PRZESTRZEŃ ZDARZEŃ ELEMENTARNYCH BŁĘDÓW PODSTAWOWYCH. CIAŁO ZDARZEŃ — PODZBIORY BŁĘDÓW ODNOSZĄCE SIĘ DO RÓŻNYCH RODZAI I WARUNKÓW POMIARÓW

Szeroko rozumianymi tu wynikami doświadczeń będą błędy podstawowe (instrumentalne) występujące podczas pomiarów przyrządami różnych typów, o różnym czasie eksploatacji, różnych wielkości i wartości. Zbiór tych błędów — podstawowych (narzędziowych), wszystkich przyrządów eksploatowanych we współczesnej technice pomiarowej tworzy przestrzeń zdarzeń elementarnych, którą oznacza się tu symbolem V. Ciałem zdarzeń $\mathcal{F}$ będą odpowiednio dobrane rodziny podzbiorów błędów ze zbioru $\{\Delta x\}$, odpowiadające potrzebom praktyki.

W technice pomiarowej, zbiór zdarzeń elementarnych V_Ω wskazanym jest na wstępie podzielić na podzbiory przyrządów i związanych z nimi błędów w zależności od rodzaju mierzonej wielkości np. napięcia V_1, prądu V_2, mocy V_3 itp.

Otzymamy zatem podzbiory

$$V_1, V_2, \ldots, V_d \tag{2}$$

które można zapisać w postaci

$$V_k \text{ dla } k = 1, 2, \ldots d \tag{3}$$

należące do przestrzeni zdarzeń elementarnych

$$V_k \in V_\Omega \tag{4}$$

gdzie k – numer, oznaczenie wielkości mierzonej.
Suma tych podzbiorów tworzy przestrzeń zdarzeń elementarnych

$$V_1 \cup V_2 \cup \ldots \cup V_d = V_\Omega \tag{5}$$

przy czym d jest liczbą rozróżnianych wielkości mierzonych.
Wystąpi zatem podział błędów na zbiory

V_1 – zbiór $\{\Delta x_{V_1}\}$ błędów zależ. od $\Delta x_{V_1}(v, x, t_e, e)$

V_2 – zbiór $\{\Delta x_{V_2}\}$ błędów zależ. od $\Delta x_{V_2}(v, x, t_e, e)$

. . . . . (6)

V_d – zbiór $\{\Delta x_{V_d}\}$ błędów zależ. od $\Delta x_{V_d}(v, x, t_e, e)$.

Poszczególne zbiory (podzbiory zdarzeń elementarnych) $V_1, V_2 \ldots V_d$ charakteryzują się następującymi rozkładami

$$p(\Delta x_{V_1}), p(\Delta x_{V_2}), \ldots, p(\Delta x_{V_d}). \tag{6a}$$

Pomiar pośredni $y = f(x_1, x_2 \ldots, x_n)$ ogranicza się przynajmniej do kilku wielkości mierzonych n. Korzysta się wówczas z kilku podzbiorów zdarzeń elementarnych typu V_k których rozkłady oznacza się

$$p_1(\Delta x_{1,V_k}), p_2(\Delta x_{2,V_k}), \ldots, p_n(\Delta x_{n,V_k}) \tag{6b}$$

gdzie; w symbolu V_k, symbol k przyjmuje nie kolejne ale jedną z możliwych wartości

$$k = 1, 2, \ldots d.$$

Rozkład błędów pomiaru pośredniego jest splotem rozkładów 6b

$$p(\Delta y) = p_1(\Delta x_{1V_k}) * p_2(\Delta x_{2,V_k}) * \ldots * p_n(\Delta x_{n,V_k}). \tag{7}$$

Parametry rozkładu błędów wyniku pomiaru $p(\Delta y)$ są funkcjami parametrów rozkładów splatanych.

Najprościej jest zbadać rozkłady błędów rodzin egzeplarzy wielu typów przyrządów i sporządzić rozkłady gęstości prawdopodobieństwa z parametrów rozkładów błędów poszczególnych typów
gdzie:

$$p(q); p(r) \tag{8}$$

$$r = \frac{s}{|G|}; \; q = \frac{\Delta x_{sr}}{|G|} = \frac{c}{|G|} \tag{9}$$

Δx_{sr} — wartość średnia błędów w rodzinie przyrządów tego samego typu ($c = \Delta x_{sr}$),
s — odchylenie średnie kwadratowe z rozkładów błędów przyrządów tego samego typu,
G — błąd graniczny określony klasą.

Rozkłady $p(q)$, $p(r)$ charakteryzują błędy przestrzeni zdarzeń elementarnych V_Ω tzn. charakteryzują dokładność narzędzi obecnie eksploatowanych.
Podzbiory V_k wskazanym jest podzielić na mniejsze zbiory, zbiory błędów w rodzinach przyrządów poszczególnych typów, oznaczając je symbolem V_{kl}, gdzie; k — oznacza mierzoną wielkość, a l — typ przyrządu odpowiednio oznakowany.

Otrzymamy zatem dla każdej wielkości mierzonej następujące podzbiory:

$$V_{k1}, V_{k2}, \ldots V_{kh} \tag{10}$$

$$\text{dla } k = 1, 2, \ldots d.$$

Podział V na zbiory V_{kl} można zapisać w postaci

$$\begin{aligned} &V_{k1} \text{ — zbiór bł. } \{\Delta x_{V_{k1}}\} \text{ zależnych od } \Delta x_{V_{k1}}(x, t_e, e) \\ &V_{k2} \text{ — zbiór bł. } \{\Delta x_{V_{k2}}\} \text{ zależnych od } \Delta x_{V_{k2}}(x, t_e, e) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &V_{kh} \text{ — zbiór bł. } \{\Delta x_{V_{kh}}\} \text{ zależnych od } \Delta x_{V_{kh}}(x, t_e, e). \end{aligned} \tag{11}$$

Populacje V_{kl}, których liczba równa jest liczbie eksploatowanych typów kh, przedstawia macierz.

$$\begin{aligned} &V_{11}\ V_{12} \ldots V_{1h} \\ &V_{21}\ V_{22} \ldots V_{2h} \\ &\ldots\ldots\ldots\ldots \\ &V_{d1}\ V_{d2} \ldots V_{dh}. \end{aligned} \tag{12}$$

Można zatem sporządzić rozkłady

$$p(\Delta x_{V_{kl}}) \tag{13}$$

$$\text{dla } k = 1, 2, \ldots, d;\ l = 1, 2, \ldots h$$

Zbiorami (11), (12) i rozkładami (13) należy posługiwać się przy obliczaniu błędów pomiarów pośrednich kiedy losowo dobieramy egzemplarze przyrządów (lub przetworniki systemu pomiarowego).

Podział przestrzeni V_Ω na coraz mniejsze podzbiory zestępujące, należy dalej kontynuować. Podzbiory błędów V_{kl} rodzin egzemplarzy tego samego typu można podzielić na mniejsze zbiory, zawierające tylko błędy związane z poszczególnymi egzemplarzami przyrządów, które opisują wyrażenia:

$$\begin{aligned} &V_{kl1} \text{ —zbiór bł. } \{\Delta x_{V_{kl1}}\} \text{ zależnych od } \Delta x_{V_{kl1}}(x, t_e) \\ &V_{kl2} \text{ — zbiór bł. } \{\Delta x_{V_{kl2}}\} \text{ zależnych od } \Delta x_{V_{kl2}}(x, t_e) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned} \tag{14}$$

$$V_{klw} \text{ – zbiór bł. } \{\Delta x_{V_{klw}}\} \text{ zależnych od } \Delta x_{V_{klw}}(x, t_e), \qquad (14 \text{ c. d.})$$

gdzie: $k = 1, 2, \ldots, d$; $l = 1, 2, \ldots, h$

Dla wyodrębnionych zbiorów (14) otrzymuje się następujące rozkłady gęstości prawdopodobieństwa:

$$p(\Delta x_{V_{klz}}) \qquad (15)$$

$$\text{dla } k = 1, 2, \ldots, d$$
$$l = 1, 2, \ldots, h$$
$$z = 1, 2, \ldots, w.$$

Są to rozkłady błędów występujących w określonych przyrządach, błędy te zmieniają się według funkcji losowej $\Delta x = f(x)$ dla $t_e = \text{const}$. Kiedy dokonujemy wielu pomiarów pośrednich, całej serii, ciągle tymi samymi przyrządami wówczas zbiór zdarzeń elementarnych pomiarów bezpośrednich ogranicza się do zbiorów błędów w użytych przyrządach (nie zmienionych w układzie), $\{\Delta x_{V_{klz}}\}$.

Analizowana na początku przestrzeń zdarzeń elementarnych V_Ω jest przestrzenią przeliczalną. Do ciała $\mathcal{F}$tej przestrzeni należą rodziny wydzielonych zbiorów odpowiednio dobranych do potrzeb pomiarowych

$$V_\Omega = \{\Delta x_\Omega(v, x, t_e, e)\} \qquad (16)$$

$$V_\Omega = \{V_k\} = \{V_1, V_2, \ldots, V_d\} \qquad (17)$$

$$V_\Omega = \{V_{kl}\} = \{V_{k1}, V_{k2}, \ldots, V_{kh}\} \qquad (18)$$

$$\text{dla } k = 1, 2, \ldots d$$

$$V_\Omega = \{V_{klz}\} = \{V_{k1}, V_{k2}, \ldots, V_{klw}\} \qquad (19)$$

$$\text{dla } k = 1, 2, \ldots d;\ l = 1, 2, \ldots h.$$

Opiszmy je dokładniej:

V_Ω – zbiór błędów podstawowych (instrumentalnych) współcześnie eksploatowanych przyrządów, wszystkich typów i egzemplarzy – przestrzeń, zdarzeń elementarnych,

V_k – zbiór błedów typów i egzemplarzy przyrządów służących do pomiaru wielkości k (np. napięcia elektrycznego)

V_{kl} – zbiór błędów rodziny przyrządów jednego typu, oznaczonego nr V_{kl} (np. woltomierzy Mu200)

V_{klz} – zbiór błedów występujących w przyrządzie pomiarowym, określonym egzemplarzu oznaczonym symbolem *klz*.

Wyodrębnione zbiory zdarzeń elementarnych charakteryzują rozkłady gęstości prawdopodobieństwa

$$p(\Delta x_\Omega) \qquad (20)$$

$$p(\Delta x_{Vk}) \qquad (21)$$

$$p(\Delta x_{V_{kl}}) \qquad (22)$$

$$p(\Delta x_{Vklz}) \qquad (23)$$

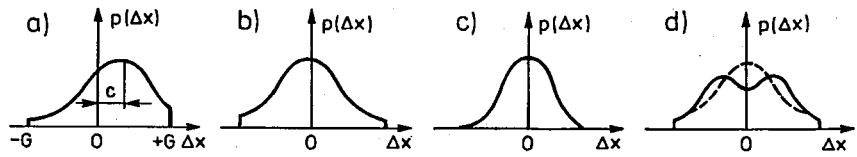
4. ZAŁOŻONE GRANICZNE KSZTAŁTY ROZKŁADÓW BŁĘDÓW WYDZIELONYCH ZBIORÓW ZDARZEŃ ELEMENTARNYCH.

Graniczne rozkłady, ze względu na poziomy ufności sumowania błędów cząstkowych pomiaru pośredniego, wykazują następujące własności:

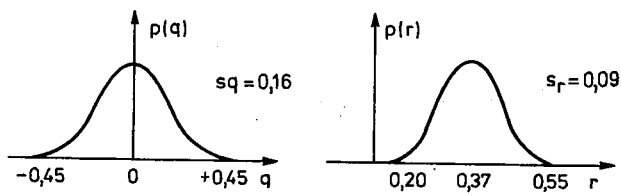
a) Przestrzeń zdarzeń elementarnych V_Ω, charakteryzują rozkłady $p(q)$ i $p(r)$ parametrów z rozkładów błędów populacji typów przyrządów współcześnie eksploatowanych [1, 5]. Wynika z nich, że rozkłady $p(\Delta x_\Omega)$ są normalne i nieprzesunięte co potwierdzają badania eksperymentalne [1, 5].

b) Rozkłady z poszczególnych zbiorów zdarzeń elementarnych V_k, oznaczone symbolem $p(\Delta x_{V_k})$, są również normalne nieprzesunięte, co wynika z badań [5, 1].

c) Rozkłady zbiorów V_{kl}, $p(\Delta x_{V_{kl}})$ — rodzin przyrządów poszczególnych typów są normalne odcięte i przesunięte (z nielicznymi wyjątkami rozkładów dwumodolnych i splotów jednostajnego z normalnym). Przykłady takich rozkładów przytoczono na rys. 1, a rozkłady ich asymetrii q i kształtów r pokazano na rys. 2.



Rys. 1. Przykłady rozkładów $p(\Delta x_{vkl})$ — rodzin przyrządów czterech typów. a, b, c — rozkłady typowe, powszechnie występujące, d — przykład rozkładu nietypowego, który może być aproskymowany rozkładem typowym (normalnym) co nie wpłynie istotnie na rozkład splotu kilku różnych rozkładów

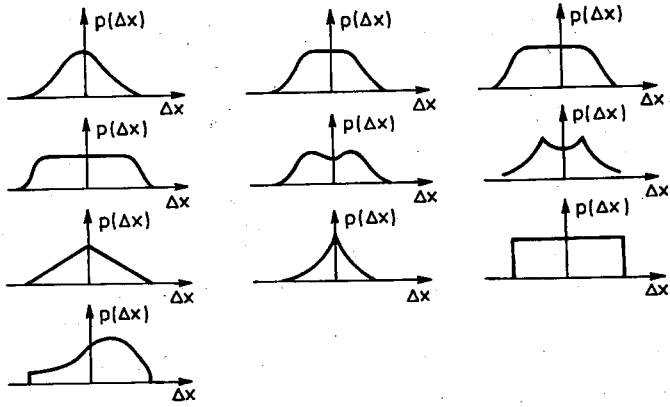


Rys. 2. Rozkład parametru asymetrii $p(q)$ i kształtu $p(r)$ populacji typów przyrządów. Charakterystyka błędów przestrzeni zdarzeń elementarnych V_Ω [5,1,8]

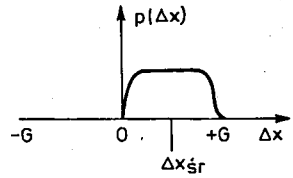
d) Rozkłady błędów przyrządów, pojedyńczych egzemplarzy, ze zbiorów $V_{klz} = \{\Delta x_{V_{klz}}\}$, które zbadał i opisał P. W. Novicki [6] przyjmują najczęściej kształty przedstawione na rys. 3.

Na podstawie badań rozkładów błędów wielu przyrządów $p(\Delta x_{V_{klz}})$ [1, 5, 6], a także rozkładów rodzin przyrządów $p(\Delta x_{V_{kl}})$ można za graniczny niekorzystny (dla sumowania błędów w pomiarze pośrednim) przyjąć rozkład pokazany na rys. 4, o asymetrii wynoszącej $\Delta x_{śr} = \frac{1}{2} G$ i kształcie zbliżonym do jednostajnego, rys. 4.

Dla większej od dwóch liczby zmiennych pomiaru pośredniego, asymetria średnia — $q_{śr}$ i asymetria splotu rozkładów zanika i poziomy ufności sumowania błędów będą



Rys. 3. Rozkłady błędów przyrządów, poszczególnych egzemplarzy, ze zbiorów $V_{klz}=\{\Delta x_{vklz}\}$ [6,5,8]



Rys. 4. Założony graniczny rozkład błędów narzędzi pomiarowych, dla pomiaru pośredniego dwoma przyrządami, $y=f(x_1, x_2)$. Dotyczy zbiorów $V_{klz} = \{\Delta x_{Vklz}\}$

korzystniejsze – wzrastają ze wzrostem n. Stąd ufności dla rozkładów przyjętych na rys. 4 i dwóch zmiennych należy uznać za najmniej korzystne i w oparciu o ten przypadek dobierać wyrażenia na błędy graniczne.

5. ZASADY DOBORU POZIOMÓW UFNOŚCI I WYRAŻEŃ NA BŁĘDY GRANICZNE

Obiektem zainteresowania jest składanie błędów:

a) w układach przetworników,

b) w pomiarach pośrednich.

Zakłada się, że błędy graniczne przetworników, wynikające z ich klasy są znane. Funkcją wiążącą zależności między wielkościami, np. mierzonymi bezpośrednio i pośrednio jest

$$y = f(x_1, x_2, \ldots, x_n) \tag{24}$$

Błąd oblicza się za pomocą wyrażenia

$$\Delta y = \frac{\partial y}{\partial x_1}\Delta x_1 + \frac{\partial y}{\partial x_2}\Delta x_2 + \ldots + \frac{\partial y}{\partial x_n}\Delta x_n \tag{25}$$

Aby dokładnie obliczyć Δy należy znać rozkłady błędów instrumentalnych, podstawowych:

$$p(\Delta x_1), p(\Delta x_2), \ldots, p(\Delta x_n). \tag{26}$$

Przy czym rozkłady błędów narzędziowych będą zależały od tego do jakiego zbioru wyników pomiarów (zbioru zdarzeń elementarnych) odnosimy nasze obliczenia. Dokonując pomiarów w tym samym układzie może zachodzić konieczność posłużenia się różnymi rozkładami, dla wielu pomiarów — serii — rozkładami $p(\Delta x_{V_{kn}})$, a dla jednego pomiaru w tym samym układzie, jeżeli przyrządy dobierano losowo — rozkładem $p(\Delta x_{V_u})$.

Rozkład błędów wyniku pomiaru pośredniego $p(\Delta y)$ jest splotem rozkładów błędów cząstkowych (26) sumowanych

$$p(\Delta y) = p(\Delta x_1) * p(\Delta x_2) * \ldots * p(\Delta x_n). \tag{27}$$

Znając rozkład p(Δy) można wyznaczyć graniczne błędy $\pm G_y$ dla złożonych wymaganych poziomów ufności $P(G_y)$ z warunku

$$\int_{-G_y}^{G_y} p(\Delta y)\mathrm{d}\Delta y \geqslant P(G_y). \tag{28}$$

Zagadnienie sprowadza się do doboru wyrażeń na błędy graniczne $\pm G_y$, które byłyby zachowane na wymaganym poziomie ufności 95% lub 99%. Wyrażenia na G_y zależą także od kształtów rozkładów $p(\Delta x_i)$, $i = 1, 2, \ldots, n$.

Jak wiadomo z przedstawionymi w tym artykule zbiorami zdarzeń elementarnych, wiążą się określone rozkłady. Rozpatrzmy zatym, po kolei, ogólne zasady obliczania poziomów ufności i dobierania wyrażenia na G_y w zależności od rodzaju pomiaru i odpowiadającej mu przestrzeni zdarzeń.

a) Pomiary pośrednie w układach do których typy przyrządów dobiera się losowo (pod względem parametrów ich rozkładów) i potem w drugim etapie losuje się egzemplarze narzędzi pod względem wartości błędów leżących w przedziale $G > \Delta x > -G$. Wynikami tego typu pomiarów—doświadczeń jest zbiór zdarzeń elementarnych V_Ω lub suma kilku zbiorów V_k, które charakteryzują się normalnymi i nieprzesuniętymi rozkładami, symetrycznymi względem układu — GOG.

Rozkład błędów wyniku takiego pomiaru jest splotem

$$p_\Omega(\Delta y) = p_1(\Delta x_{1,V_k}) * p_2(\Delta x_{2,V_k}) * \ldots * p_n(\Delta x_{n,V_k}) \tag{29}$$

gdzie: $p_1(\Delta x_{1,V_k}), \ldots, p_n(\Delta x_{n,V_k})$ są rozkładami błędów cząstkowych pomiarów bezpośrednich wielkości $x_1, x_2, \ldots, x_n$.
Poziom ufności oblicza się z warunku

$$\int_{-G_{y_\Omega}}^{G_{y_\Omega}} p_\Omega(\Delta y)\,\mathrm{d}\,\Delta y \geqslant P(G_{y_\Omega}) \tag{30}$$

Z warunku (30) oblicza się również wyrażenie na błąd graniczny G_{y_Ω} o ile założymy wymagany poziom ufności $P(G_{y_\Omega})$, np. 95 lub 99%.

b) Pomiary w układach do których egzemplarze przyrządów dobiera się losowo, natomiast typy przyrządów pozostają stałe dla całej serii pomiarów. Zdarzeniom tym odpowiadają zbiory wyników doświadczenia V_{kl}, dla których rozkłady są normalne obcięte i przesunięte (opisane w rozdz. 4 w punkcie c.). Rozkład błędów wyniku pomiaru pośredniego (lub błąd układu przetworników) jest splotem rozkładów rodzin typów narzędzi (normalnych obciętych i przesuniętych).

$$p_{V_{kl}}(\Delta y) = p_1(\Delta x_{1,V_{kl}}) * p_2(\Delta x_{2,V_{kl}}) * \ldots * p_n(\Delta x_{n,V_{kl}}) \tag{31}$$

Poziom ufności lub wyrażenia na błąd graniczny $\pm G_{y,V_{kl}}$ oblicza się z warunku

$$\int_{-G_{y,V_{kl}}}^{G_{y,V_{kl}}} p_{V_{kl}}(\Delta y)\mathrm{d}\Delta y \geqslant P(Gy, V_{kl}) \tag{32}$$

c) Pomiary w stałych układach, stale tymi samymi przyrządami. Wtedy nie występuje zjawisko losowego doboru typu ani też egzemplarza przyrządu, w przypadku kiedy obliczania odnosimy do serii wielu pomiarów w jednym układzie. Zbiory zdarzeń elementarnych ograniczają się wówczas do zbiorów błędów w poszczególnych przyrządach V_{klz}.

Rozkłady tych zdarzeń $p(\Delta x_{V_{klz}})$ przyjmują różne kształty: zbliżone do jednostajnych, normalne, dwumodalne, trójkątne i inne [6, 5]. Są one różne, ale posiadają pewne cechy wspólne i można je aproksymować kilkoma typowymi kształtami.

Rozkład błędów wyniku pomiaru pośredniego $p_{V_{klz}}(\Delta y)$ jest zatem splotem różnych rozkładów błędów cząstkowych.

$$p_{V_{klz}}(\Delta y) = p_1(\Delta x_{1,V_{klz}}) * \ldots * p_n(\Delta x_{n,V_{klz}}). \tag{33}$$

Dla przyjętego wyrażenia na błędy graniczne $G_{y,V_{klz}}$, poziom ufności oblicza się z warunku

$$\int_{-G_{y,V_{klz}}}^{G_{y,V_{klz}}} p_{V_{klz}}(\Delta y)\ \mathrm{d}\Delta y \geqslant p(G_{y,V_{klz}}), \tag{34}$$

lub odwrotnie, zakładając pożądaną ufność można dobrać wyrażenie na błąd graniczny pomiaru pośredniego lub układu przetworników.

Rozkłady $p(\Delta x_{V_{klz}})$ przyjmują różne kształty (rys. 3) w obrębie kilku kształtów, którymi można je aproksymować. Stąd dla tego typu pomiarów wygodnie jest założyć graniczny rozkład (ze względu na ufność) jednakowy dla dwóch przyrządów w pomiarze pośrednim $y = f(x_1, x_2)$. Z badań eksperymentalnych wynika, że może nim być rozkład bardzo zbliżony do jednostajnego i położony asymetrycznie (rys. 4). Otrzymuje się wówczas większą wartość błędów, ale uwalnia się obliczającego od konieczności pozyskiwania informacji co do kształtów rozkładów $p(\Delta x_{V_{klz}})$.

W przedstawionej analizie założono że w przypadkach użytych do pomiaru, uważanych za sprawne, nie występuje przekroczenie klasy, tzn. że:

$$\begin{aligned}
&-G_1 \leqslant \Delta x_1 \leqslant +G_1 \text{ z prawdopodob. } p = 1\\
&-G_2 \leqslant \Delta x_2 \leqslant +G_2 \text{ z prawdopodob. } p = 1\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&-G_n \leqslant \Delta x_n \leqslant G_n \quad \text{z prawdopodob. } p = 1
\end{aligned} \tag{35}$$

W praktyce przekroczenia klasy występują i często sięgają wartości 5% przyrządów uważanych za sprawne. Ze względu na wpływ operacji splotu przy sumowaniu błędów, ich niekorzystny skutek można pominąć [5]. Np. w pomiarach $y = f(x_1, x_2, x_3)$ dla trzech zmiennych, o założonych przekroczeniach klasy wynoszących w każdym przyrządzie po 5%, operacja splotu spowoduje, że przekroczenie błędu pomiaru pośredniego, wyniesie zaledwie $0{,}2 \div 0{,}1\%$ [5].

Warto zuważyć, że poziom ufności jest wartością względną w zależności od tego do jakiego zbioru zdarzeń elementarnych (wyników doświadczeń) odnosimy obliczenia G_y. Np. poziom ufności błędów granicznych pomiaru pośredniego w układzie zmontowanym do jednego pomiaru (do którego przyrządy dobrano losowo), określa się na zbiorze zdarzeń elementarnych $\{\Delta x_{V_{kl}}\}$ i rozkładzie $p(\Delta x_{V_{kl}})$. Natomiast jeżeli tym samym układem wykonamy wiele pomiarów pośrednich (serię) wówczas obliczenia będziemy odnosić do zbiorów zdarzeń $\{\Delta x_{V_{kl}}\}$, które charakteryzują się innymi rozkładami, co zmieni także poziom ufności przy tym samym wyrażeniu na G_y. Zagadnienie względności poziomów ufności błędów granicznych dla pomiarów pośrednich w tym samym układzie autor zamierza opisać w oddzielnej publikacji pt.: „Względność poziomów ufności sumowania błędów pomiarowych".

Przedstawiona w artykule przestrzeń probabilistyczna odnosząca się do sumowania błędów narzędzi pomiarowych umożliwia usystematyzowanie i klarowny opis wielu metod szczegółowych obliczania błędów, przystosowanych do określonych przypadków występujących w technice pomiarowej. Natomiast sformułowane ogólne zasady określania ufności i dobierania wyrażeń na błędy graniczne mogą ułatwić obliczanie błędów w różnych przypadkach szczegółowych.

BIBLIOGRAFIA

1. Z. Orzeszkowski: *Wybrane działy z teorii pomiarów cz. I Politechnika Wrocławska,* 1974
2. J. Piotrowski: *Teoria pomiarów.* Pomiary w fizyce i technice. Warszawa: PWN, 1986
3. J. Jaworski: *Matematyczne podstawy metrologii.* Warszawa: PWN, 1979
4. J. Dudziewicz: *Niepewności wyników pomiarów.* Biuletyn Informacyjny Instytutu Łączności 1990, Nr 11–12
5. R. Rakowski: *Obliczanie niepewności pomiarów statycznych.* WOSR Jelenia Góra 1989
6. P. W. Novicki, I. A. Zograf: *Ocenka pagresznostiej rezultatów izmierenić,* Leningrad: ELO 1985
7. H. Poleck: *Die Ergebnissicherheigt bei quadratischer Addition von Fehlergrenzen.* VDI Zeitschrift 1965, z. 28
8. R. Rakowski: *Szacowanie błędów w pomiarach technicznych.* WOSR Jelenia Góra, 1989

R. RAKOWSKI

PROBABILISTIC SPACE OF ADDING INSTRUMENTAL ERRORS

Summary

Probabilistic space appropriate to statistical assemblage of the partial errors of indirect measurement and systems of converters has been presented. Boundary errors distributions for several essential files of accidents have been defined or assumed on the base of experiments. General procedures of accounting levels of confidence and selecting expressions on boundary errors have been submited.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1992, 38, z. 2., ss. 455–485

(621.382.066)

Synteza neuralnych systemów komórkowych dla zagadnień podejmowania decyzji

Część I: Systemy selekcji — identyfikacji

ANDRZEJ SZATKOWSKI

Instytut Technologii Elektronicznej, Politechnika Gdańska

Otrzymano 1990. 07. 12

Autoryzowano do druku 1991. 05. 10

Komórkowe systemy neuralne są nieliniowymi systemami analogowymi dla informacyjnego przetwarzania sygnałów, realizowanymi jako układy specjalizowane — systemy kontroli i sterowania, selekcji danych, przetwarzania obrazów. Struktura komórkowa dla sieci neuralnych wynika z zastosowania w ich konstrukcji funktorów selekcji. Komórkowe systemy neuralne są układami dynamicznymi pracującymi w czasie rzeczywistym ciągłym lub jako systemy iteracyjne — w czasie dyskretyzowanym. Posiadają przy tym właściwość charakterystyczną dla układów parametrycznych. W części I zaproponowano pewną uniwersalną strukturę komórkowego systemu neuralnego selekcji — identyfikacji z wykorzystaniem funktora 1-wymiarowego. Podano też szereg uwag ogólnych odnośnie systemów dynamicznych selecji — identyfikacji. Przedstawiono szereg wniosków dotyczących implementacji elektronicznej.

1. WSTĘP

W pracy rozważana jest klasa nieliniowych systemów analogowych, określanych w części I jako systemy dynamiczne selekcji — identyfikacji, a w części II pracy jako systemy mnogościowo — logiczne i ogólniej — dynamiczne systemy komórkowe typu neuralnego. Struktura komórkowa dla rozpatrywanych sieci wynika z zastosowania w ich konstrukcji układów selekcyjnych. W przypadku sieci neuralnej Hopfielda funktorem elementarnym selekcji — identyfikacji jest układ flip–flop [10] — [12], [15]. Jest to układ dynamiczny 2-wymiarowy (wymiar przestrzeni stanu jest równy 2) z dwoma obszarami selekcyjnymi. Proponowany w pracy funktor elementarny jest systemem dynamicznym 1-wymiarowym z obszarem selekcji pozytywnej w postaci odcinka $(-\varepsilon, \varepsilon)$, $\varepsilon > 0$, osi rzeczywistej i obszarem selekcji negatywnej będącym dopełnieniem przedziału $[-\varepsilon, \varepsilon]$. Przy spełnieniu odpowiednich wymagań przez

zewnętrzną sieć funkcyjną tworzącą układ sprzężeń pomiędzy elementami funktorów oraz węzłami wejść i wyjść, rozważane w pracy systemy posiadają podstawową cechę charakterystyczną dla neuralnych sieci komórkowych realizując w stanie ustalonym, dla $t \to \infty$,odwzorowanie continuum stanów początkowych, zadanych przez sygnał zewnętrzny w chwili $t = 0$, w pewien zbiór dyskretny liczb lub ogólniej — symboli. Zbiór osiąganych stanów ustalonych jest przy tym z punktu widzenia dynamiki systemu zbiorem jego asymptotycznie stabilnych stanów równowagi. Z punktu widzenia informacyjnego przetwarzania sygnałów neuralna sięć komórkowa dokonuje zatem konwersji zbiorów danych numerycznych w pewien skończony zbiór symboli. Jest to funkcja analogiczna do przetwarzania analogowo—cyfrowego w systemach cyfrowych i jest w istocie procedurą selekcji informacji, realizowaną na przykład w układach perceptronów [13, [20], [22], [26].

Proponowany w pracy wybór funktora podstawowego w postaci 1-wymiarowego systemu dynamicznego selekcji upraszcza w znacznym stopniu procedurę projektowania sieci neuralnej, czyniąc ją bardziej uniwersalną i przejrzystą. Uzyskuje się prostsze sterowanie rozmiarami obszarów selekcyjnych (co jest istotne w syntezie sieci adaptacyjnych i uczących się) oraz łatwą realizację funktorów wielowymiarowych (o obszarach selekcyjnych wielowymiarowych, dowolnego kształtu).

W koncepcji neuralnej sieci komórkowej przedstawionej w [3], [4],. w opisie definiującym każdą z komórek (w jej odpowiednim równaniu stanu) uwzględniona jest całość oddziaływań komórki z pozostałą częścią sieci. Z założenia, oddziaływanie to jest ograniczone do ustalonego, choć dowolnie dużego, otoczenia każdej z komórek elementarnych. W konsekwencji, przedstawiona w pracach [3], [4] procedura syntezy jest stosunkowo skomplikowana i zalecana raczej w specjalnych zastosowaniach.

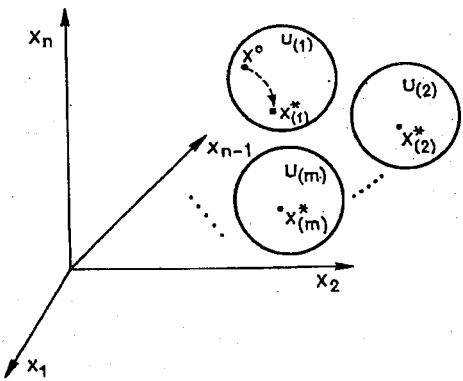
W szeregu procedurach syntezy sieci neuralnych elementarny funktor selekcji — identyfikacji jest rozważany jako element funkcyjny (zob. np. prace [13], [20], [25]). Otrzymuje się wówczas systemy o strukturze na ogół hierarchiczno — szeregowej. Nie ma też przy takim podejściu możliwości rozważania równoległego przetwarzania informacji. Ponadto w wyniku działających zakłóceń stan sieci może ulec zmianie w przypadku, gdy niektóre spośród składowych wektora stanu sieci są położone w bliskim sąsiedztwie punktów przełączania odpowiednich funktorów selekcyjnych. Charakterystyka selektora elementarnego jest tu opisana funkcją schodkową.

Funktor selekcji — identyfikacji reprezentowany przez układ dynamiczny charakteryzuje się natomiast asymptotycznie stabilnymi stanami równowagi. Punkty identyfikujące (będące stanami asymptotycznie stabilnymi) mogą być ulokowane we wnętrzu odpowiednich obszarów selekcyjnych w sposób optymalny, przy uwzględnieniu różnorodnych potrzeb i kryteriów [7], [11], [12], [13], [15], [27]. Ponadto, cała sieć jest analizowana w dziedzinie czasu, nie tylko w kontekście wyznaczania stanów ustalonych, ale również w zakresie analizy stanów przejściowych. Otrzymuje się w wyniku informację odnośnie stanów przejściowych w sieci, co jest istotne przy określaniu czasów ustalania się wartości sygnałów wyjściowych i przy określaniu wartości ekstremalnych przebiegów w stanie nieustalonym. Zmianę parametrów funktorów selekcji można tu powiązać dynamicznie ze stanem sieci w bieżącej chwili czasu, co ma podstawowe znaczenie w syntezie neuralnych systemów uczących się [14].

W szeregu zastosowań (np. przetwarzanie obrazów) analiza stanów przejściowych jest równie istotna jak analiza stanów ustalonych [7]. Dla potrzeb algorytmów neuralnych wykorzystuje się często wartości sygnałów w stanach przejściowych w odpowiednio zsyntetyzowanych sieciach neuralnych, przy zastosowanej dyskretyzacji zmiennej czasowej.

W części I pracy rozważane jest zagadnienie syntezy układu dynamicznego, którego zachowanie się odpowiada realizacji sformułowanego ogólnie problemu selekcji i identyfikacji. Część II pracy, w której omawiane są neuralne systemy komórkowe mnogościowo — logiczne, stanowi rozwinięcie tematu części I. W części II jest też przedstawiony ogólny model matematyczny sieci neuralnej z funktorami 1-wymiarowymi selekcji - identyfikacji jako elementami podstawowymi i dowolną nieliniową siecią funkcyjną.

Zagadnienie selekcji i identyfikacji jest sformułowane ogólnie w następujący sposób. Niech w przestrzeni zmiennych — danych wejściowych (przestrzeni parametrów opisujących pewien realnie istniejący obiekt), posiadającej z założenia strukturę przestrzeni R^n i będącej jednocześnie przestrzenią stanu zsyntetyzowanego systemu dynamicznego, zadana będzie pewna rodzina (spójnych) podzbiorów otwartych $U_{(1)}$, $U_{(2)}, \ldots, U_{(m)}$, parami rozłącznych. W każdym ze zbiorów $U_{(k)}$, $k = 1, 2, \ldots, m$, wyróżniony jest pewien punkt $x^*_{(k)}$. W kontekście rozważanego tu problemu selekcji — identyfikacji, podzbiory $U_{(k)}$ wyznaczają w przestrzeni R^n zbiory punktów równoważnych. Wyróżniony w $U_{(k)}$ punkt $x^*_{(k)}$ identyfikuje elementy należące do $U_{(k)}$ jako równoważne.



Rys. 1. Obszary selekcyjne $U_{(k)}$ i punkty identyfikujące $x^*_{(k)}$ w systemie selekcji — identyfikacji

Obszary $U_{(1)}, \ldots, U_{(m)}$ nazywane są obszarami selekcyjnymi. Rozmiary i kształt każdego z nich zostają ukształtowane w wyniku przeprowadzonego uprzednio procesu uczenia, procesu zdobywania doświadczeń odnośnie danego zagadnienia klasyfikacji — selekcji. Identyfikacja punktów z każdego z obszarów selekcyjnych $U_{(k)}$ z wyróżnionym w $U_{(k)}$ punktem $x^*_{(k)}$ jest realizowana w wyniku granicznej, dla $t \to \infty$, identyfikacji trajektorii odpowiedniego zsyntetyzowanego systemu dynamicznego φ^*,

z punktem $x^*_{(k)}$. To znaczy, dla każdego $x^0 \in U_{(k)}$, dodatnia pół–trajektoria $x(x^0, t)$, $t \geqslant 0$, systemu dynamicznego φ^* o równaniu stanu

$$\dot{x} = f^*(x), \left(\dot{x} = \frac{dx}{dt}\right), \tag{1}$$

gdzie $x = (x_1, \ldots, x_n) \in R^n$ i $f^*(\cdot) : R^n \to R^n$ jest pewną funkcją otrzymaną w wyniku przeprowadzonej procedury syntezy, jest określona jednoznacznie i zmierza asymptomatycznie dla $t \to \infty$ do wyróżnionego w $U_{(k)}$ punktu $x^*_{(k)}$. Każdy z punktów $x^*_{(k)}$, $k = 1, 2, \ldots, m$, jest asymptotycznie stabilnym punktem równowagi zsyntetyzowanego systemu dynamicznego (1), z obszarem przyciągania równym $U_{(k)}$.

Tak więc w zależności od położenia punktu początkowego x^0, którego współrzędne zawierają dane — informację o pewnym realnie istniejącym (bądź zaistniałym) obiekcie, układ dynamiczny osiąga asymptotycznie, dla $t \to \infty$, odpowiedni wyróżniony stan $x^*_{(k)}$, jeżeli $x^0 \in U_{(k)}$.

Tak określony system dokonuje selekcji i identyfikacji jednostopniowej i rozłącznej (obszary selekcyjne są podzbiorami rozłącznymi). Wielostopniowy system selekcji — identyfikacji otrzymuje się w wyniku dalszego wyróżnienia w każdym z obszarów $U_{(K)}$ pewnej rodziny podzbiorów rozłącznych $U_{(K,1)} \ldots\ldots U_{(K,1_K)}$ i następnie kolejne powtórzenie tej procedury określoną ilość razy w odniesieniu do podobszarów selekcyjnych kolejnego wyższego stopnia selekcji.

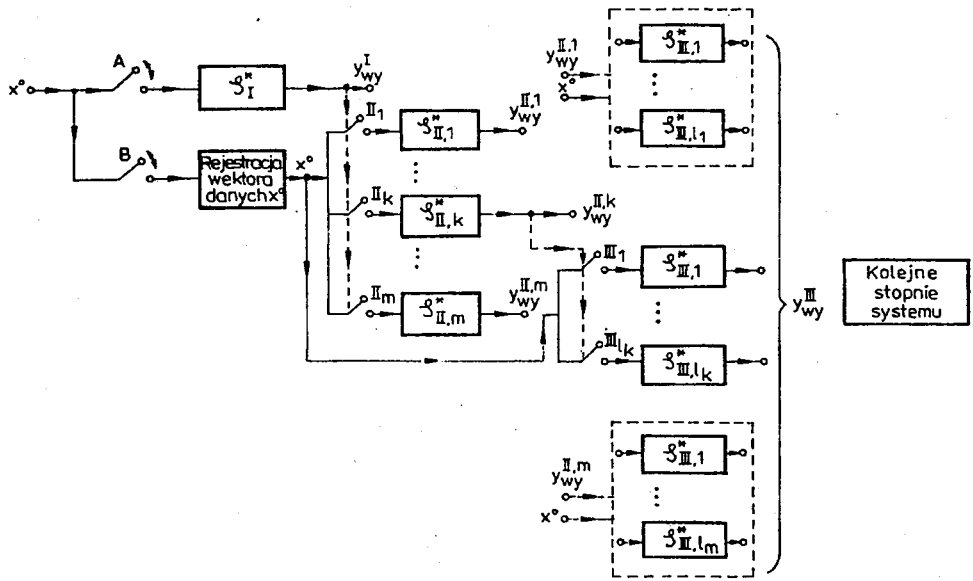
Zagadnienie syntezy systemów wielostopniowych pozostaje poza zakresem pracy, natomiast komórkowe systemy neuralne omawiane w części II odnoszą się do problemu selekcji — identyfikacji nie koniecznie rozłącznej. Ogólną strukturę systemu wielostopniowego przedstawia Rys. 2.

W rozdziale 2 podano ogólną formułę matematyczną dla realizacji syntezy systemu dynamicznego selekcji — identyfikacji. Równanie stanu systemu ma postać równania różniczkowego zwyczajnego (1), i jest równaniem typu gradientowego. Przedstawiono sposób konstrukcji odpowiedniego systemu (1) oraz podano twierdzenie charakteryzujące przebieg trajektorii.

Rozważania dotyczące syntezy systemu dynamicznego selekcji — identyfikacji są prowadzone z wykorzystaniem koncepcji skojarzonej z systemem przestrzeni energo — fazowej. W pracy, a szczególnie w rozdziale 2, wykorzystywane są metody dynamiki geometrycznej [1], [24], to znaczy geometrycznej teorii systemów dynamicznych. Badanie właściwości systemu dynamicznego sprowadza się tu do badania skojarzonej z systemem hiperpowierzchni energii zanurzonej w przestrzeni energo — fazowej systemu. Brzeg każdego z obszarów selekcyjnych jest wyznaczony przez rzut odpowiedniej hiperpowierzchni izoenergetycznej, wyznaczonej w przestrzeni energo — fazowej systemu jako przekrój hiperpowierzchni energii systemu hiperpłaszczyzną stałej energii, przy ustalonej odpowiedniej wartości energii, na przestrzeń stanu systemu. Punkty identyfikujące $x^*_{(1)}, \ldots, x^*_{(m)}$ są punktami krytycznymi funkcji energii (to znaczy punktami, w których gradient funkcji energii jest wektorem zerowym). Każdy z punktów identyfikujących jest położony we wnętrzu odpowiedniego obszaru selekcyjnego i jest z założenia asymptotycznie stabilnym punktem równowagi systemu dynamicznego selekcji — identyfikacji, z obszarem selekcyjnym zadanym w konkretnym problemie syntezy.

Rozważania ogólne są kontynuowane w rozdziale 3 pracy, gdzie zaproponowano praktyczną procedurę realizacyjną zadania syntezy, w tym z uwzględnieniem możliwości implementacji w technice układów elektronicznych. Obszary selekcyjne są aproksymowane rodzinami n-wymiarowych obszarów prostopadłościennych, sumą mnogościową lub iloczynem mnogościowym odpowiednio dobranych tego rodzaju obszarów w R^n. Przedstawiony schemat systemu dynamicznego selekcji — identyfikacji jest uniwersalny i cechuje się prostą strukturą. Inną wersję schematu syntezy podano w rozdziale 2, cz. II pracy, dla przypadku, gdy obszary selekcyjne zadane są za pośrednictwem układu więzów nierównościowych.

Podstawowym elementem realizacyjnym syntezy systemu według proponowanych schematów jest 1-wymiarowy funktor selekcji — identyfikacji. Podano uniwersalną strukturę neuralnego systemu komórkowego selekcji — identyfikacji z zastosowaniem funktora 1-wymiarowego. Natomiast przedstawiony model układu elektronicznego realizującego zadanie funktora 1-wymiarowego czyni procedurę syntezy atrakcyjną z praktycznego punktu widzenia (odnośnie tego tematu zob. też [8], [18], [19]). Istotna jest tu między innymi możliwość uwzględnienia zmiany rozmiarów obszarów selekcyjnych, co określa adaptacyjność systemu.



Rys. 2. Struktura wielopoziomowego systemu selekcji — identyfikacji. W chwili początkowej stan węzłów wyjściowych jest ustalony jako zerowy. Stopień pierwszy φ_I^* jest układem selekcji — identyfikacji opisanym w tekście. Przyjęto, że dla $x^0 \in U_{(k)}$, $y^I_{wy}(t \to \infty) = k$ oraz $y^I_{wy}(t) = 0$ dla każdego t, jeżeli $x^0 \notin \bigcup_{k=1}^{m} \bar{U}_{(k)}$. Gdy $y^I_{wy}(t) \simeq k$, następuje zamknięcie klucza II_k na przedział czasu niezbędny dla przekazania wartości składowych wektora danych x^0 do podukładu φ^*_{II} drugiego stopnia selekcji — identyfikacji. Ma to miejsce tylko wtedy, gdy $x^0 \in U_{(k)}$. Jeżeli $y^I_{wy}(t) = 0$ dla każdego t, klucze pozostają otwarte. Opis funkcjonowania kolejnych stopni systemu jest analogiczny

Systemy — mnogościowo logiczne są uogólnieniem systemów selekcji — identyfikacji. Koncepcja systemu dynamicznego mnogościowo — logicznego st następująca. Dana jest rodzina $U_{(1)}, \ldots, U_{(m)}$ podzbiorów w R^n oraz pewien zbiór (sekwencja) operacji mnogościowych, które przypisują rodzinie $U_{(1)}, \ldots, U_{(m)}$ pewien podzbiór U w R^n (np. $U: = \bigcup_{k=1}^{m} U_{(k)}$. Dla zadanego warunku początkowego $x^0 \in R^n$ reprezentującego zbiór, wektor danych wejściowych) system dynamiczny mnogościowo—logiczny zsyntetyzowany dla zadanej rodziny podzbiorów $U_{(k)}$ i sekwencji operacji mnogościowych na tych podzbiorach osiąga asymptotycznie stan wyjścia „1", jeżeli $x^0 \in U$ i stan „0", gdy $x^0 \notin \bar{U}$. W realizacji praktycznej istotna jest przy tym adaptacyjność systemu wyrażająca się w możliwości uwzględniania zmiany rozmiarów obszarów $U_{(1)}, \ldots, U_{(m)}$.

System dynamiczny mnogościowo—logiczny stanowi więc realizację systemową pewnego określonego funktora mnogościowo—logicznego. W rozdziale 3, cz. II pracy, zaproponowany został ogólny model matematyczny neuralnej sieci komórkowej. Model ten obejmuje swoim zakresem wszystkie rozważane uprzednio w pracy systemy.

2. SYNTEZA SYSTEMU DYNAMICZNEGO SELEKCJI — IDENTYFIKACJI

W przestrzeni R^n (przestrzeni zmiennych — danych wejściowych, będącej przestrzenią parametrów odnoszących się do pewnej klasy realnie istniejących obiektów) zadana jest rodzina $U_{(1)}, \ldots, U_{(m)}$ podzbiorów otwartych (niepustych) i spójnych.

Podzbiory $U_{(k)}$, $k = 1, 2, \ldots, m$, są tu interpretowane jako obszary selekcyjne dla pewnego problemu selekcji — identyfikacji. Zakłada się, że są one ograniczone oraz dla każdej pary $U_{(k)}$, $U_{(l)}$, $k \neq l$, zachodzi

$$\bar{U}_{(k)} \cap \bar{U}_{(l)} = \emptyset, \tag{2}$$

gdzie $\bar{U}$ oznacza domknięcie podzbioru U w R^n.

Niech brzeg $\Gamma_{(k)}$ każdego z obszarów $U_{(k)}$ będzie zwartą [6] $(n-1)$ -wymiarową hiperpowierzchnią regularną w R^n [6] zadaną równaniem

$$h_{(k)}(x) = 0, \tag{3}$$

gdzie $h_{(k)}(\cdot) \in C^2(R^n, R)$, oraz

$$\nabla h_{(k)}(x) \neq 0 \text{ dla } x \in \Gamma_{(k)}. \tag{4}$$

Bez zmniejszania ogólności zakłada się też, że

$$U_{(k)} = \{x \in R^n : h_{(k)}(x) < 0\}. \tag{5}$$

Przy założeniu (2), dla każdego $U_{(k)}$, $k = 1, 2, \ldots, m$, istnieje podzbiór otwarty i ograniczony $V_{(k)}$ w R^n zawierający $\bar{U}_{(k)}$ i taki, że

$$U_{(l)} \subset R^n \setminus \bar{V}_k \text{ dla każdego } l \neq k. \tag{6}$$

Podzbiory $V_{(k)}$, $k = 1, 2, \ldots, m$, można ustalić tak, aby

$$\bar{V}_{(k)} \cap \bar{V}_{(l)} = \emptyset, \text{ dla } k \neq l. \tag{7}$$

Istnieją też funkcje $\lambda_{(k)}(\cdot) : R^n \to R$, $k = 1, 2, \ldots, m$, klasy C^2 takie, że [2], [6]

$$\lambda_{(k)}(x) = 1, \text{ dla } x \in U_{(k)}, \tag{8}$$

oraz

$$\lambda_{(k)}(x) = 0, \text{ dla } x \in R^n \setminus V_{(k)}. \tag{9}$$

Niech funkcja $h(\cdot) : R^n \to R$ będzie dana wzorem

$$h(x) = \sum_{k=1}^{m} \lambda_{(k)}(x) \cdot h_{(k)}(x) \tag{10}$$

i niech φ^* będzie różniczkowym systemem dynamicznym mającym równanie stanu postaci

$$\dot{x} = \alpha(x) \cdot h(x) \cdot \nabla h(x), \tag{11}$$

gdzie $\alpha(\cdot) \in C^1(R^n, R)$ oraz $\alpha(x) > 0$ dla $x \in \bigcup_{k=1}^{m} \bar{U}_{(k)}$.

System dynamiczny (11) realizujący, jak to zostanie wykazane w Twierdzeniu 1, zadanie syntezy systemu selekcji dla zadanych obszarów $U_{(1)}, \ldots, U_{(m)}$, można zapisać równoważnie w postaci

$$\dot{x} = \frac{1}{2} \alpha(x) \cdot \nabla h^2(x). \tag{12}$$

Zawężenie systemu dynamicznego (11), (12) do dowolnego z obszarów $U_{(k)}$ daje w wyniku system gradientowy

$$\dot{x} = \frac{1}{2} \nabla h^2_{(k)}(x).$$

Można zatem w sposób naturalny zinterpretować funkcję $h(\cdot)$ jako funkcję energii skojarzoną z systemem (11), (12). Brzeg $\Gamma_{(k)}$ każdego z obszarów selekcyjnych $U_{(k)}$ jest, przy takiej interpretacji, $(n-1)$-wymiarową hiperpowierzchnią w przestrzeni stanu R^n systemu, zawartą w podzbiorze izoenergetycznym

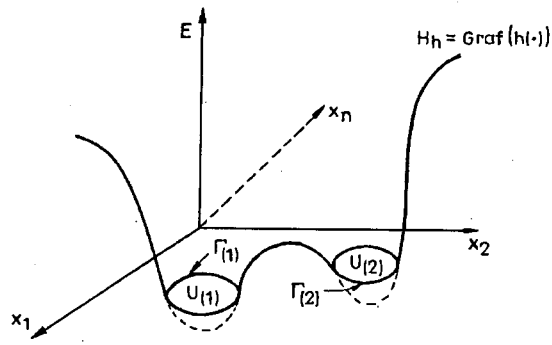
$$\{x \in R^n : h(x) = 0\}. \tag{13}$$

Podzbiór izoenergetyczny (13), odpowiadający stałej wartości $E = 0$ energii, jest wyznaczony jako rzut na przestrzeń stanu przekroju hiperpowierzchni energii

$$H_h = \text{Graf}(h(\cdot)),$$

$\text{Graf}(h(\cdot)) = \bigcup_{x \in R^n} (x, h(x))$, odpowiadającej funkcji energii $h(\cdot)$ systemu, z n-wymiarową hiperpłaszczyzną $E = 0$ w R^{n+1}, którą możemy tu identyfikować z przestrzenią stanu R^n systemu φ^* (11), (12).

Zachodzi następujące twierdzenie charakteryzujące zachowanie się trajektorii różniczkowego systemu dynamicznego φ^* danego równaniami (11), (12).



Rys. 3. Przestrzeń energo-fazowa $R^n \times R$ systemu φ^* (11), (12) i wykreślona w niej hiperpowierzchnia energii H_h odpowiadająca pewnej funkcji energii $h(\cdot)$ jak w (10). Symbolem E oznaczono współrzędną energii

Twierdzenie 1

Dla każdego warunku początkowego $x^0 \in R^n$ rozwiązanie $x = x(x^0, t)$, $t \geqslant 0$, systemu (11), (12) φ^* istnieje i jest określone jednoznacznie, oraz zmierza asymptotycznie dla $t \to \infty$ do zbioru złożonego z punktów równowagi systemu.Dla każdego $k = 1, 2, \ldots, m$ i warunku początkowego $x^0 \in U_{(k)}$ dodatnia półtrajektoria $x(x^0, t)$, $t \geqslant 0$, systemu φ^* jest zawarta całkowicie w $U_{(k)}$ i zmierza asymptotycznie, dla $t \to \infty$, do zbioru punktów równowagi systemu położonych w $U_{(k)}$. Punkty równowagi systemu φ^* (11), (12) położone w obszarze $U_{(k)}$ są dokładnie punktami krytycznymi funkcji $h(\cdot)$ w $U_{(k)}$ (to znaczy punktami x^*, w których $\nabla h(x^*) = 0$).

Ponadto, dla $x^0 \notin \bigcup_{k=1}^{m} \bar{U}_{(k)}$ trajektoria $x(x^0, t)$, $t \geqslant 0$, pozostaje cały czas w obszarze dopełnienia sumy (mnogościowej) obszarów $\bar{U}_{(k)}$ zmierzając asymptotycznie, dla $t \to \infty$, do zbioru złożonego z punktów równowagi systemu φ^* położonych w $R^n \setminus \bigcup_{k=1}^{m} \bar{U}_{(k)}$.

Następnie, brzeg $\Gamma_{(k)}$ każdego z obszarów $U_{(k)}$ jest zbiorem złożonym z punktów równowagi systemu i dla każdego dowolnie małego zaburzenia $\tilde{x}^0$ położenia początkowego $x^0 = x(t=0) \in \Gamma_{(k)}$, które nie zachodzi stycznie do $\Gamma_{(k)}$, rozwiązanie $x = x(\tilde{x}^0, t)$ zmierza asymptotycznie do zbioru punktów równowagi systemu położonego w $U_{(k)}$ lub w obszarze dopełnienia $R^n \setminus \bigcup_{k=1}^{m} \bar{U}_{(k)}$. □

Dowód Twierdzenia 1 jest podany w Dodatku. Poniżej przedstawione zostaną natomiast wnioski i uwagi związane z jego treścią.Jeżeli każda z funkcji $h_{(k)}(\cdot)$ ma w odpowiednim obszarze $U_{(k)}$ dokładnie jeden punkt krytyczny, oznaczony tu przez $x^*_{(k)}$, $\nabla h_{(k)}(x^*_{(k)}) = 0$, to system różniczkowy (11), (12) z funkcją energii $h(\cdot)$ jak we wzorze (10) realizuje zadanie syntezy systemu dynamicznego selekcji — identyfikacji z obszarami selekcyjnymi $U_{(1)}, \ldots, U_{(m)}$ i punktami identyfikującymi zlokalizowanymi w $x^*_{(1)}, \ldots, x^*_{(m)}$. Punkt $x^*_{(k)} \in U_{(k)} : \nabla h_{(k)}(x^*_{(k)}) = 0$ jest asymptotycznie stabilnym

punktem równowagi systemu dynamicznego φ^* (11), (12) o obszarze przyciągania równym obszarowi selekcyjnemu $U_{(k)}$.

Punkty równowagi systemu φ^*, z funkcją energii daną wzorem (10), które są położone w obszarze brzegowym $\bigcup_{k=1}^{m} \Gamma_{(k)}$ są punktami równowagi nietrwałej. W systemie fizycznym opisanym równaniem stanu (11), (12), trajektoria $x(x^0, t)$ z punktem początkowym x^0 należącym do obszaru brzegowego opuszcza ten obszar w wyniku zaburzenia położenia równowagi będącego skutkiem działających zawsze czynników zakłócających. Zbiór brzegowy określany jest więc jako obszar braku decyzji rozstrzygającej.

Obszar dopełnienia

$$R^n \setminus \bigcup_{k=1}^{m} \bar{U}_{(k)}$$

określany jest jako obszar selekcji negatywnej.

Ze względu na zastosowania, system dynamiczny selekcji — identyfikacji φ^*, realizujący zadanie syntezy dla zadanej rodziny obszarów selekcyjnych i wyróżnionych w nich punkatach identyfikujących, uzupełniany jest blokiem identyfikatora decyzji systemu. Transmitancja bloku identyfikatora dana jest funkcją $\chi(\cdot) : R^n \to R$ spełniającą nastepujące wymagania:

$$\chi(x^*_{(k)} \neq \chi(x^*_{(l)},$$

dla każdej pary punktów $x^*_{(k)} \neq x^*_{(l)}$, oraz $\chi(x) = \text{const}$, dla $x \in R^n \setminus \bigcup_{k=1}^{m} U_{(k)}$. W dalszym ciągu używane są oznaczenia

$$\chi_{(k)} := \chi(x^*_{(k)}), \; k = 1, 2, \ldots, m,$$

oraz

$$\chi_{(0)} := \chi(x), \text{ dla } x \in R^n \setminus \bigcup_{k=1}^{m} U_{(k)}.$$

Wprowadzenie bloku identyfikatora prowadzi w wyniku do identyfikacji położenia asymptotycznego punktu trajektorii $x(x^0, t)$ systemu za pośrednictwem zbioru liczb $\chi_{(0)}, \chi_{(1)}, \ldots, \chi_{(m)}$ (i ogólniej pewnego skończonego zbioru symboli). Funkcja

$$\tilde{\chi}(t) := \chi(x(x^0, t))$$

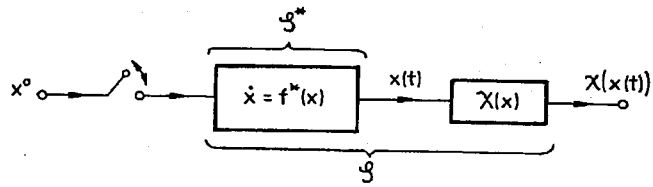
przyjmuje asymptotycznie jedną z wartości:

$$\tilde{\chi}(t \to \infty) = \chi_{(k)}, \text{ gdy } x^0 \in U_{(k)},$$

lub

$$\tilde{\chi}(t \to \infty) = \chi_{(0)}, \text{ gdy } x^0 \notin \bigcup_{k=1}^{m} U_{(k)}.$$

Uwaga 1. Dla zadanego obszaru otwartego, spójnego i ograniczonego $U_{(k)}$, jak w przedstawionym problemie syntezy systemu selekcji — identyfikacji, o brzegu $\Gamma_{(k)}$



Rys. 4. Ogólna struktura blokowa systemu selekcji-identyfikacji φ z uwzględnieniem bloku identyfikatora $\chi(\cdot)$

będącym $(n-1)$-wymiarową gładką hiperpowierzchnią regularną w R^n, przedstawienie za pośrednictwem równania więzów (3) może nie istnieć. Przedstawienie (3) istnieje zawsze lokalnie, w otoczeniu dowolnego punktu $x \in \Gamma_{(k)}$, ale nie koniecznie w odniesieniu do całego brzegu $\Gamma_{(k)}$ [2], [6], t. 2.

Opis za pośrednictwem równania więzów jak we wzorze (3) istnieje jednak zawsze dla brzegu obszaru $U_{(k)}$ będącego obrazem dyfeomorficznym kuli jednostkowej w R^n. W kontekście rozważanego tu problemu, podzbiory tej postaci stanowią, jak się wydaje, wystarczająco szeroką klasę podzbiorów R^n.

Dokładniej, niech każdy z obszarów $U_{(k)}$, ogólnie — obszar U, będzie dany za pośrednictwem pewnego odwzorowania dyfeomorficznego $g(\cdot)$ przestrzeni R^n na R^n,

$$U = g(D_n),$$

gdzie D_n jest otwartą kulą jednostkową w R^n o środku w początku układu współrzędnych ($D_n = \{x \in R^n : \|x\| < 1\}$). Wówczas brzeg Γ obszaru U posiada przedstawienie dane równaniem więzów postaci (3):

$$\Gamma = \left\{ x \in R^n : \sum_{i=1}^{n} ((g)_i^{-1}(x))^2 - 1 = 0 \right\}, \tag{15}$$

gdzie $g(\cdot) = (g_1, g_2, \ldots, g_n)(\cdot)$ i $g^{-1}(\cdot) = ((g^{-1})_1, (g^{-1})_2, \ldots, (g^{-1})_n)(\cdot)$.

Kładąc

$$h(\cdot) \triangleq \left(\sum_{i=1}^{n} \left((g)_i^{-1}(\cdot) \right)^2 \right) - 1,$$

i zapisując (15) w postaci

$$\Gamma = \{x \in R^n : h(x) = 0\}$$

nietrudno zauważyć, że

$$\nabla h(x) = 0$$

w dokładnie jednym punkcie $x = g(0)$. Istnieje więc w tym przypadku dobrze określona funkcja energii $h(\cdot)$ lokalne funkcje energii $h_{(k)}(\cdot)$, (spełniające wymagania wymienione dla procedury syntezy systemu φ^* według formuły (11), (12)). Jeżeli dla danego obszaru U istnieje dyfeomorfizm $g(\cdot)$ jak w (14) (jeżeli istnieje taki dyfeomorfizm dla każdego z obszarów $U_{(k)}$, $k = 1, 2, \ldots, m$), to istnieje też dyfeomorfizm $\bar{g}(\cdot)$: $R^n \underset{na}{\rightarrow} R^n$ taki, że

$$U = \bar{g}(D_n)$$

oraz

$$\bar{g}(0) = x^*,$$

gdzie x^* jest ustalonym dowolnie punktem w U [6], t. 2. Można zatem w tym przypadku, przynajmniej z teoretycznego punktu widzenia, zadać położenie punktów identyfikujących w obszarach selekcyjnych w sposób dowolny — dogodny z punktu widzenia wymagań praktyki. Przy zadanych U i x^*, otwartym problemem pozostaje znalezienie odpowiedniego wyrażenia dla dyfeomorfizmu $\bar{g}(\cdot)$. W pewnych zastosowaniach wyrażenie dla $\bar{g}(\cdot)$ może być dane.

W rozdziale 3 pracy zostanie przedstawiona pewna praktyczna metoda aproksymacji obszarów selekcyjnych, wystarczająca ze względu na dokładność a jednocześnie umożliwiająca syntezę systemu w przypadku ogólnym z możliwością jej kontynuacji aż do poziomu systemu fizycznego. Inna metoda jest ponadto proponowana w rozdziale 2, cz. II pracy. □

Zaproponowana formuła (10) dla syntezy funkcji energii $h(\cdot)$ w postaci „sklejenia" z lokalnych funkcji $h_{(k)}(\cdot)$ odnoszących się do poszczególnych obszarów selekcyjnych $U_{(k)}$, a następnie synteza systemu dynamicznego φ^* w postaci określonej w (11), (12) mogą być związane z trudnościami w realizacji funkcji rozdzielających $\lambda_{(k)}(\cdot)$ — w szczególności, gdy obszary selekcyjne $U_{(k)}$ są rozlokowane w bliskim sąsiedztwie.

Nie zawsze też można oczekiwać znalezienia zwartego przepisu dla funkcji energii $h(\cdot)$, której pewien poziom izoenergetyczny $h(x) = E_0$ wyznacza w przestrzeni stanu R^n rodzinę $\Gamma_{(k)}$, $k = 1, 2, \ldots, m$, brzegów obszarów selekcyjnych $U_{(k)}$ — zadanych w problemie syntezy, oraz dla której zachodzi

$$\bigcup_{k=1}^{m} U_{(k)} = \{x \in R^n : h(x) < E_0\}$$

i $\nabla h(x) = 0$ w zadanych punktach $x^*_{(k)} \in U_{(k)}$.

Zaproponowana zostanie bardziej praktyczna „segmentowa" konstrukcja systemu φ. Niech $U_{(k)}$, $k = 1, 2, \ldots, m$, będzie rodziną obszarów w R^n — spełniających wymagania wymienione na początku rozdziału 2, wraz z wyróżnionymi punktami $x^*_{(k)} \in U_{(k)}$. Odnośnie każdej z funkcji $h_{(k)}(\cdot)$ (klasy C^2) opisującej brzeg $\Gamma_{(k)}$ odpowiedniego obszaru $U_{(k)}$ wymaga się tu aby:

$$\nabla h_{(k)}(x) \neq 0, \text{ dla } x \in \bar{U}_{(k)} \backslash x^*_{(k)}, \tag{16}$$

oraz

$$\nabla h_{(k)}(x^*_{(k)}) = 0. \tag{17}$$

Ponadto, lokalne funkcje energii $h_{(k)}(\cdot)$ modyfikowane są tak, aby

$$h_{(k)}(x^*_{(k)}) = 0, \tag{18}$$

oraz

$$h_{(k)}(x) > 0, \text{ dla każdego } x \in U_{(k)} \backslash x^*_{(k)}. \tag{19}$$

Każda z funkcji $h_{(k)}(\cdot)$ przyjmuje więc w odpowiednim obszarze $U_{(k)}$ wartość minimalną w punkcie $x^*_{(k)}$ i wartość ta wynosi 0. Ze względu na (18) i (19), relacje (3) i (5) opisujące brzeg $\Gamma_{(k)}$ i obszar $U_{(k)}$ przyjmują następującą zmodyfikowaną postać:

$$\Gamma_{(k)} = \{x \in R^n : h_{(k)}(x) = E_{(k)}\}, \tag{20}$$

gdzie $E_{(k)} > 0$, i

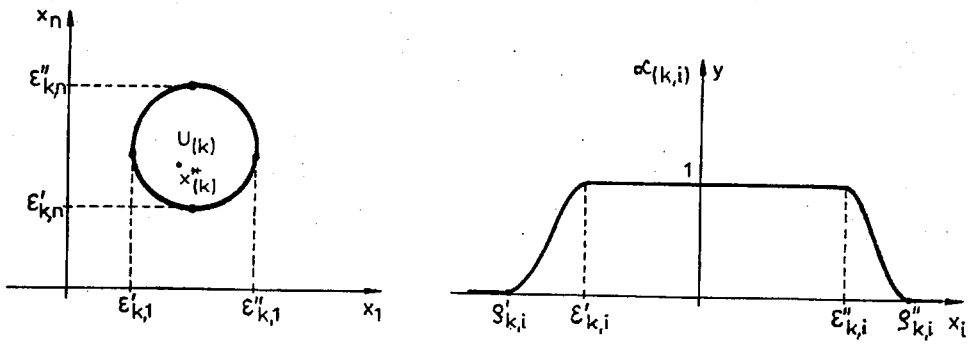
$$U_{(k)} = \{x \in R^n : h_{(k)}(x) < E_{(k)}\}, \tag{21}$$

$k = 1, 2, \ldots, m$.

Dla każdego z obszarów selekcyjnych $U_{(k)}$, $k = 1, 2, \ldots, m$, syntetyzowany jest system dynamiczny $\varphi^*_{(k)}$ według następującej formuły:

$$\dot{x} = \alpha_{(k)}(x) \cdot [h_{(k)}(x) - E_{(k)}] \cdot \nabla h_{(k)}(x), \tag{22}$$

gdzie $\alpha_{(k)}(\cdot)$ jest funkcją klasy C^1 w R^n, oraz $\alpha_{(k)}(x) > 0$ dla $x \in \bar{U}_{(k)}$ i $\alpha_{(k)}(x) = 0$, dla x na zewnątrz pewnego podzbioru ograniczonego w R^n zawierającego $\bar{U}_{(k)}$. Zachowanie się trajektorii każdego z tak otrzymanych systemów $\varphi^*_{(k)}$ jest opisane w treści Twierdzenia 1. Można powiedzieć, że $\varphi^*_{(k)}$ realizuje zadanie syntezy systemu selekcji — identyfikacji dla jednego zadanego obszaru selekcyjnego $U_{(k)}$.



Rys. 5. Ilustracja dla konstrukcji funkcji $\alpha_{(k)}(\cdot)$ w postaci iloczynowej (23)

Każda z funkcji $\alpha_{(k)}(\cdot)$ jest konstruowana w postaci iloczynowej:

$$\alpha_{(k)}(x) = \alpha_{(k,1)}(x_1) \cdot \ldots \cdot \alpha_{(k,n)}(x_n). \tag{23}$$

Funkcja identyfikatora decyzji systemu $\chi(\cdot)$ zadana jest przepisem:

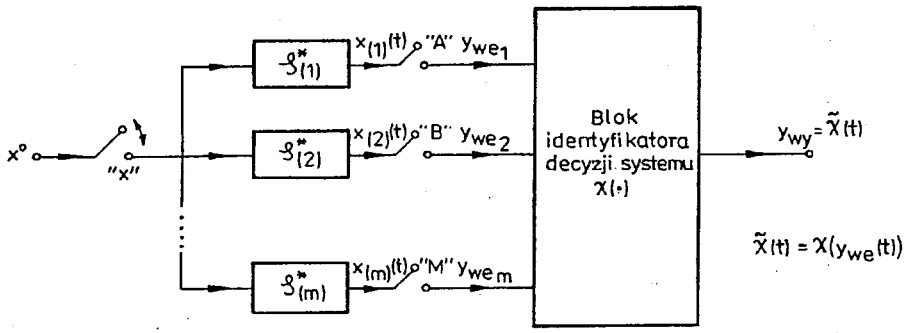
$$\chi(x^*_{(k)}) = k, \; k = 1, 2, \ldots, m,$$

oraz

$$\chi(x) = 0,$$

dla $x \in R^n \setminus \bigcup_{k=1}^{m} U_{(k)}$. Całość konstrukcji systemu selekcji — identyfikacji w postaci „segmentowej" przedstawiona jest na Rys. 6.

Klucz „X" jest zamykany na przedział czasu niezbędny dla przekazania systemowi informacji o wektorze początkowym x^0 (wektorze danych wejściowych). Przyjmuje



Rys. 6. Schemat konstrukcji systemu selekcji-identyfikacji w układzie segmentowym. W chwili początkowej $t=0$ klucze „A"—„M" są otwarte

się, że wartość funkcji $\tilde{X}(\cdot)$ w chwili $t = 0$ wynosi zero. Następnie, gdy w pewnej chwili czasu $t \geqslant 0$

$$h_{(k)}(t) \cong 0, \text{ dla pewnego } k = 1, 2, \ldots, m,$$

następuje zamknięcie klucza „K". W bloku identyfikatora wyznaczana jest nowa wartość funkcji $\chi(\cdot)$, dla przekazanej do bloku $\chi(\cdot)$ wartości wektora $y_{we_t} = x^*_{(k)}$, równa $\chi(x^*_{(k)}) = k$. Po tym stan systemu nie ulega już dalszej zmianie.

Jeżeli natomiast $x^0 \in R^n \setminus \bigcup_{k=1}^{m} \bar{U}_{(k)}$, to każda z trajektorii $x_{(k)}(t)$, $t \geqslant 0$, $x_{(k)}(0) = x^0$, pozostaje na zewnątrz odpowiedniego obszaru selekcyjnego $\bar{U}_{(k)}$, zmierzając asymptotycznie do pewnego punktu równowagi w $R^n \setminus \bar{U}_{(k)}$. Funkcja identyfikatora $\chi(\cdot)$ zachowuje przez cały czas ($t \geqslant 0$) stałą wartość początkową $\tilde{\chi} = 0$.

3. APROKSYMACJA OBSZARÓW SELEKCYJNYCH OBSZARAMI PROSTOPADŁOŚCIENNYMI. SYNTEZA SYSTEMU SELEKCJI — IDENTYFIKACJI Z WYKORZYSTANIEM FUNKTORA 1-WYMIAROWEGO

W systemach selekcji — identyfikacji rozważanych w praktyce (i odpowiadających sformułowaniu ogólnemu podanemu w rozdziale 1) nie zawsze istnieje możliwość znalezienia dla każdego z obszarów selekcyjnych $U_{(k)}$ przedstawienia za pośrednictwem równania więzów jak w (20) i (21), przy spełnionych jednocześnie warunkach (16) i (17) dla odpowiednich lokalnych funkcji energii $h_{(k)}(\cdot)$. Obszary selekcyjne dane są często w postaci graficznej lub za pośrednictwem pół—analitycznego przepisu, i jeżeli nawet istnieje dla nich opis w postaci (20), (21), to może on być trudny do ustalenia lub mieć zbyt złożoną formułę matematyczną (szereg ogólniejszych rozważań dotyczących rozwiązywania zagadnień geometrycznych z wykorzystaniem komputerów jest zawartych w ref. [16], [21]).

W tym rozdziale proponowane są dwie metody syntezy systemu selekcji — identyfikacji, odpowiadające względom praktycznym. W pierwszej z nich każdy z obszarów

selekcyjnych $U_{(k)}$ jest aproksymowany rodziną zbiorów o stosunkowo prostej strukturze, natomiast w drugiej stosuje się do opisu obszarów selekcyjnych więzy w postaci układu nierówności. W metodzie pierwszej, każdy z obszarów $U_{(k)}$ jest aproksymowany rodziną przedziałów otwartych $P_{k,1}, P_{k,2}, \ldots, P_{k,r_k}$ — n-wymiarowych obszarów prostopadłościennych, według formuły

$$U_{(k)} \sim \bigcup_{j=1}^{r_k} P_{k,j}, \tag{24}$$

lub alternatywnej formuły

$$U_{(k)} \sim \bigcap_{j=1}^{r_k} P_{k,j}, \tag{25}$$

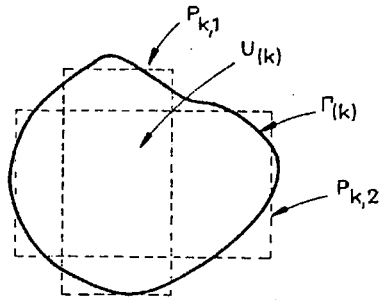
przy czym jako kryterium jakości aproksymacji przyjmuje się n-wymiarową objętość obszaru danego jako różnica symetryczna

$$U_{(k)} \div \bigcup_{j=1}^{r_k} P_{k,j} \tag{26}$$

dla aproksymacji (24), i odpowiednio n-wymiarową objętość obszaru

$$U_{(k)} \div \bigcap_{j=1}^{r_k} P_{k,j} \tag{26'}$$

dla aproksymacji (25) (zakłada się mierzalność każdego z obszarów $U_{(k)}$). Liczba przedziałów aproksymujących powinna być jak najmniejsza.



Rys. 7. Przykład aproksymacji obszaru $U_{(k)}$ dwoma obszarami prostokątnymi $P_{k,1}$ i $P_{k,2}$ według formuły (24)

Kryterium selekcji $x^0 \in U_{(k)}$ jest zastępowane teraz kryterium przybiżonym

$$x^0 \in P_{k,1} \vee x^0 \in P_{k,2} \vee \ldots \vee x^0 \in P_{k,r_k} \tag{27}$$

w przypadku aproksymacji (24), oraz kryterium przybliżonym

$$x^0 \in P_{k,1} \wedge x^0 \in P_{k,2} \wedge \ldots \wedge x^0 \in P_{k,r_k} \tag{28}$$

w przypadku aproksymacji (25).

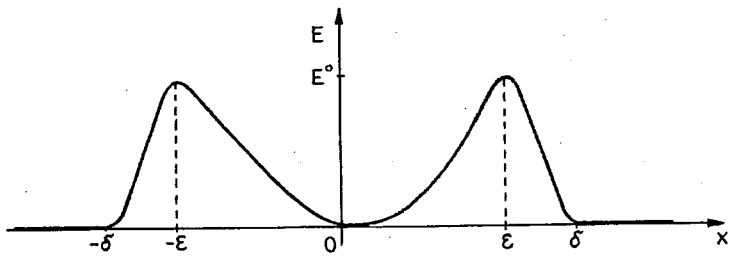
W dalszym ciągu roważana jest synteza systemów selekcji – identyfikacji z obszarami selekcyjnymi przedstawionymi jak w (24) lub (25). Elementarną strukturę składową systemu w proponowanej procedurze syntezy stanowi 1-wymiarowy system dynamiczny selekcji – identyfikacji, 1-wymiarowy funktor selekcji – identyfikacji, o równaniu stanu

$$\dot{x} = f^*(x), f^*(\cdot) \in C^1(R, R) \tag{29}$$

gdzie

$$f^*(x) = -\nabla h(x), h(\cdot) \in C^2(R, R) \tag{30}$$

natomiast funkcja energii $h(\cdot)$ ma postać daną na wykresie na Rys. 8.



Rys. 8. Wykres (znormalizowany) funkcji energii $h(\cdot)$ 1-wymiarowego funktora selekcji-identyfikacji φ_ε^*. Ogólnie, $E_0 > h(x)$ dla $x \in (-\varepsilon, \varepsilon)$ a wartość E^0 funkcji $h(\cdot)$ dla x w przedziałach $(-\infty, -\delta]$ i $[\delta, \infty)$ może być ustalona dowolnie przy spełnieniu $E_0 < E^0$

Funktor ten jest oznaczany symbolem φ_ε^* (ε jest tu liczbą dodatnią).

System φ_ε^* realizuje zadanie selekcji – identyfikacji z jednym obszarem selekcyjnym $U = (-\varepsilon, \varepsilon)$ i punktem identyfikującym $x^* = 0$. Punkt $x^* = 0$ jest asymptotycznie stabilnym punktem równowagi systemu z obszarem przyciągania $(-\varepsilon, \varepsilon)$. Natomiast dla $x^0 = x(t = 0)$ w przedziale $(-\infty, -\varepsilon)$, dodatnia pół-trajektoria $x(x^0, t)$ zmierza asymptotycznie do pewnego punktu równowagi w przedziale $(-\infty, -\delta]$, a dla $x^0 \in (\varepsilon, \infty)$ trajektoria $x(x^0, t)$ systemu osiąga asymptotycznie pewien punkt równowagi zawarty w przedziale $[\delta, \infty)$ (nietrudno jest tu ustalić lokalizację odpowiednich punktów równowagi). Punkty $x = -\varepsilon$ i $x = \varepsilon$ są punktami równowagi nietrwałej. Przy konstruowaniu 1-wymiarowego funktora elementarnego należy zapewnić odpowiednio dużą wartość parametru δ, w celu wystarczającego w praktyce odseparowania położenia obszaru asymptotycznego selekcji negatywnej $(-\infty, -\delta] \cup [\delta, \infty)$ od punktu $x^* = 0$.

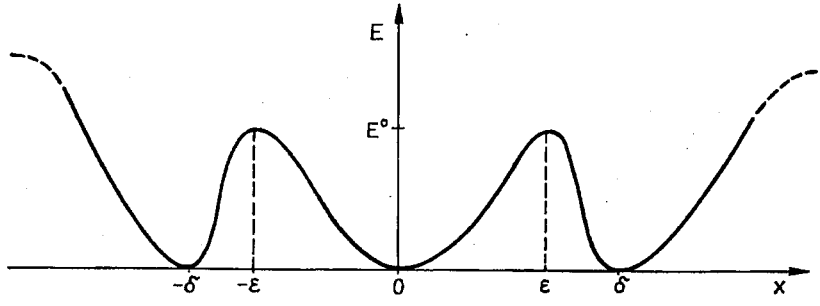
W realizacji praktycznej istotne jest też zapewnienie możliwości stosunkowo łatwej zmiany parametru ε wyznaczającego wielkość obszaru selekcyjnego $U = (-\varepsilon, \varepsilon)$. Ten problem jest omawiany w rozdziale 4 poświęconym zagadnieniom związanym z realizacją elementów systemu selekcji – identyfikacji w technice układów elektronicznych.

Funktor elementarny, przedstawiony powyżej, realizuje zadanie rozdziału zbioru wszystkich stanów wejściowych x^0 wypełniających całą oś rzeczywistą R na dwa obszary selekcyjne. Przy tym, obszar selekcji negatywnej jest dwuczęściowy; składa się

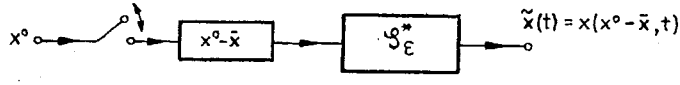
z dwóch rozdzielnych odcinków półprostych $(-\infty, -\varepsilon)$ i (ε, ∞). W istocie jest to więc funktor jednowymiarowy o trzech rozłącznych obszarach selekcyjnych, przy utożsamieniu typu decyzji jako identycznej dla dwóch spośród nich. (Dla porównania, funktor elementarny sieci Hopfielda posiada dwa obszary selekcyjne, jest jednak funktorem 2-wymiarowym-I jest układem dynamicznym o wymiarze przestrzeni stanu równym 2). Punkty identyfikujące dla każdego z obszarów są asymptotycznie stabilnymi punktami równowagi o obszarze przyciągania równym obszarowi selekcyjnemu.

W tym kontekście, dla funktora φ_ε^* punkty równowagi w obszarach $(-\infty, -\delta)$ i (δ, ∞) są punktami stabilnymi, ale nie asymptotycznie stabilnymi. Punkty $x = -\delta$ i $x = \delta$ są asymptotycznie stabilne przy $x^0 \in [-\delta, -\varepsilon)$ i $x^0 \in (\varepsilon, \delta]$, odpowiednio. Taka uproszczona wersja konstrukcji funktora φ_ε^* wydaje się być wystarczająca w zastosowaniach będących przedmiotem cz. I pracy i rozdziału 2, cz. II pracy. Jednak w odniesieniu do ogólnych neuralnych sieci komórkowych, rozważanych w rozdziale 3, cz. II pracy, punkty $x = -\delta$ i $x = \delta$ w modelu matematycznym funktora powinny być punktami asymptotycznie trwałej równowagi, z obszarami przyciągania $(-\infty, -\varepsilon)$ i (ε, ∞), odpowiednio. W tym przypadku funkcja energii $h(\cdot)$ skojarzona z funktorem winna mieć postać jak na Rys. 9. Prostą realizację elektroniczną tego rodzaju funktora przedstawiono w rozdziale 4. Tego typu konstrukcja funktora elementarnego może być oczywiście wykorzystana w syntezie wszystkich rozważanych w tej pracy systemów, zapewniając lepsze właściwości stabilności.

Ze względu na dalszy przebieg procedury syntezy, funktor φ_ε^* jest uzupełniany układem przesuwania poziomu, jak na Rys. 10. Układ z rys. 10 jest oznaczany symbolem $\varphi_\varepsilon^*(\bar{x})$. System $\varphi_\varepsilon^*(\bar{x})$ realizuje zadanie selekcji — identyfikacji z jednym obszarem selekcyjnym $U = (\bar{x} - \varepsilon, \bar{x} + \varepsilon)$, oraz punktem identyfikującym $x^* = 0$.

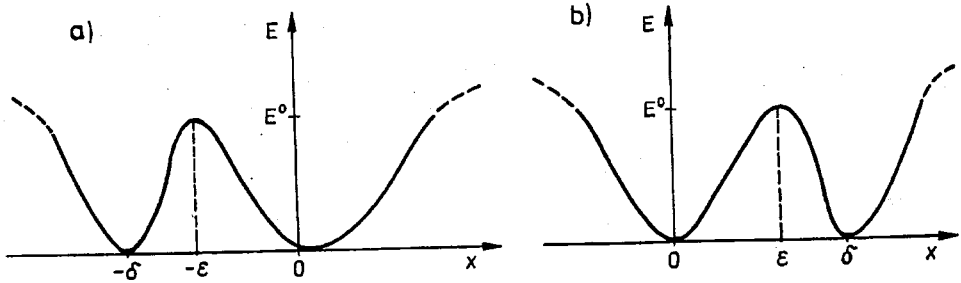


Rys. 9. Wykres (znormalizowany) funkcji energii h(·) 1-wymiarowego funktora selekcji-identyfikacji, zapewniającej asymptotyczną stabilność wszystkich punktów identyfikujących ($x^* = -\delta$, 0, δ) odpowiednich obszarów selekcyjnych



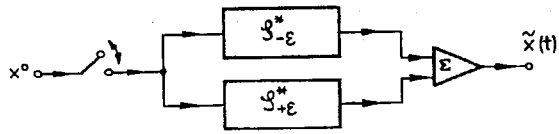
Rys. 10. Uzupełnienie funktora φ_ε^* układem przesuwania poziomu

Punkt identyfikujący jest tu więc ustalony niezależnie od położenia przedziału selekcyjnego U. Dla $x^0 \in (-\infty, \bar{x} - \varepsilon)$ dodatnia półtrajektoria $\tilde{x}(x^0, t) \triangleq x(x^0 - \bar{x}, t)$ osiąga asymptotycznie, dla $t \to \infty$, punkt równowagi w przedziale $(-\infty, -\delta]$, natomiast dla $x^0 \in (\bar{x} + \varepsilon, \infty)$ trajektoria zmierza asymptotycznie do punktu równowagi położonego w przedziale $[\delta, \infty)$. Oprócz funktora φ_ε^* wykorzystywane są funktory selekcji — identyfikacji jednostronne $\varphi_{-\varepsilon}^*$ i $\varphi_{+\varepsilon}^*$, gdzie $\varepsilon > 0$, których funkcje energii mają kształt przedstawiony na wykresach na Rys. 11a i b, odpowiednio. Funktor φ_ε^* o funkcji energii danej na wykresie



Rys. 11. Wykresy (znormalizowane) funkcji energii dla funktorów selekcji-identyfikacji jednostronnych $\varphi_{-\varepsilon}^*$ i $\varphi_{+\varepsilon}^*$, odpowiednio. Równanie elementów jest dane wzorem (30)

na Rys. 9 można otrzymać składając funktory jednostronne $\varphi_{-\varepsilon}^*$ i $\varphi_{-\varepsilon}^*$, jak na schemacie na Rys. 12.



Rys. 12. Konstrukcja układu funktora φ_ε^* z wykorzystaniem funktorów $\varphi_{-\varepsilon}^*$ i $\varphi_{+\varepsilon}^*$. Wartość graniczna zmiennej $\tilde{x}(t)$ dla $t \to \infty$ wynosi 0 wtedy, i tylko wtedy, gdy $x^0 \in (-\varepsilon, \varepsilon)$, oraz $\tilde{x}(t \to \infty) = -\delta(+\delta)$ wtedy, i tylko wtedy, gdy $x^0 \in (-\infty, -\varepsilon)$ $(x^0 \in (\varepsilon, \infty))$

Realizacje układowe funktorów $\varphi_{-\varepsilon}^*$ i $\varphi_{+\varepsilon}^*$ są przedstawione w rozdziale 4.

Przedstawiona zostanie teraz realizacja systemu dynamicznego selekcji — identyfikacji z wykorzystaniem funktora 1-wymiarowego. W pierwszej kolejności rozważana jest konstrukcja systemu z zadanym jednym obszarem selekcyjnym w postaci przedziału P w R^n,

$$P : \varepsilon_i' < x_i < \varepsilon_i'', \; i = 1, 2, \ldots, n,$$

oraz z wyróżnionym punktem identyfikującym x^* położonym w początku układu współrzędnych w R^n.

Niech

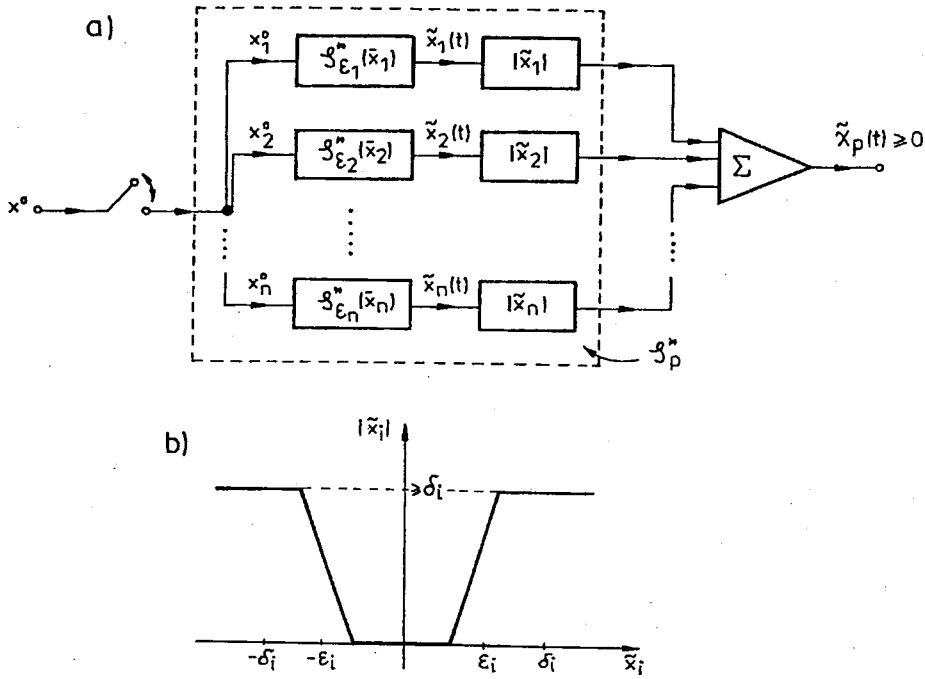
$$\bar{x}_i \triangleq \frac{\varepsilon_i' + \varepsilon_i''}{2}$$

oraz

$$\varepsilon_i \triangleq \frac{\varepsilon_i'' - \varepsilon_i'}{2},$$

dla $i = 1, 2, \ldots, n$.

Schemat blokowy systemu selekcji — identyfikacji φ_P przedstawiony jest na Rys. 13. Charakterystykę elementu funkcyjnego $|\tilde{x}_i|$ podano na Rys. 13b.



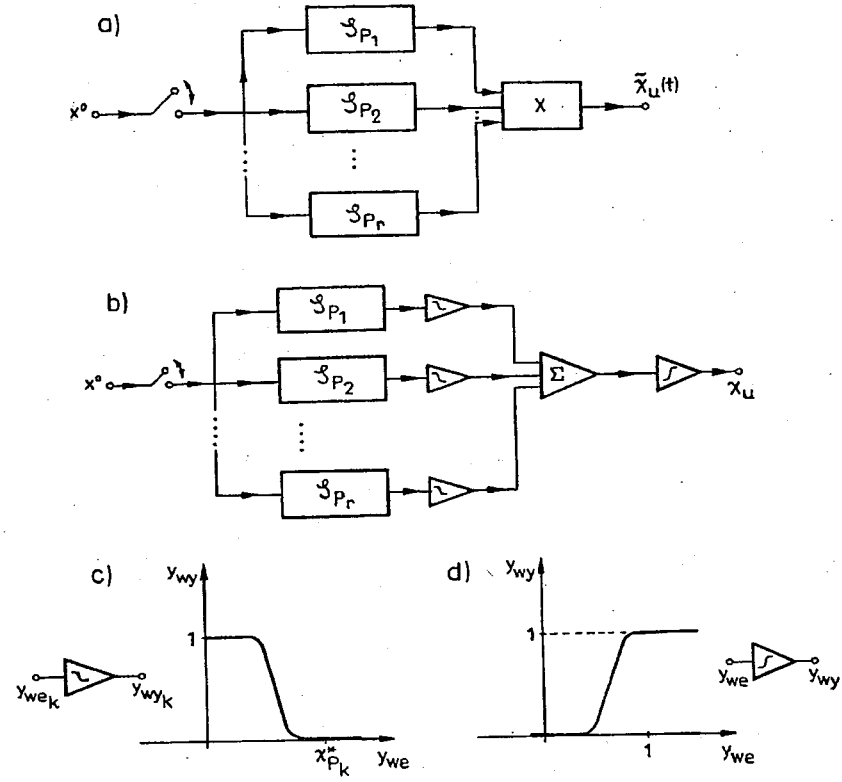
Rys. 13. a. Schemat systemu φ_P. Fragment układu zakreślony linią przerywaną oznaczono symbolem φ_P^*. b. Charakterystyka elementu funkcyjnego $|\tilde{x}_i|$

Z opisu elementów składowych systemu φ_P i sposobu ich połączenia wynika, że $x^0 \in P$ wtedy, i tylko wtedy, gdy $\tilde{\chi}_P(t) \cong 0$ dla pewnej chwili czasu $t \geq 0$. Jeżeli natomiast $x^0 \notin P$, to

$$\tilde{\chi}_P(t) \geq \min_{i=1,2,\ldots,n} \delta_i, \text{ dla } t \geq 0, \tag{31}$$

gdzie δ_i jest parametrem δ odpowiedniego funktora $\varphi_{\varepsilon_i}^*$ (zob. Rys. 8 i 9). Jak widać, przy konstruowaniu funktorów elementarnych wchodzących w skład każdego z bloków $\varphi_{\varepsilon_i}^*(\bar{x}_i)$ należy zapewnić odpowiednio dużą wartość minimalną parametrów δ_i tak, aby wartość $\tilde{\chi}_P(t \to \infty)$ różniła się wystarczająco dużo od 0 w przypadku, gdy $x^0 \notin P$. Inne realizacje systemu φ_P można otrzymać stosując np. funktory jednostronne oraz elementy wyznaczające (x_i^0) na wejściu układu. Można też wykorzystać elementy przesuwania poziomu wartości sygnału.

Następnie, niech dany będzie jeden obszar selekcyjny U w przestrzeni zmiennych – danych wejściowych (przestrzeni parametrów) R^n, otwarty, ograniczony i spójny. Obszar U jest aproksymowany rodziną przedziałów otwartych w R^n według formuły (24) a kryterium selekcyjne $x^0 \in U$ jest zastępowane kryterium przybliżonym (27). W odniesieniu do każdego z przedziałów aproksymujących P_j, $j = 1, 2, \ldots, r$, konstruowany jest odpowiedni system φ_{P_j}, według przedstawionego wyżej schematu (Rys. 13). Całość konstrukcji systemu φ_U jest podana na schemacie na Rys. 14.a.



Rys. 14. System selekcji-identyfikacji φ_U; a. wersja z elementem mnożącym $\times$; b. wersja układu z zastosowaniem sumatora; c,d. charakterystyki elementów formujących. $\chi^*_{P_k}$ oznacza tu wartość progową równą $\min_{i=1,2,\ldots,n} \delta_{k,i}$, charakteryzującą blok φ_{P_k}

Dla systemu φ_U otrzymuje się:

$$x^0 \in U \text{ wtedy, i tylko wtedy, gdy } \tilde{\chi}_U(t) \cong 0 \tag{32}$$

dla pewnej chwili czasu $t \geqslant 0$. Korzystając natomiast z oszacowania (31) otrzymuje się:

$$\text{jeżeli } x^0 \notin \bigcup_{j=1}^{r} P_j, \text{ to } \tilde{\chi}_U(t) \geqslant \prod_{j=1}^{r} \left(\min_{i=1,2,\ldots,n} \delta_{j,i} \right),$$

gdzie założono, że parametry $\delta_{j,i}$ są ustalone tak, że $\delta_{j,i} \geqslant 1$ dla każdego $j = 1, 2, \ldots, r$ oraz $i = 1, 2, \ldots, n$.

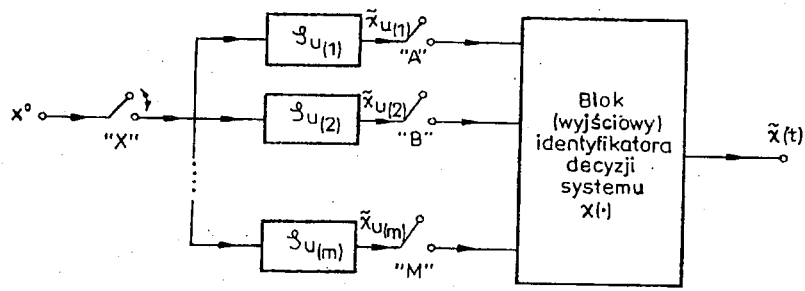
Odnotować należy, że realizacja wielowejściowego elementu mnożącego w układzie z Rys. 14.a. może być w praktyce układów elektronicznych nietechnologiczna. Konstrukcję odpowiedniego układu bez stosowania elementu mnożącego przedstawia Rys. 14.b. Dla realizacji systemu φ_U jak na Rys. 14.b. zachodzi: $x^0 \in U$ wtedy, i tylko wtedy, gdy $\chi_U(t) = 1$ dla pewnego $t \geqslant 0$ oraz $x^0 \notin U$ wtedy, i tylko wtedy, gdy $\chi_U(t) = 0$ dla pewnego $t \geqslant 0$. Stosując formułę aproksymacyjną (25) oraz odpowiednie przybliżone kryterium selekcyjne (28) otrzymujemy wersję dualną φ'_U systemu φ_U zastępując element mnożący w schemacie z Rys. 14.a. wielowejściowym sumatorem. Kryterium selekcyjne (32) dla systemu aproksymującego pozostaje bez zmiany. Zmianie ulega natomiast kryterium (33). Dla systemu φ'_U

$$\text{jeżeli } x^0 \notin \bigcap_{j=1}^{r} P_j, \text{ to } \chi'_U(t) \geqslant \min_{j=1,2,\ldots,r} \left(\min_{i=1,2,\ldots n} \delta_{j,i} \right) \geqslant 1 \tag{34}$$

(przy $\delta_{j,i} \geqslant 1$).

Konstrukcja całości systemu selekcji — identyfikacji, dla zadanej rodziny obszarów selekcyjnych $U_{(1)}, U_{(2)}, \ldots, U_{(m)}$, otwartych, spójnych i ograniczonych oraz parami rozłącznych wraz z brzegami, przebiega teraz następująco. Dla każdego z oszarów $U_{(k)}$ dokonuje się jego aproksymacji według formuły (24) lub (25). Zakłada się, że domknięcia zbiorów aproksymujących dla różnych obszarów selekcyjnych są parami rozłącznymi podzbiorami R^n. Dla różnych podzbiorów selekcyjnych można przy tym stosować różne aproksymacje (według przepisu (24) lub (25)).

Następnie, dla przyjętych aproksymacji obszarów $U_{(k)}$ konstruuje się w odniesieniu do każdej z nich odpowiedni system dynamiczny selekcji — identyfikacji $\varphi_{U_{(k)}}$ lub $\varphi'_{U_{(k)}}$, $k = 1, 2, \ldots, m$. Konstrukcję całości systemu otrzymuje się jak na schemacie na Rys. 15. (por. schemat z Rys. 6).

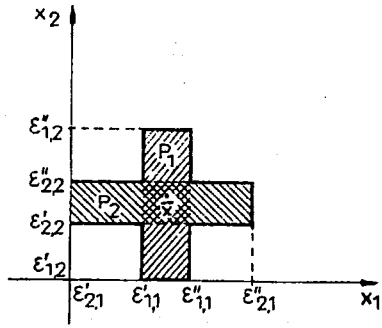


Rys. 15. Schemat konstrukcji systemu selekcji-identyfikacji z blokami selektorów typu φ_U lub φ'_U, odpowiadającymi odpowiednio aproksymacji (24) lub (25). W chwili $t=0$ klucze „A"—„M" są otwarte

Wartość funkcji $\tilde{\chi}(\cdot)$ w chwili $t = 0$ wynosi 0. Jeżeli x^0 nie należy do jednego z obszarów aproksymujących, przypisanych każdemu z obszarów $U_{(k)}$, to klucze „A" — „M" pozostają otwarte dla wszystkich $t \geqslant 0$, a funkcja $\tilde{\chi}(\cdot)$ utrzymuje stałą wartość początkową $\chi = 0$.

Jeżeli natomiast $x^0 \in U_{(k)}$, dla pewnego $k = 1, 2, \ldots, m$, to $\tilde{\chi}_{U_{(k)}}(t)$ przyjmuje wartość asymptotyczną dla $t \to \infty$ równą zero. W pewnej chwili czasu $t > 0$ następuje zamknięcie klucza „K" (każdy z kluczy „K" jest sterowany wartością $\tilde{\chi}_{U_{(k)}}(t)$ i jest zwierany w chwili, gdy $\tilde{\chi}_{U_{(k)}}(t) \cong 0$). Blok identyfikatora wyznacza teraz odpowiednią wartość funkcji $\chi(\cdot)$ równą k. Po tym stan systemu nie ulega już dalszej zmianie.

Przykład: Niech obszar selekcyjny dla pewnego 2-wymiarowego zadania selekcji – identyfikacji dany będzie jak na Rys. 16.



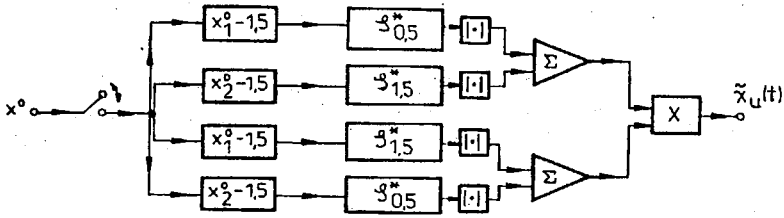
Rys. 16. Obszar selekcyjny U dany w postaci sumy mnogościowej dwóch przedziałów (obszarów prostokątnych) $P_1 = \{(x_1, x_2)$: $\varepsilon'_{1,1} < x_1 < \varepsilon''_{1,1}$, $\varepsilon'_{1,2} < x_2 < \varepsilon''_{1,2}\}$ i $P_2 = \{(x_1, x_2)$: $\varepsilon'_{2,1} < x_1 < \varepsilon''_{2,1}$, $\varepsilon'_{2,2} < x_2 < \varepsilon''_{2,2}\}$. Przyjęto $\varepsilon'_{1,1} = 1$, $\varepsilon''_{1,1} = 2$, $\varepsilon'_{1,2} = 0$, $\varepsilon''_{1,2} = 3$, $\varepsilon'_{2,1} = 0$, $\varepsilon''_{2,1} = 3$, $\varepsilon'_{2,2} = 1$, $\varepsilon''_{2,2} = 2$

Punkt $\bar{x}$ ma współrzędne $\bar{x}_1 = 1.5$, $\bar{x}_2 = 1.5$. I dalej, parametry $\varepsilon_{j,i}$ obszaru U przesuniętego do początku układu współrzędnych tak, aby punkt $\bar{x}$ pokrył się z punktem 0 są równe:

$$\varepsilon_{1,1} = 0.5,\ \varepsilon_{1,2} = 1.5,$$

$$\varepsilon_{2,1} = 1.5,\ \varepsilon_{2,2} = 0.5.$$

Konstrukcję systemu dynamicznego φ_U realizującego powyższe zadanie selekcji – identyfikacji podano na Rys. 17.



Rys. 17. System realizujący postawione zadanie selekcji-identyfikacji. $\tilde{x} \in U$ wtedy; i tylko wtedy, gdy $\tilde{\chi}_U(t) = 0$ dla $t \to \infty$

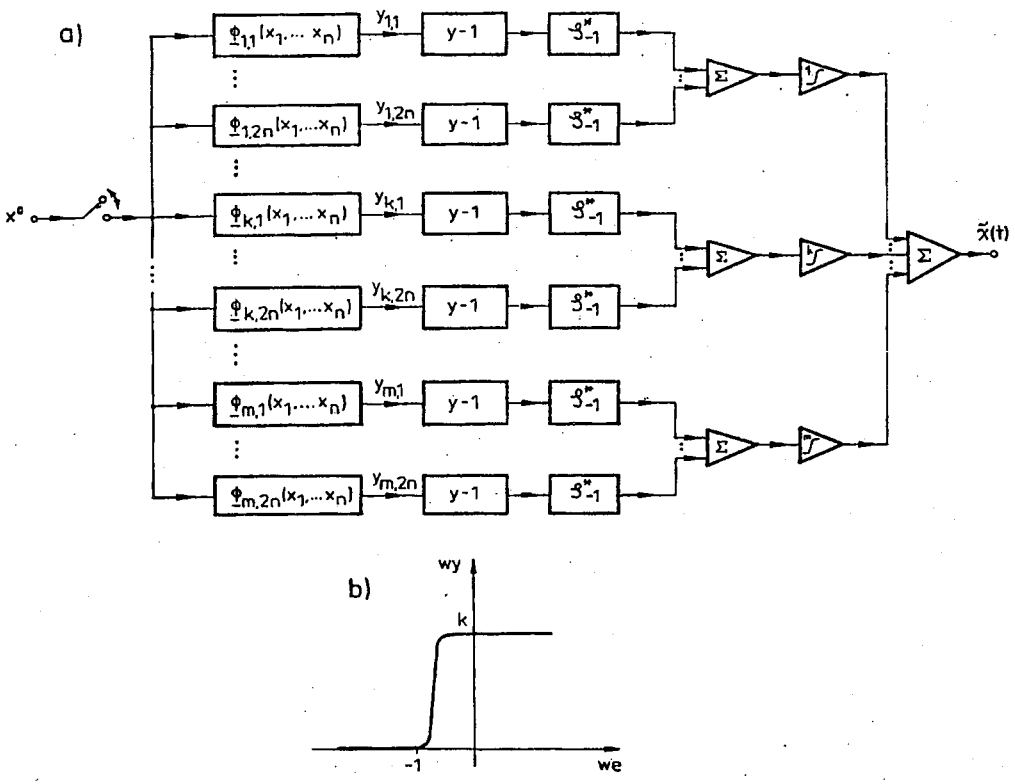
□

W drugiej z proponowanych metod syntezy systemu selekcji – identyfikacji każdy z obszarów selekcyjnych $U_{(k)}$, $k = 1,2, \ldots, m$, jest przedstawiany za pośrednictwem rodziny więzów nierównościowych,

$$U_{(k)} = \{(x_1, \ldots, x_n) \in R^n : \underline{\Phi}_{k,1}(x_1, \ldots, x_n) > 0, \ldots, \underline{\Phi}_{k,2n}(x_1, \ldots, x_n) > 0\}. \quad (35)$$

Powyższe przdstawienie może być formułą dokładną lyb też stanowić pewną aproksymację. Odnośnie podzbiorów $U_{(k)}$ zakłada się, że są parami rozłączne (dotyczy to również aproksymacji, jeżeli przedstawienia (35) są pewnymi przybliżeniami opisów dokładnych).

Strukturę systemu dynamicznego realizującego zadanie selekcji — identyfikacji z zadaną rodziną obszarów selekcyjnych $U_{(k)}$ otzrymuje się teraz w postaci przedstawionej na Rys. 18.



Rys. 18. a. Struktura systemu dynamicznego selekcji-identyfikacji otrzymana dla przedstawienia (35) dla obszarów selekcyjnych. W układzie wykorzystano funktor jednostronny φ^*_{-1} o funkcji energii jak na Rys. 11.a. Charakterystyka elementu formującego z indeksem k, $k=1,2,\ldots,m$, jest przedstawiona na Rys. b

Dla systemu z Rys. 18 zachodzi: $x^0 \in U_{(k)}$ wtedy, i tylko wtedy, gdy $\tilde{\chi}(t \to \infty) = k$ oraz

$$x^0 \notin \bigcup_{k=1}^{m} U_{(k)}$$

wtedy, i tylko wtedy, gdy $\tilde{\chi}(t) = 0$, dla każdego $t > 0$.

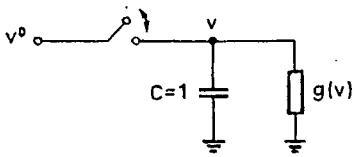
4. UWAGI ODNOŚNIE REALIZACJI ELEMENTÓW SYSTEMU SELEKCJI – IDENTYFIKACJI W TECHNICE UKŁADÓW ELEKTRONICZNYCH

A. Synteza układu 1-wymiarowego funktora selekcji – identyfikacji

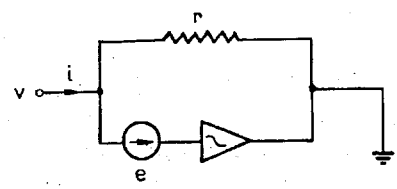
Opis matematyczny tego funktora podano we wzorach (29) i (30). Na rysunkach 8, 9 i 11 przedstawiono natomiast pożądany kształt funkcji energii $h(\cdot)$ skojarzonej z funktorem, dla różnych jego realizacji. Funktory $\varphi^*_{-\varepsilon}$, $\varphi^*_{+\varepsilon}$ i φ^*_{ε} mają podstawowe znaczenie w przedstawionym w rozdziale 3 schemacie syntezy systemu selekcji identyfikacji, realizując funkcję selekcji danych w obszarze 1-wymiarowym. Ze względu na ich prostą konstrukcję oraz dalsze zastosowania są one też nazywane elementarnymi funktorami mnogościowo–logicznymi.

Funktory $\varphi^*_{-\varepsilon}$ i $\varphi^*_{+\varepsilon}$ są układami dynamicznymi bistabilnymi. Tego rodzaju układy są dobrze znane w technice układów impulsowych [9]. Funktor $\varphi^*_{-\varepsilon}$ posiada dwa asymptotycznie stabilne stany równowagi, stan równowagi $x^* = 0$ o obszarze przyciągania $(-\varepsilon, \infty)$, oraz stan równowagi $x^* = -\delta$ o obszarze przyciągania $(-\infty, -\varepsilon)$. Punkt $x^* = -\varepsilon$ jest stanem równowagi nietrwałej. W kontekście rozważanego w pracy zagadnienia selekcji – identyfikacji przedział $(-\varepsilon, \infty)$ jest interpretowany jako obszar selekcji pozytywnej, z punktem identyfikującym $x^* = 0$, natomiast przedział $(-\infty, -\varepsilon)$ wyznacza obszar selekcji negatywnej, z punktem identyfikującym położonym w $x = -\delta$.

Opis dynamiki funktora $\varphi^*_{+\varepsilon}$ jest analogiczny do przedstawionego opisu dla $\varphi^*_{-\varepsilon}$. Przedział $(-\infty, \varepsilon)$ jest tu obszarem selekcji pozytywnej, natomiast (ε, ∞) jest przedziałem selekcji negatywnej. Zważywszy, że konstrukcję układu funktora symetrycznego otrzymuje się w prosty sposób z wykorzystaniem funktorów jednostronnych $\varphi^*_{-\varepsilon}$ i $\varphi^*_{+\varepsilon}$ (zob. schemat z Rys. 12), w dalszym ciągu rozważana będzie realizacja układowa dla $\varphi^*_{-\varepsilon}$ i $\varphi^*_{+\varepsilon}$. Przy interpretacji zmiennej wejściowej x jako napięcia na pewnym dwójniku, układ elektryczny dla funktora $\varphi^*_{-\varepsilon}$ otrzymuje się jak na Rys. 19. Konstrukcja dwójnika bezinercyjnego $i = g(v)$ jest dana na schemacie na Rys. 20.

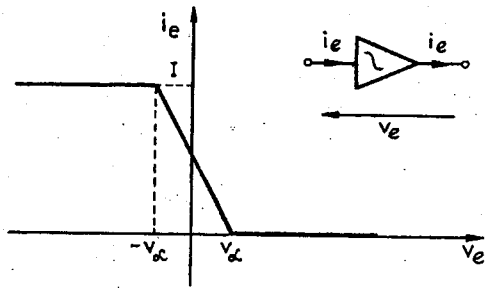


Rys. 19. Schemat układu elektrycznego realizującego zadanie funktora $\varphi^*_{-\varepsilon}$

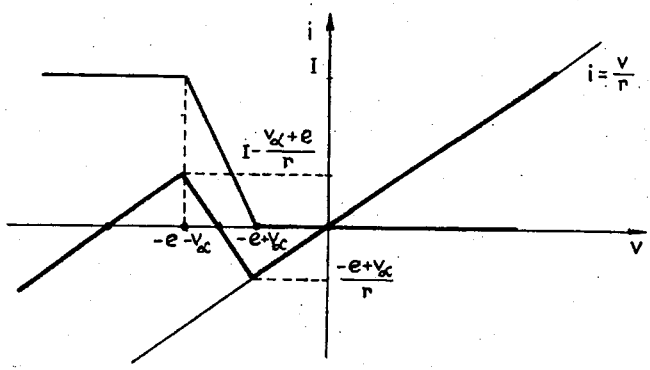


Rys. 20. Schemat elektryczny dwójnika $i=g(v)$

Charakterystyka elementu nieliniowego jest przedstawiona na Rys. 21, natomiast na Rys. 22 wyznaczono charakterystykę całego dwójnika. Zakłada się tu, że



Rys. 21. Charakterystyka elementu nieliniowego. Podano tu aproksymację odcinkowo-liniową



Rys. 22. Charakterystyka napięciowo-prądowa $i=g(v)$ dwójnika z Rys. 20

$$r > \frac{v_\alpha + e}{I} \tag{36}$$

oraz

$$E > v_\alpha. \tag{37}$$

Równanie dynamiczne dla układu z Rys. 19 ma postać

$$\dot{v} = -g(v).$$

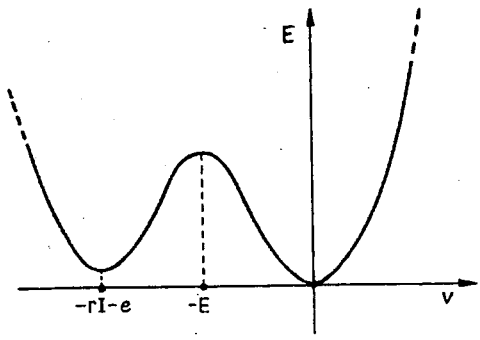
Dla układu tego istnieje dobrze określona funkcja energii $h(\cdot)$, której wykres przedstawiono na Rys. 23. Równoważna postać równania dynamicznego układu jest następująca

$$\dot{v} = -\nabla h(v),$$

skąd nietrudno wywnioskować, że:

— punkt $v^* = 0$ jest asymptotycznie stabilnym punktem równowagi układu z obszarem przyciągania $(-E, \infty)$;

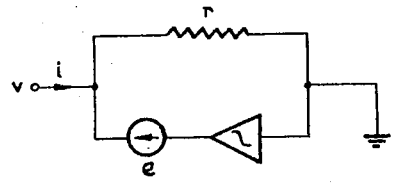
— punkt $v = -E$ jest niestabilnym punktem równowagi;



Rys. 23. Wykres funkcji energii dla układu z Rys. 19

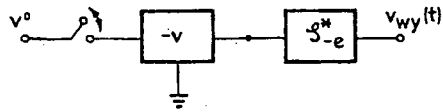
– punkt $v = -r \cdot I - e$ jest asymptotycznie stabilnym punktem równowagi układu z obszarem przyciągania $(-\infty, -E)$.

Układ z Rys. 19 realizuje zatem funkcję selekcji – identyfikacji odpowiednią dla funktora φ^*_{-e} (przy uwzględnieniu ograniczenia (37)). Zmianę rozmiaru obszaru selekcji pozytywnej otrzymuje się tu poprzez regulację wydajności źródła napięciowego e w zakresie uwzględniającym ograniczenie $e > v_\alpha$. Schemat układu elektrycznego odpowiedniego dwójnika $i = g(v)$ dla funktora φ^*_{+e} jest przedstawiony na Rys. 24. Charakterystyka rezystora nieliniowego jest jak na Rys. 21.



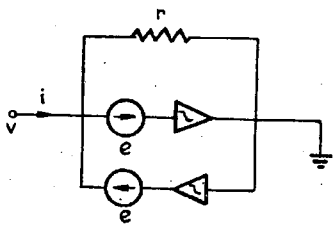
Rys. 24. Schemat elektryczny dwójnika $i=g(v)$ dla funktora φ^*_{+e}

Inny sposób konstrukcji układu funktora φ^*_{+e} z wykorzystaniem odpowiedniego układu dla φ^*_{-e} przedstawia Rys. 25.



Rys. 25. Konstrukcja układu funktora φ^*_{+e} z wykorzystaniem elementu φ^*_{-e}

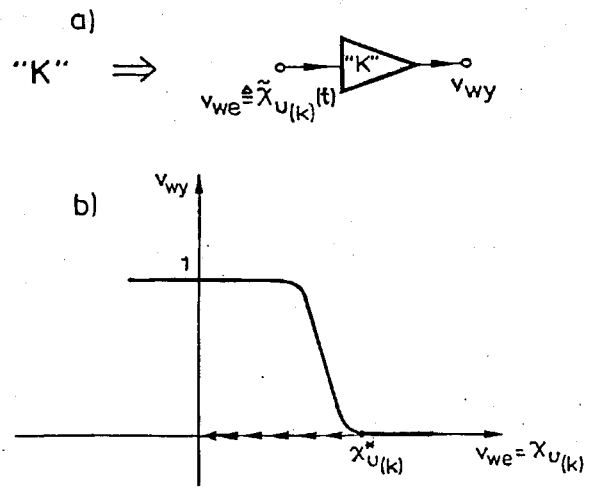
Natomiast na Rys. 26 podano schemat układu elektrycznego dla funktora φ^*_e z użyciem dwóch elementów $i = g(v)$ o charakterystyce jak na rys. 21.



Rys. 26. Schemat układu elektrycznego dla funktora φ_e^*

B. Realizacja kluczy „A" – „M" w układzie selekcji – identyfikacji φ

Do wyjścia każdego z bloków $\varphi_{U_{(k)}}$ systemu φ przedstawionego na Rys. 15 (zob. też system przedstawiony na Rys. 6) dołączony jest klucz „K" sterowany wartością funkcji identyfikatora $\tilde{\chi}_{U_{(k)}}(\cdot)$ w bieżącej chwili czasu $t > 0$. Realizację klucza „K" proponuje się w postaci układu przedstawionego na Rys. 27. Zakłada się reprezentaję napięciową dla $\chi_{U_{(k)}}$.



Rys. 27. Na wykresie b. przedstawiono charakterystykę przetwornika $v_{we} \to v_{wy}$ realizującego funkcję klucza „K". $\chi^*_{U_k}$ oznacza tu wartość progową daną wzorami (33) lub (34), w zależności od sposobu aproksymacji obszarów $U_{(k)}$

Dla $v^0 \in U_{(k)}$ $(v^0 = (v_1^0, \ldots, v_n^0))$ napięcie $v_{we}(t) = \tilde{\chi}_{U_{(k)}}(t)$ osiąga asymptotycznie wartość 0 (zachodzi przy tym $v_{we}(t) \geqslant 0$, dla każdego $t \geqslant 0$). Gdy natomiast $v^0 \in R^n \backslash \bar{U}_{(k)}$, wartość napięcia $v_{we} = \tilde{\chi}_{U_{(k)}}(t)$ pozostaje powyżej wartości progowej $\chi^*_{U_{(k)}}$ oszacowanej we wzorach (33) i (34), w zależności od sposobu aproksymacji przyjętego dla obszaru selekcyjnego $U_{(k)}$. A zatem, dla $v^0 \in U_{(k)}$, $v_{wy}(t)$ przyjmuje wartość 1 w granicy dla $t \to \infty$, oraz $v_{wy} = 0$, gdy $v^0 \in R^n \backslash U_{(k)}$, dla wszystkich $t \geqslant 0$.

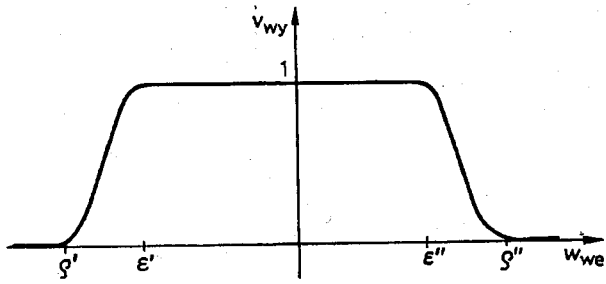
Blok wyjściowy identyfikatora decyzji systemu $\chi(\cdot)$ na schemacie z Rys. 15 konstruowany jest zgodnie z następującym przepisem:

a). jeżeli dla pewego $t \geqslant 0$ napięcie $v_{wy}(t)$ na wyjściu klucza „K" przyjmuje wartość 1, to $\chi(t)$ przyjmuje wartość równą k;

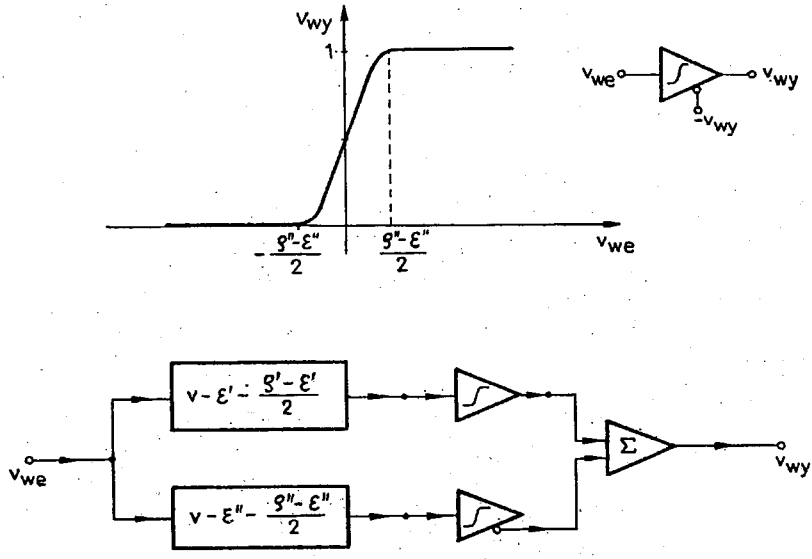
b). jeżeli $v_{wy} = 0$ na wyjściu wszystkich kluczy „A" – „M", to $\chi(t)$ utrzymuje wartość początkową równą 0.

C. Synteza charakterystyki $\alpha(\cdot)$

W opisanym w rozdziale 2 schemacie syntezy systemu selekcji – identyfikacji φ (wzory (16) – (23)) stosowany jest element funkcyjny $\alpha(\cdot)$ opisany wykresem przedstawionym na Rys. 5. Element $\alpha(\cdot)$ jest wykorzystywany w celu zapewnienia ograniczoności trajektorii każdego z podsystemów $\varphi^*_{(k)}$ (równanie stanu (22)) wchodzących w skład całego systemu φ, Rys. 6.



Rys. 28. Wykres charakterystyki $\alpha(\cdot)$ (przyjmuje się $\rho''-\varepsilon''=\varepsilon'-\rho'$)



Rys. 29. Układ realizujący charakterystykę $\alpha(\cdot)$ z Rys. 28

Syntezę elementu o charakterystyce (v_{we}, v_{wy}) jak na Rys. 28 można wykonać z wykorzystaniem dwóch elementów o charakterystykach typu sigmoidalnego.

DODATEK. DOWÓD TWIERDZENIA 1.

Niech Q oznacza zbiór punktów równowagi systemu (11), (12). Q jest sumą mnogościową następujących podzbiorów przestrzeni stanu R^n systemu:

$$Q = Q^{(+)} \cup \left(\bigcup_{k=1}^{m} \Gamma_{(k)}\right) \cup Q^{(1)}, \tag{D. 1}$$

gdzie $Q^{(+)}$ jest zbiorem punktów równowagi położonych w obszarze dopełnienia $R^n \setminus \bigcup_{k=1}^{m} \bar{U}_{(k)}$, natomiast $Q^{(-)} \triangleq \bigcup_{k=1}^{m} Q_{(k)}$, gdzie $Q_{(k)}$ oznacza zbiór złożony z punktów równowagi systemu, położonych w obszarze $U_{(k)}$. Z (5) i (11) wynika bezpośrednio, że każdy z podzbiorów $Q_{(k)}$, $k = 1, 2, \ldots, m$, jest zbiorem punktów krytycznych funkcji $h(\cdot)$ zawartych w $U_{(k)}$. Korzystając z faktu, że $\nabla h_{(k)}(x) \neq 0$ dla każdego $x \in \Gamma_{(k)}$, $k = 1, 2, \ldots, m$, oraz biorąc pod uwagę wzór (10) definiujący funkcję $h(\cdot)$ otrzymuje się, że $Q^{(+)}$ i $Q^{(-)}$ są podzbiorami domkniętymi w R^n, i zatem

$$\mathrm{d}Q^{(+)} \cap \left(\bigcup_{k=1}^{m} \Gamma_{(k)}\right) = \emptyset \text{ oraz } Q^{(-)} \cap \left(\bigcup_{k=1}^{m} \Gamma_{(k)}\right) = \emptyset.$$

Niech $x^0 = x(t = 0) \in U_{(k)}$, dla pewnego ustalonego k, $k = 1, 2, \ldots, m$. Wzdłuż dodatniej półtrajektorii $x(x^0, t)$, $t \geqslant 0$, systemu (11), (12), odpowiadającej punktowi początkowemu x^0, zachodzi:

$$\dot{h}(t) = \alpha(x) \cdot h(x) \cdot [\nabla h(x)]^2 \leqslant 0, \left(\dot{h}(t) \triangleq \frac{\mathrm{d}}{\mathrm{d}t} h(x(x^0, t))\right), \tag{D.2}$$

dla każdego $t \geqslant 0$, oraz $\dot{h}(t) = 0$ wtedy, i tylko wtedy, gdy $x^0 \in Q_{(k)}$. Korzystając z opisu zbiorów $U_{(k)}$ i ich brzegów $\Gamma_{(k)}$, założonego odpowiednio w postaci danej relacjami (5) i (3) otrzymuje się stąd, że półtrajektoria $x(x^0, t)$, $x^0 \in U_{(k)}$, jest zawarta całkowicie w obszarze $U_{(k)}$. Definiując funkcję $h(\cdot)$ jako funkcję Lapunowa systemu (11), (12) i korzystając z kryteriów Lapunowskich stabilności [5], [17], [23] wnioskuje się na podstawie powyższego, że $x(x^0, t)$ zmierza asymptotycznie, dla $t \to \infty$, do zbioru złożonego z punktów równowagi systemu zawartych w obszarze $U_{(k)}$ (do zbioru $Q_{(k)}$). Niech następnie $x^0 \in R^n \setminus \bigcup_{k=1}^{m} \bar{U}_{(k)}$. Nietrudno zauważyć, biorąc pod uwagę opis (3), (5) podzbiorów $U_{(k)}$ i ich brzegów $\Gamma_{(k)}$ oraz korzystając z wyrażenia na $\dot{h}(t)$, że

$$\dot{h}(t) \geqslant 0 \tag{D. 3}$$

wzdłuż półtrajektorii $x(x^0, t)$, $t \geqslant 0$, przy czym

$$\dot{h}(t) = 0 \tag{D. 4}$$

dla pewnego $t \geqslant 0$ wtedy, i tylko wtedy gdy $x^0 \in Q^{(+)}$, i wówczas $x(x^0, t) = x^0$, dla wszystkich $t \in R$. Dodatnia półtrajektoria $x(x^0, t)$ jest więc, dla każdego warunku początkowego $x^0 = x(t = 0)$ w obszarze dopełnienia $R^n \setminus \bigcup_{k=1}^{m} \bar{U}_{(k)}$ zawarta tam całkowicie. Z wyrażenia (10) podającego konstrukcję funkcji $h(\cdot)$ wynika natomiast, że trajektoria $x(x^0,t)$, $t \geqslant 0$, jest ograniczona, tzn.

$$\sup_{t \geqslant 0} \|x(x^0, t)\| < \infty.$$

Korzystając z wymienionych właściwości trajektorii $x(x^0, t)$, $x^0 \in R^n \setminus \bigcup_{k=1}^{m} \bar{U}_{(k)}$, $t \geqslant 0$, wnioskuje się [5], [17], że jej zbiór ω- graniczny Ω_{x^0} jest zwartym (i niepustym) podzbiorem zawartym w

$$\mathrm{cl}\left(R^n \setminus \bigcup_{k=1}^{m} \bar{U}_{(k)}\right) = R^n \setminus \bigcup_{k=1}^{m} U_{(k)}.$$

Zostanie teraz wykazane, że Ω_{x^0} jest zbiorem złożonym z punktów równowagi systemu (11), (12) należących do $Q^{(+)}$.

Wobec zwartości zbioru ω-granicznego Ω_{x^0} (jako podzbioru R^n), istnieje punkt $\bar{x} \in \Omega_{x^0}$ taki, że

$$h(\bar{x}) = \inf_{x \in \Omega_{x^0}} h(x).$$

Z relacji (D.3) i (D.4) dla dowolnej dodatniej półtrajektorii w obszarze dopełnienia $R^n \setminus \bigcup_{k=1}^{m} \bar{U}_{(k)}$ oraz wobec niezmienniczości zbioru ω-granicznego Ω_{x^0} (względem przepływu generowanego przez rozwiązania na systemu na R^n) wynika natychmiast, że $\bar{x}$ jest punktem równowagi systemu (11), (12), $\bar{x} \in Q^{(+)}$.

Korzystając z tych samych argumentów wykazuje się też, że istnieje w Ω_{x^0} punkt $\bar{\bar{x}}$ taki, że

$$h(\bar{\bar{x}}) = \sup_{x \in \Omega_{x^0}} h(x),$$

oraz, że $\bar{\bar{x}} \in Q^{(+)}$. Z relacji (D. 3) wynika,że

$$h(\bar{x}) = h(\bar{\bar{x}}). \tag{D. 5}$$

Istotnie, niech bowiem zachodzi nierówność

$$h(\bar{x}) \neq h(\bar{\bar{x}}). \tag{D. 6}$$

Na mocy definicji zbioru ω-granicznego trajektorii, istnieją dwa ciągi

$$0 \leqslant t'_1 < t'_2 < \ldots < t'_k \to \infty, \text{ dla } k \to \infty,$$

oraz

$$0 \leqslant t''_1 < t''_2 < \ldots < t''_l \to \infty, \text{ dla } l \to \infty,$$

takie, że $\lim_{k\to\infty} x(x^0, t'_k) = \bar{x}$ oraz $\lim_{l\to\infty} x(x^0, t''_l) = \bar{\bar{x}}$. Przy spełnionej nierówności (D. 6) istniałyby wówczas takie wartości $t > 0$, dla których

$$\dot{h}(t) < 0,$$

wzdłuż trajektorii $x(x^0, t)$, co jest w sprzeczności z (D. 3).

A zatem $h(x) = \text{const}$, dla $x \in \Omega_{x^0}$, skąd wobec niezmienniczości podzbioru Ω_{x^0},

$$\frac{\mathrm{d}}{\mathrm{d}t} h(x(\tilde{x}^0, t)) = 0$$

dla każdego punktu $\tilde{x}^0 \in \Omega_{x^0}$, dla każdego t. Zbiór ω-graniczny Ω_{x^0} jest więc podzbiorem $Q^{(+)}$. Końcowy fragment tezy twierdzenia wynika bezpośrednio z konstrukcji systemu (11), (12) oraz z wykazanych właściwości jego rozwiązań.

BIBLIOGRAFIA

1. W. I. Arnold: *Teoria równań różniczkowych.* Warszawa: PWN, 1983
2. L. Auslander, R. E. MacKenzie: *Rozmaitości różniczkowalne.* Warszawa: PWN, 1969
3. L. O. Chua, L. Yang: *Cellular neural networks: theory.* IEEE Trans. Circ. and Systems, 1988, vol. CAS-35, nr 10, pp. 1257–1271
4. L. O. Chua, L. Yang: *Cellular neural networks: applications.* IEEE Trans. Circ. and Systems, 1988, vol. CAS–35, nr 10, 1273–1290
5. B. P. Demidowicz: *Matematyczna teoria stabilności.* Warszawa: WNT, 1972
6. R. Duda: *Wprowadzenie do topologii, część I. Topologia ogólna., część II. Topologia algebraiczna i topologia rozmaitości.* Warszawa: PWN, 1986
7. N. El–Leithy, R. W. Newcomb (Eds.): *Special Issue on Neural Networks.* IEEE Trans. Circuits and Systems, 1989, vol. CAS–36, nr 5.
8. N. El–Leithy, R. W. Newcomb, M. Zaghloul: *A basic MOS neural–type junction: a prospective on neural–type microsystems.* Proc. of ICNN'87, San Diego, CA, June 1987, pp. III–469 – III–477
9. L. M. Goldenberg: *Teoria i obliczanie półprzewodnikowych układów impulsowych.* Warszawa: WNT, 1972
10. J. J. Hopfield: *Neural networkss and physical systems with emergent computational abilities.* Proc. Natl. Acad. Sci. USA., 1982, vol. 79, pp. 2554–2558
11. J. J. Hopfield, D. W. Tank: *„Neural" computation of decisions in optimization problems.* Biol. Cybern., 1985, vol. 52, pp. 141–152
12. J. J. Hopfield, D. W. Tank: *Computing with neural circuits: a model.* Science (USA), 1986, vol. 233, no 4764, pp. 625–633
13. W. Y. Huang, R. P. Lippmann: *Neural net and traditional classiffiers.* Presented at the IEEE Conf. on Neural Information Processing Systems – Natural and Synthetic, Denver, CO, Nov. 1987
14. T. Kohonen: *Self – Organization and Associative Memory.* Springer – Verlag, New York, 1984
15. R. P. Lippmann: *An Introduction to computing with neural networks.* IEEE ASSP Mag., Apr. 1987, pp. 4–22.
16. R. Miller, Q. F. Stout: *Mesh computer algorithms for computational geometry.* IEEE Trans. on Computers, 1989, vol. C–38, no 3, pp. 321–340.
17. V. V. Nemytskii, V. V. Stepanov: *Qualitative Theory of Differential Equations.* Princeton University Press, N. J., 1972

18. R. W. Newcomb: *Neural – type microsystems: some circuits and considerations.* Proc. IEEE Int. Conf. on Circuits and Computers, ICCC'80, Port Chester, NY, Oct. 1980, pp. 1072–1074
19. R. W. Newcomb: *Neural – type microsystems – circuit status.* Proc. IEEE Int. Symp. on Circuits and Systems, Chicago, IL, Apr. 1981, pp. 97–100.
20. K. M. Passino, M. A. Sartori, P. J. Antsaklis: *Neural computing for numeric–to–symbolic conversion in control systems.* IEEE Control Systems Magazine, 1989, vol. 7, no 3, pp. 44–52
21. F. P. Preparata, D. T. Lee: *Computational geometry – a survey.* IEEE Trans. Comp., 1984, vol. C–33, no 12, pp. 1072–1100
22. K. Preston, Jr., M. J. B. Duff: *Modern Cellular Automata: Theory and Applications.* Plenum, New York, 1984
23. J. La Salle, S. Lefschetz: *Zarys teorii stabilności Lapunowa i jego metody bezpośredniej.* Warszawa: PWN, 1966
24. J. M. Skowroński: *Elementy dynamiki geometrycznej.* Warszawa: PWN, 1972
25. *Special Section on Neural Networks for Systems and Control.* IEEE Contr. Syst Mag., 1988, vol. 8, nr 2
26. S. Wolfram (Eds): *Theory and Applications of Cellular Automata.* World Scientific, New York, 1986
27. J. M. Zurada, M. J. Kang: *Summing networks using neural optimization concept.* Electr. Lett., 1988, vol. 24, no 10, pp. 616–617

A. SZATKOWSKI

NEURAL CELLULAR SYSTEMS SYNTHESIS FOR DECISION MAKING PART I: SELECTION - IDENTIFICATION SYSTEMS

Summary

Neural cellular systems are nonlinear continuous – time or discrete – time systems for information processing. These are the large – scale integrated electronic circuits as well as specialized systems for control, data selection and image processing. The structure of a cellular system is assured by the use of selection functors which are the basic building elements. In the Part I there has been proposed a general structure of neural cellular system for selection – identification tasks, with 1-dimensional selection functors. A number of results concerning dynamical selection – identification systems have been enclosed. Electronic implementation has been proposed.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1992, 38, z. 2, ss. 487–512

621.382.066

Synteza neuralnych systemów komórkowych dla zagadnień podejmowania decyzji

Część II: Systemy mnogościowo–logiczne oraz ogólne komórkowe sieci neuralne

ANDRZEJ SZATKOWSKI

Instytut Technologii Elektronicznej, Politechnika Gdańska

Otrzymano 1990. 07. 12.

Autoryzowano do druku 1991. 05. 10

W części I pracy została zaproponowana pewna uniwersalna struktura komórkowego systemu neuralnego selekcji – identyfikacji z wykorzystaniem układu dynamicznego funktora 1-wymiarowego. Komórkowe systemy neuralne mnogościowo–logiczne, którym poświęcona jest cz. II są rozwinięciem systemów omawianych w cz. I. Stanowią one realizację określonych funktorów mnogościowo–logicznych. Przedstawiony jest też ogólny model matematyczny komórkowej sieci neuralnej, z funktorami 1-wymiarowymi jako elementami podstawowymi i dowolną nieliniową siecią funkcyjną realizującą sprzężenia.

1. WSTĘP

W części I zaproponowana została pewna uniwersalna struktura komórkowego systemu neuralnego selekcji – identyfikacji z wykorzystaniem układu dynamicznego funktora 1-wymiarowego. Przedstawiono też szereg uwag ogólnych dotyczących systemów dynamicznych selekcji – identyfikacji, jak również rozważano zagadnienia związane z implementacją elektroniczną.

Komórkowe systemy neuralne mnogościowo – logiczne, którym poświęcona jest część II pracy są rozwinięciem systemów omawianych w cz. I. Rozwój techniki sieci neuralnych wpłynął w istotny sposób na zmiany w metodologii projektowania systemów komputerowych, co znalazło swój wyraz między innymi w szerszym powiązaniu z koncepcjami teorii zbiorów i logiki wielowartościowej.

Elementem łączącym algebrę zbiorów z algebrą logiki (dwuwatrościowej) jest funktor mnogościowo – logiczny $F(\cdot)$ opisany następująco:

$$\left.\begin{array}{c} F(a \in A) = 1 \text{ prawda} \\ \text{i} \\ F(a \notin A) = 0 \text{ nieprawda} \end{array}\right\} \tag{1}$$

W teorii systemów funktor $F(\cdot)$ jest realizowany w postaci odpowiedniego systemu selekcji – identyfikacji, na przykład w postaci odpowiedniego systemu dynamicznego φ_A, którego schemat syntezy został przedstawiony w części I pracy.

Funktor mnogościowo – logiczny (1) jest też jednym z elementów podstawowych reprezentowanych w strukturze systemów działających w ramach logiki zero – jedynkowej. Jest tam jednak rozważany w postaci szczególnej, ze względu na ściśle określone kryteria doboru położenia i rozmiarów podzbiorów A. W przedstawionym w pracy ujęciu obszary selekcyjne są dowolnie ustalonymi podzbiorami przestrzeni zmiennych – danych R^n. Funktor mnogościowo – logiczny jest przyjmowany w następującym sformułowaniu.

Dana jest rodzina $U_{(1)}, U_{(2)}, \ldots, U_{(m)}$ podzbiorów R^n oraz pewien zbiór (sekwencja) $\mathscr{F}$ operacji mnogościowych, które przypisują rodzinie $\{U_{(k)}$ pewien podzbiór U w R^n. Funktor mnogościowo – logiczny jest odwzorowaniem

$$F(x^0 \in \mathscr{F}(U_{(1)}, \ldots, U_{(m)}))$$

przypisującym elementom $x^0, U_{(1)}, \ldots, U_{(m)}$ liczbę 1 lub 0, w zależności od spełnienia relacji przynależności punktu $x^0 \in R^n$ do zbioru $\mathscr{F}(U_{(1)}, \ldots, U_{(m)})$. System dynamiczny mnogościowo – logiczny stanowi realizację systemową funktora mnogościowo – logicznego.

W rozdziale 2 przedstawiona jest realizacja dla podstawowych funktorów mnogościowo – logicznych. Elementem podstawowym konstruowanych systemów jest układ dynamiczny φ_ε^* 1-wymiarowego funktora selekcji – identyfikacji.

Zaproponowany w rozdziale 3 model matematyczny neuralnej sieci komórkowej obejmuje swoim zakresem wszystkie rozważane wcześniej w pracy systemy. Sformułowano odpowiednie warunki zapewniające posiadanie przez system podstawowej cechy neuralnego systemu komórkowego wyrażającej się w realizacji odwzorowania (w stanie ustalonym) continuum stanów początkowych w dyskretny zbiór asymptotycznie stabilnych stanów równowagi, rozważanych jako stany wyjściowe systemu.

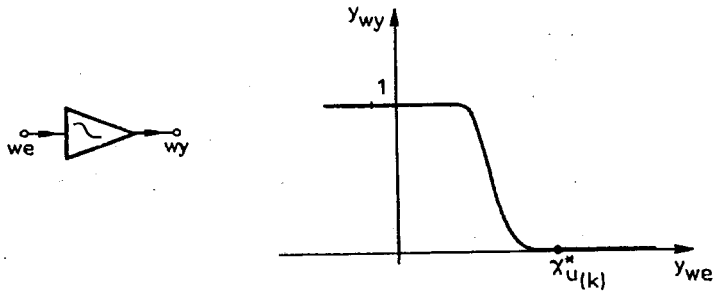
2. REALIZACJA PODSTAWOWYCH SYSTEMÓW DYNAMICZNYCH MNOGOŚCIOWO – LOGICZNYCH Z WYKORZYSTANIEM FUNKTORA φ_ε^*

A. Realizacja systemowa funktora mnogościowo – logicznego

$$F(x^0 \in \bigcup_{k=1}^{m} U_{(k)}) \tag{2}$$

Dana jest rodzina $U_{(1)}, U_{(2)}, \ldots, U_{(m)}$ podzbiorów otwartych, spójnych i ograniczonych w R^n. Odnośnie podzbiorów $U_{(k)}$ nie zakłada się, że są parami rozłączne. Dla

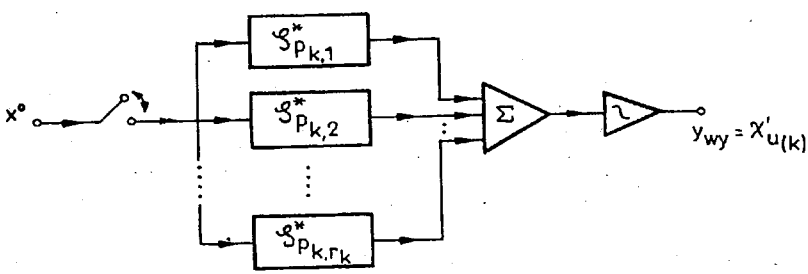
każdego z obszarów $U_{(k)}$, $k = 1, 2, \ldots, m$, dokonuje się jego aproksymacji odpowiednią rodziną przedziałów otwartych P_{k,j_k}, $j_k = 1, 2, \ldots, r_k$, w R^n według formuły (24) lub (25). Dla każdego z przedziałów P_{k,j_k} konstruowany jest system $\varphi_{P_{k,j_k}}$ według schematu przedstawionego na Rys. 13, cz. I pracy, a następnie dla każdego z podzbiorów $U_{(k)}$ otrzymuje się odpowiedni system $\varphi_{U_{(k)}}$ lub $\varphi'_{U_{(k)}}$, w zależności od zastosowanej formuły aproksymacyjnej (24) lub (25).



Rys. 1. Transmitancja elementu dołączonego na wyjściu systemu $\varphi_{U_{(k)}}$. $\chi^*_{U_{(k)}}$ oznacza wartość progową daną wzorem (33). $x^0 \in U_{(k)}$ wtedy, i tylko wtedy, gdy $y_{wy}(t) = 1$, dla pewnej chwili czasu $t \geqslant 0$

Schemat systemu $\varphi_{U_{(k)}}$ jest identyczny ze schematem systemu φ_U przedstawionym na Rys. 14.a (cz. I), uzupełnionym na wyjściu elementem o transmitancji jak na Rys. 1, lub jest zastosowany schemat z Rys. 14.b, bez użycia elementu mnożącego.

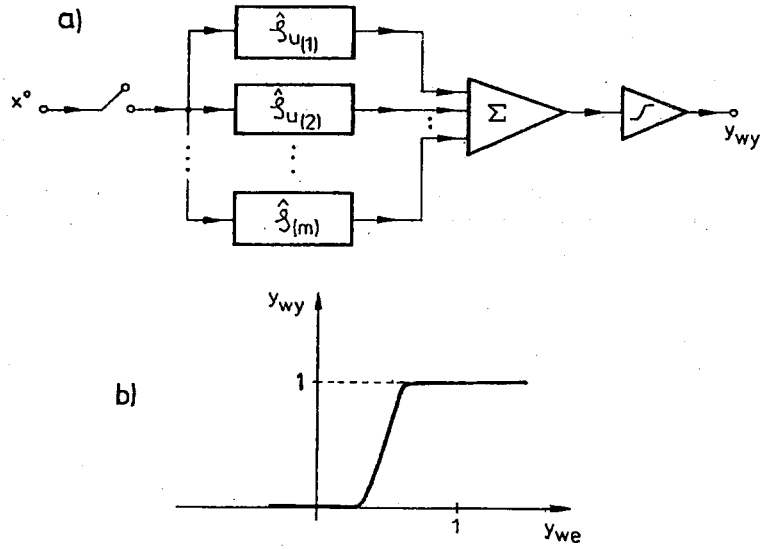
Schemat systemu $\varphi'_{U_{(k)}}$ (odpowiadający aproksymacji (25) podzbioru $U_{(k)}$) jest przedstawiony na Rys. 2.



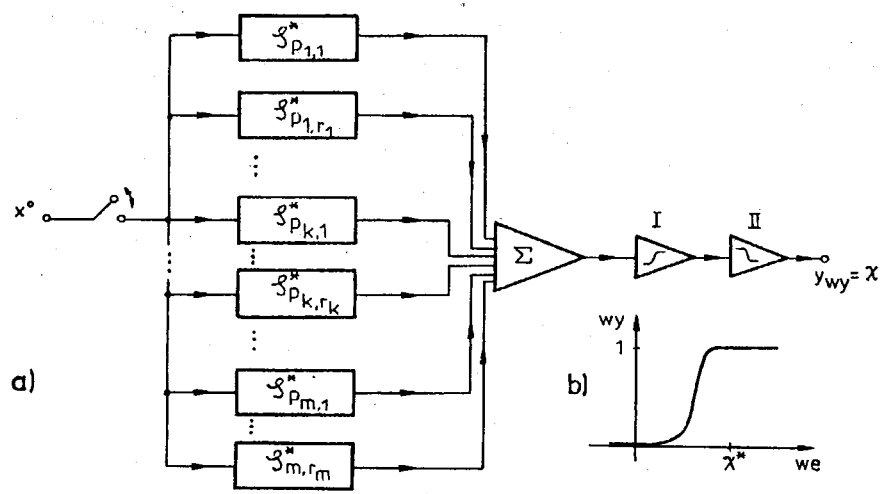
Rys. 2. Realizacja systemu $\varphi'_{U_{(k)}}$. $x^0 \in U_{(k)}$ wtedy, i tylko wtedy, gdy $y_{wy}(t) = \tilde{\chi}'_{U_{(k)}}(t) = 1$, dla pewnej chwili czasu $t \geqslant 0$. Wartość progowa $\chi'_{U_{(k)}}$ dla elementu normującego na wyjściu jest dana wzorem (34). φ^*_P oznacza część systemu φ_P odgraniczoną na Rys. 13 linią przerywaną

Konstrukcja całości systemu φ^U realizującego zadanie funktora mnogościowo – logicznego (2) jest podana na Rys. 3, gdzie $\hat{\varphi}_{U_{(k)}}$ oznacza system $\varphi_{U_{(k)}}$ lub $\varphi'_{U_{(k)}}$.

Należy odnotować, że zarówno przy zastosowaniu aproksymacji (24) (sumacyjnej) jak i aproksymacji (25) (iloczynowej) podzbiorów $U_{(1)}, U_{(2)}, \ldots, U_{(m)}$, konstrukcję systemu dynamicznego realizującego zadanie funktora (2) otrzymuje się tu z wykorzystaniem elementu normującego o charakterystyce sigmoidalnej.



Rys. 3. $x^0 \in \bigcup_{k=1}^{m} U_{(k)}$ wtedy, i tylko wtedy, gdy $y_{wy}(t)=1$, dla pewnego $t \geqslant 0$. Na rys. b przedstawiona jest charakterystyka elementu normującego na wyjściu systemu

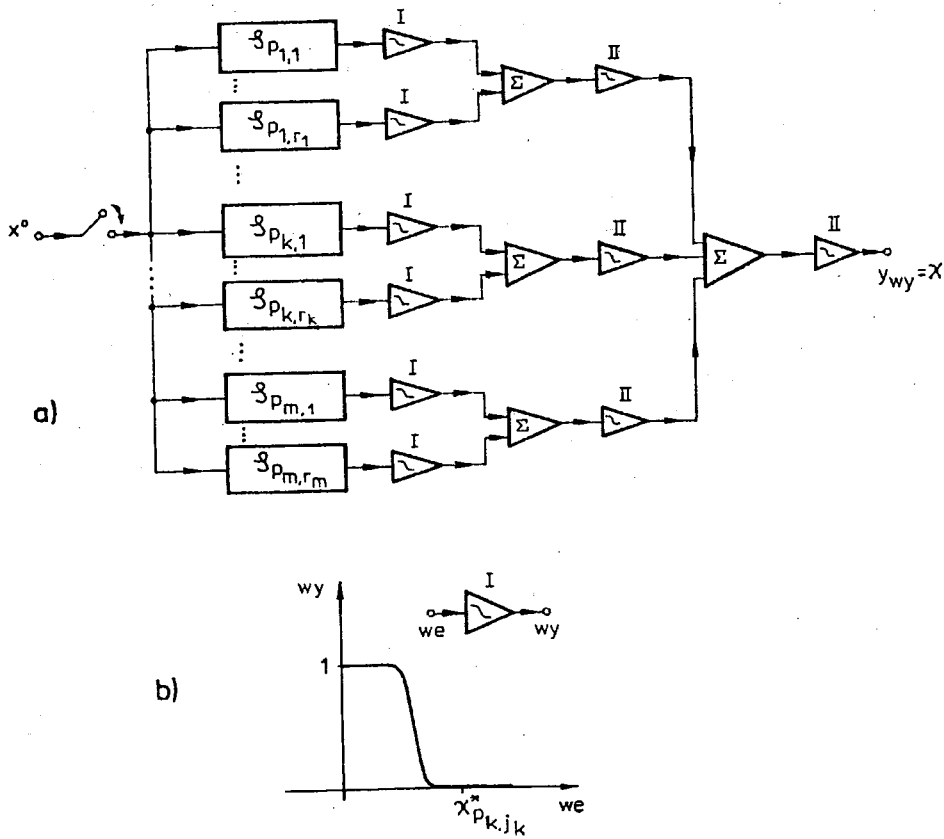


Rys. 4. Realizacja systemowa dla funktora mnogościowo-logicznego (3). $x^0 \in \bigcap_{k=1}^{m} U_{(k)}$ (odpowiednio, $x^0 \notin \bigcap_{k=1}^{m} U_{(k)}$) wtedy, i tylko wtedy, gdy $\tilde{\chi}(t)=1$, (odpowiednio $\tilde{\chi}(t)=0$), dla pewnej chwili czasu t. Charakterystyka elementu normującego I jest przedstawiona na rys. b, gdzie χ^* jest wartością progową daną wzorem $\chi^* = \min_{k=1,2,\ldots m;\ j_k=1,2,\ldots,r_k;\ i=1,2,\ldots,n} \min \min \delta_{j_k,i}$ (zob, wzór (34) w części I). Charakterystyka elementu normującego II jest dana na rys. 3.b

Jeżeli podzbiory $U_{(1)}, \ldots, U_{(m)}$ zadane są za pośrednictwem układów więzów nierównościowych jak w (35), cz. I pracy, system φ^U otrzymuje się w postaci jak na Rys. 18.a, zastępując tam elementy formujące z indeksem $k > 1$, dołączone do wyjść kolejnych sumatorów, elementami z indeksem $k = 1$ oraz dołączając na wyjściu układu element normujący o charakterystyce jak na Rys. 3.b. W tym przypadku nie jest wymagana ograniczoność podzbiorów $U_{(k)}$.

B. Realizacja systemowa funktora mnogościowo – logicznego

$$F(x^0 \in \bigcap_{k=1}^{m} U_{(k)}), \tag{3}$$



Rys. 5. Schemat systemu φ^n realizującego zadanie syntezy funktora (3) przy aproksymacji sumacyjnej (24) obszarów $U_{(k)}$, $k=1,2,\ldots,m$. $x^0 \in \bigcap_{k=1}^{m} U_{(k)}$ (i odpowiednio, $x^0 \notin \bigcap_{k=1}^{m} U_{(k)}$) wtedy, i tylko wtedy, gdy $\tilde{\chi}(t)=1$ (odpowiednio, $\tilde{\chi}(t)=0$) dla pewnej chwili czasu $t \geqslant 0$. Na rys. b przedstawiono wykres charakterystyki elementu normującego I. Wartość progowa $\chi^*_{P_{k,j_k}}$ wynosi $\min_{i=1,2,\ldots n} \delta_{j_k,i}$. Charakterystyka elementu normującego II jest jak na rys. 3.b

gdzie $U_{(1)}, \ldots, U_{(m)}$ są podzbiorami otwartymi, spójnymi i ograniczonymi w R^n.

Przy zastosowaniu do każdego z obszarów $U_{(k)}$ aproksymacji iloczynowej (25) kryterium selekcyjne

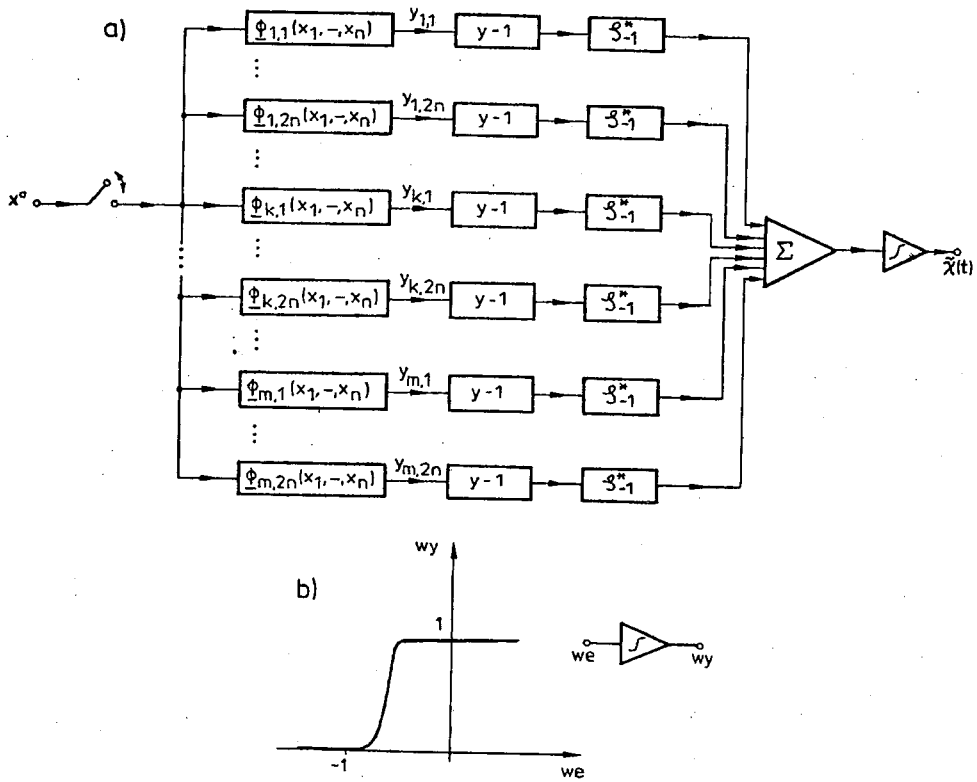
$$x^0 \in \bigcap_{k=1}^{m} U_{(k)} \tag{4}$$

przyjmuje postać

$$x^0 \in \bigcap_{k=1}^{m} \bigcap_{j_k=1}^{r_k} P_{k,j_k}.$$

Schemat systemu φ^n realizującego zadanie funktora mnogościowo — logicznego (3) przedstawia Rys. 4 (znaczenie symbolu $\varphi^*_{P_{k,j_k}}$ jest jak na schemacie na Rys. 13).

Natomiast przy zastosowaniu aproksymacji sumacyjnej (24) dla każdego z obszarów $U_{(k)}$ otrzymuje się dla funktora (3) realizację systemową: podaną na Rys. 5.



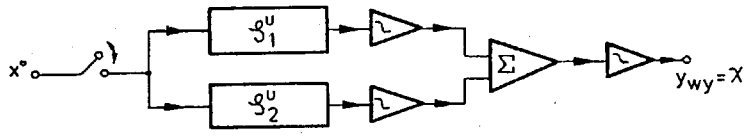
Rys. 6. a. Schemat systemu φ^n otrzymany dla przypadku opisu podzbiorów $U_{(1)},\ldots,U_{(m)}$ za pośrednictwem układu więzów nierównościowych (35). $x^0 \in \bigcap_{k=1}^{m} U_{(k)}$ (odpowiednio, $x^0 \notin \bigcap_{k=1}^{m} U_{(k)}$) wtedy, i tylko wtedy, gdy $\tilde{\chi}(t)=1$ (odpowiednio, $\tilde{\chi}(t)=0$) dla pewnej chwili czasu t. Na rys. b przedstawiono charakterystykę elementu normującego załączonego na wyjściu układu

Jeżeli podzbiory $U_{(1)}, \ldots, U_{(m)}$ zadane są za pośrednictwem układu więzów nierównościowych (35), cz. I pracy, strukturę systemu φ^n otrzymuje się jak na Rys. 6. W tym przypadku nie jest wymagana ograniczoność podzbiorów $U_{(1)}, \ldots, U_{(m)}$.

C. Realizacja systemowa funktora mnogościowo – logicznego

$$F\left(x^o \in \left(\bigcup_{j=1}^{m} U_{(j)}\right) \cap \left(\bigcup_{l=1}^{m_2} U_{(l)}\right)\right). \tag{5}$$

W pierwszej wersji realizowane są systemy φ_1^n i φ_2^n według jednego ze schematów podanych w punkcie A. Schemat całości systemu φ realizującego postawione zadanie przedstawia Rys. 7. Charakterystyki elementów normujących są jak na wykresie na Rys. 3. b. Należy przy tym odnotować, że w konkretnej ustalonej realizacji podsystemów φ_1^n i φ_2^n niektóre pary elementów normujących mogą być zredukowane.



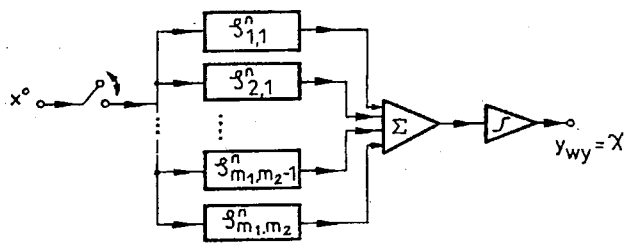
Rys. 7. Schemat systemu φ realizującego zadanie funktora (5)

W wersji drugiej korzysta się z wyrażenia po prawej stronie tożsamości

$$\left(\bigcup_{j=1}^{m_1} U_{(j)}\right) \cap \left(\bigcup_{l=1}^{m_2} U_{(l)}\right) = \bigcup_{j=1}^{m_1} \bigcup_{l=1}^{m_2} \left(U_{(j)} \cap U_{(l)}\right). \tag{6}$$

Każdy z obszarów $U_{(j)}, U_{(l)}, j = 1, 2, \ldots, m_1, l = 1, 2, \ldots, m_2$, jest aproksymowany zgodnie z formułą iloczynową (25). Dla otrzymanej w wyniku aproksymacji iloczynowej każdego z podzbiorów $U_{(j)} \cap U_{(l)}$ konstruowany jest odpowiedni system $\varphi^n \overset{\text{ozn}}{=} \varphi_{j,l}^n$ według schematu z Rys. 4 (w tym przypadku $m = 2$; odnośnie podzbiorów $U_{(j)}$, $U_{(l)}$ zakłada się, że są ograniczone). Jeżeli natomiast obszary $U_{(j)}$, $U_{(l)}$ zadane są za pośrednictwem układu więzów nierównościowych (35), cz. I pracy, dla syntezy systemu $\varphi_{j,l}^n$ stosuje się odpowiedni schemat jak na Rys. 6.

Schemat całości systemu φ realizującego postawione zadanie przedstawia Rys. 8.

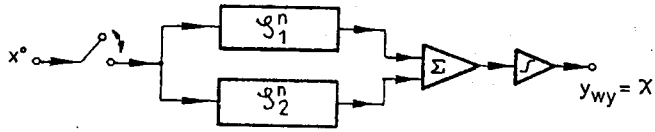


Rys. 8. Wersja konstrukcji systemu φ realizującego zadanie syntezy funktora (5) według formuły równoważnej (6). Charakterystyka elementu normującego jest jak na wykresie na rys. 3.b

D. Realizacja systemowa funktora mnogościowo — logicznego

$$F\left(x^0 \in \left(\bigcap_{j=1}^{m_1} U_{(j)}\right) \cup \left(\bigcap_{l=1}^{m_2} U_{(l)}\right)\right). \tag{7}$$

W pierwszej wersji realizowane są systemy φ_1^n i φ_2^n według schematu z Rys. 4, dla aproksymacji iloczynowej każdego z podzbiorów $U_{(j)}$, $U_{(k)}$, lub według schematu z Rys. 5, dla aproksymacji sumacyjnej tych podzbiorów. Jeżeli natomiast podzbiory $U_{(j)}$, $U_{(l)}$ dane są za pośrednictwem układu więzów nierównościoych postaci (35), cz. I pracy, systemy φ_1^n i φ_2^n otrzymuje się zgodnie ze schematem z Rys. 6. Schemat całości systemu φ realizującego postawione zadanie przedstawia Rys. 9.

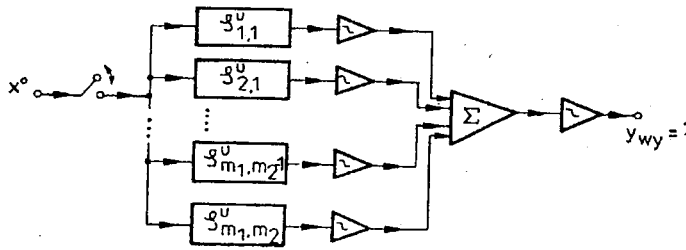


Rys. 9. Schemat systemu φ realizującego zadanie syntezy funktora (7)

W wersji drugiej korzysta się z wyrażenia po prawej stronie tożsamości

$$\left(\bigcap_{j=1}^{m_1} U_{(j)}\right) \cup \left(\bigcap_{l=1}^{m_2} U_{(l)}\right) = \bigcap_{j=1}^{m_1} \bigcap_{l=1}^{m_2} (U_{(j)} \cup U_{(l)}). \tag{8}$$

Każdy ze zbiorów $U_{(j)}$, $U_{(l)}$, $j = 1, 2, \ldots m_1$, $l = 2, \ldots m_2$, jest aprokymowany zgodnie z formułą sumacyjną (24). Dla otrzymanej w wyniku aproksymacji sumacyjnej dla każdego z podzbiorów $U_{(j)} \cup U_{(l)}$ konstruowany jest odpowiedni system $\varphi^n \overset{\text{oznacza}}{=} \varphi_{j,l}^n$ według schematu z Rys. 14.b (cz. I pracy). Schemat całości systemu φ realizującego postawione zadanie przedstawia Rys. 10.



Rys. 10. Schemat systemu φ realizującego zadanie syntezy funktora (7) według formuły równoważnej. Elementy normujące mają charakterystyki jak na wykresie na rys. 3.b

W syntezie systemów dynamicznych mnogościowo — logicznych mogą też być użyteczne tożsamości de Morgana, które prowadzą do następujących równoważnych sformułowań dla funktorów (2) i (3):

$$F\left(\bigcup_{k=1}^{m} U_{(k)}\right) = \sim F\left(\bigcap_{k=1}^{m} U'_{(k)}\right) \tag{9}$$

i

$$F\left(\bigcap_{k=1}^{m} U_{(k)}\right) = \sim F\left(\bigcup_{k=1}^{m} U'_{(k)}\right), \tag{10}$$

gdzie $U'_{(k)} \triangleq R^n \setminus U_{(k)}$. Realizację systemową funktora

$$F(x^0 \in U')$$

otrzymuje się przez dołączenie na wyjściu systemu realizującego $F(x^0 \in U)$ elementu negacji logicznej (o charakterystyce jak na Rys. 3.b).

3. OGÓLNE DYNAMICZNE SYSTEMY KOMÓRKOWE TYPU NEURALNEGO

W tym rozdziale zostanie przedstawiony model ogólny sieci komórkowej typu neuralnego, zawierający poprzednio rozważane systemy jako przypadki szczególne. Model ten jest w kontekście danych literaturowych jednym z najbardziej ogólnych sformułowań w tym zakresie (zob. np. [5, 6], [12], [14]).

Rozważania dotyczą modelu matematycznego ogólnego systemu komórkowego z układem dynamicznym funktora 1-wymiarowego selekcji – identyfikacji jako elementem podstawowym, wyróżniającym dla tej klasy systemów , jak też w dalszej części ujmują zagadnienie realizacji elektronicznej w sensie sformułowania ogólnego modelu komórkowej sieci neuralnej zrealizowanej w postaci układu (systemu) elektronicznego.

W rozdziale 3, cz. I pracy, przedstawiono odpowiedni opis dla trzech układów dynamicznych selekcji – identyfikacji $\varphi^*_{\mp\varepsilon}$, $\varphi^*_{\mp\varepsilon}$ i φ^*_{ε}, zastosowanych jako elementy podstawowe w syntezie rozważanych sieci. W zasadzie w procedurze syntezy niezbędny jest jeden z elementów $\varphi^*_{\mp\varepsilon}$ lub $\varphi^*_{\mp\varepsilon}$. Pozostałe dwa można bowiem w prosty sposób otrzymać na podstawie $\varphi^*_{\mp\varepsilon}$ lub $\varphi^*_{\mp\varepsilon}$. Odpowiednie funkcje energii dla φ^*_{ε}, $\varphi^*_{\mp\varepsilon}$ i $\varphi^*_{\mp\varepsilon}$ zostały przedstawione na Rys. 9, 11.a i 11.b, rozdział 3, cz. I pracy. W rozdziale 4, p. A, podano natomiast realizację elektroniczną dla tych funktorów. Uwzględniono przy tym możliwość bezpośredniego kształtowania rozmiarów obszarów selekcyjnych.

Równanie stanu dla funktora φ^* ma ogólnie postać

$$\dot{x} = -\nabla h^*(x). \tag{11}$$

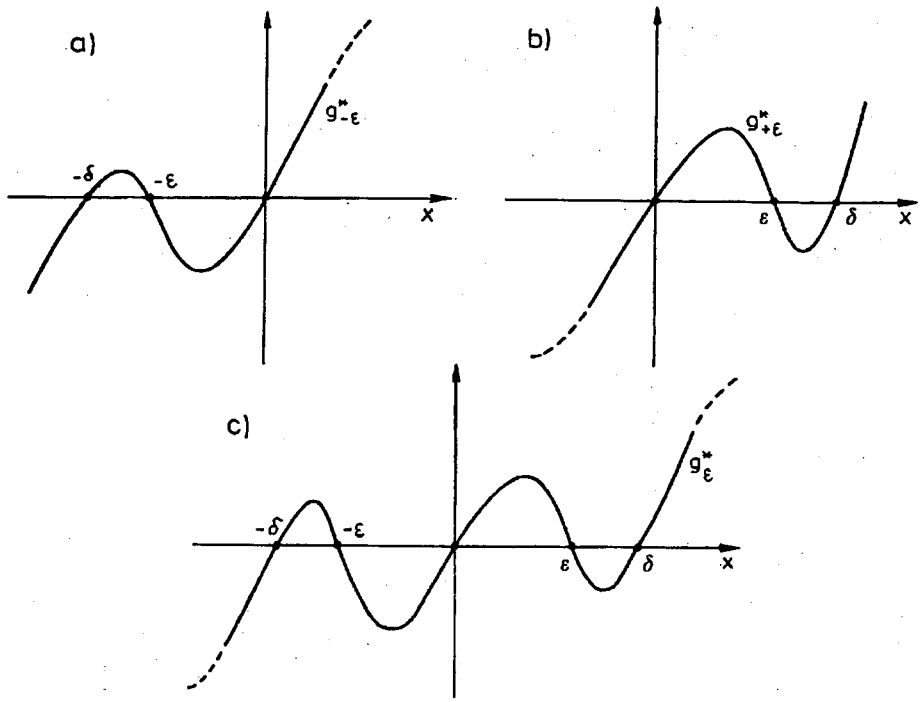
gdzie $h^*(\cdot)$ jest jedną z funkcji $h^*_{\mp\varepsilon}(\cdot)$, $h^*_{\mp\varepsilon}(\cdot)$ lub $h^*_{\varepsilon}(\cdot)$, w zależności od typu funktora.

Kładąc

$$g^*(x) = \nabla h^*(x) \tag{12}$$

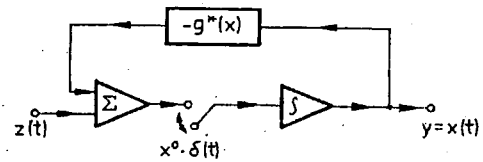
otrzymuje się inny zapis równania stanu (11),

$$\dot{x} = -g^*(x). \tag{13}$$



Rys. 11. Wykresy charakterystyk $g^*_{-\varepsilon}(\cdot)$, $g^*_{+\varepsilon}(\cdot)$ i g^*_{ε} dla funktorów $\varphi^*_{-\varepsilon}$, $\varphi^*_{+\varepsilon}$ i φ^*_{ε}, odpowiednio

Na Rys. 11 przedstawiony jest przebieg charakterystyk $g^*_{-\varepsilon}(\cdot)$, $g^*_{+\varepsilon}(\cdot)$ i $g^*_{\varepsilon}(\cdot)$ dla odpowiednich funktorów. Z uwzględnieniem połączenia z pozostałą częścią sieci, schemat funkcjonalny elementu φ^* jest przedstawiony na Rys. 12.



Rys. 12. Schemat funkcjonalny elementu φ^*. $x^0 \cdot \delta(t)$ oznacza sygnał zadający wartość początkową $x(0) = x^0$ dla zmiennej stanu x, natomiast $z(t)$ jest sygnałem (1-wymiarowym) pochodzącym z pozostałej części sieci, z którą element jest incydentny. Na schemacie nie uwzględniono możliwości zmiany parametru ε funktora

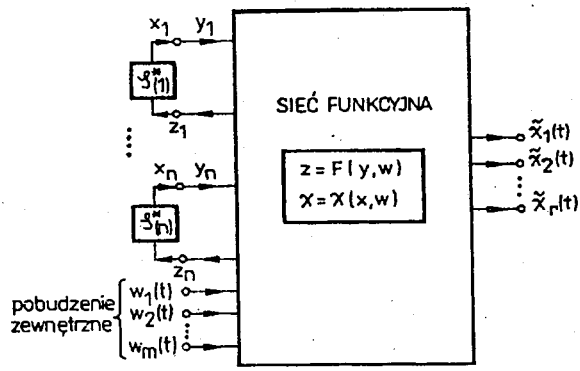
Równanie dynamiczne elementu, z uwzględnieniem zewnętrznego sygnału wejściowego, ma postać:

$$\dot{x} = -g^*(x) + z(t). \tag{14}$$

Na Rys. 13 podano reprezentację przepływowo – wrotnikową systemu komórkowego z funktorami φ^* jako elementami wyróżnionymi, dołączonymi do wrót sieci funkcyjnej.

Równanie dynamiczne sieci z Rys. 13 ma postać:

$$\dot{x}_i = -g_i^*(x_i) + F_i(x, w(t)), \; i = 1, 2, \ldots, n, \tag{15}$$



Rys. 13. Schemat sieci komórkowej z wyróżnionymi elementami $\varphi^*_{(1)},\ldots,\varphi^*_{(n)}$ typu $\varphi^*_{-\varepsilon}$, $\varphi^*_{+\varepsilon}$ lub φ^*_ε oraz wydzieloną częścią funkcyjną. Węzły $w_1,\ldots,w_m$ są węzłami wejściowymi sieci, natomiast $\chi_1,\ldots,\chi_r$ są węzłami wyjściowymi. Równanie $\chi = \chi(x,w)$ jest równaniem wyjścia systemu, natomiast równanie $z = F(y,w)$ opisuje charakterystykę sieci funkcyjnej z punktu widzenia wrót. Ponadto $y = x$

Odnotować tu należy, że przy dowolnym wyborze funkcji $F(\cdot)$ opisującej charakterystykę części bezinercyjnej sieci, klasa równań (15) obejmuje swoim zakresem wszystkie nieautonomiczne systemy różniczkowe. Można przy tym uogólnić jeszcze rozważania przyjmując, że charakterystyka sieci funkcyjnej jest pewną podrozmaitością różniczkowalną w $R^n \times R^n \times R^m$.

Aby zapewnić posiadanie przez sieć przedstawioną na Rys. 13, o równaniu stanu (15), właściwości charakterystycznych dla neuralnych systemów komórkowych należy przyjąć dodatkowe założenia odnośnie klasy dopuszczalnych w tym kontekście charakterystyk sieci funkcyjnej. Podstawową cechą wymaganą w stosunku do rozważanej sieci jest rozdział jej przestrzeni stanu na rodzinę obszarów asymptotycznej stabilności [5]–[7], [11], [12], [14]. Powinno przy tym istnieć odwzorowanie wzajemnie jednoznaczne (a nawet homeomorficzne) $\Psi(\cdot)$ przestrzeni stanu R^n na siebie takie, że:

A. $\Psi(\cdot)$ odwzorowuje zbiór punktów równowagi systemu zredukowanego

$$\dot{x}_i = -g_i^*(x_i), \; i = 1, 2, \ldots, n, \tag{16}$$

na zbiór punktów równowagi systemu (15). Z założenia więc, układy (15) i (16) posiadają tę samą liczbę punktów równowagi.

B. Odpowiadające sobie w odwzorowaniu $\Psi(\cdot)$ punkty równowagi mają ten sam typ stabilności i ponadto obszar przyciągania asymptotycznie stabilnego punktu równowagi $\Psi(x^*)$ systemu (15), gdzie x^* jest pewnym asymptotycznie stabilnym punktem równowagi systemu (16), jest dany jako obraz obszaru przyciągania punktu równowagi x^* poprzez homeomorfizm $\Phi(\cdot)$.

System zredukowany, opisany równaniami (16), otrzymuje się wydzielając z sieci na Rys. 13 jej część złożoną z elementów funktorów $\varphi^*_{(1)}, \ldots, \varphi^*_{(n)}$. Dynamikę systemu zredukowanego można w pełni przedstawić w następujących punktach:

a. punkty $x^* = (x_1^*, \ldots, x_n^*)$ o współrzędnych $x_i^* = 0$, $x_i^* = -\delta$ lub $x_i^* = \delta_i$ są asymptotycznie stabilnymi punktami równowagi systemu (jeżeli element $\varphi^*_{(i)}$ jest funktorem jednostronnym $\varphi^*_{\pm\varepsilon}$ lub $\varphi^*_{+\varepsilon}$ to w powyższym wykazie pomijamy wartości $x_i^* = \delta_i$ lub $x_i^* = -\delta_i$, odpowiednio);

b. każdy punkt $x^* = (x_1^*, \ldots, x_n^*)$, o współrzędnych przyjmujących wartości jak w punkcie a., i którego jedna ze współrzędnych wynosi $-\varepsilon_i$ lub $+\varepsilon_i$ (przy zastrzeżeniach wynikających z typu funktora $\varphi^*_{(i)}$) jest niestabilnym punktem równowagi systemu;

c. obszarem przyciągania asymptotycznie stabilnego punktu równowagi $x^* = (x_1^*, \ldots, x_n^*)$ jest (n-wymiarowy) przedział otwarty w R^n, którego krawędź wzdłuż osi x_i jest równa odpowiednio:

$$\left.\begin{array}{lll} (-\varepsilon_i, \varepsilon_i) & , \text{ gdy } \varphi^*_{(i)} = \varphi^*_{\varepsilon_i} & \\ (-\infty, \varepsilon_i) & , \text{ gdy } \varphi^*_{(i)} = \varphi^*_{+\varepsilon_j} & \\ (-\varepsilon_i, \infty) & , \text{ gdy } \varphi^*_{(i)} = \varphi^*_{-\varepsilon_i} & \end{array}\right\} \quad \text{gdy } x_i^* = 0$$

lub

$$(-\infty, -\varepsilon_i) \text{ , gdy } x_i^* = -\delta_i,$$

lub

$$(\varepsilon_i, \infty) \quad \text{, gdy } x_i^* = \delta_i.$$

Przestrzeń stanu systemu zredukowanego jest więc podzielona na przedziały będące obszarami przyciągania wymienionych w p. a. punktów równowagi. Hiperpłaszczyzny:

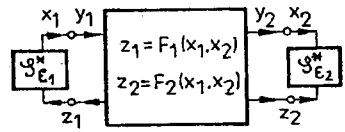
$$\left.\begin{array}{lll} x_i = \pm\,\varepsilon_i & , \text{ gdy } \varphi^*_{(i)} = \varphi^*_{\varepsilon_i} & \\ x_i = -\varepsilon_i & , \text{ gdy } \varphi^*_{(i)} = \varphi^*_{-\varepsilon_i} & \\ x_i = +\varepsilon_i & , \text{ gdy } \varphi^*_{(i)} = \varphi^*_{+\varepsilon_i} & \end{array}\right\} \quad i = 1, 2, \ldots, n,$$

wyznaczają w R^n zbiór $(n-1)$-wymiarowych ścian rozdzielających obszary asymptotycznej stabilności.

Powyższe uwagi odnośnie struktury przestrzeni stanu rozważanego modelu ogólnego neuralnej sieci komórkowej zilustrowane są przykładem dotyczącym sieci 2-wymiarowej zawierającej dwa funktory φ^*_ε, jak na schemacie na Rys. 14.

Równanie stanu systemu ma postać

$$\left.\begin{array}{l} \dot{x}_1 = -g^*_{\varepsilon_1}(x_1) + F(x_1, x_2) \\ \dot{x}_2 = -g^*_{\varepsilon_2}(x_2) + F_2(x_1, x_2) \end{array}\right\} \tag{17}$$

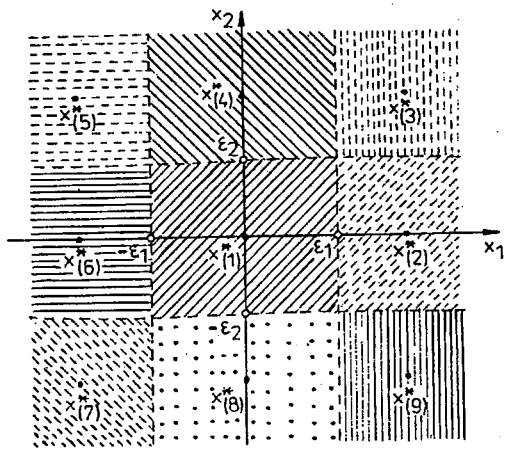


Rys. 14. Autonomiczny 2-wymiarowy neuralny system komórkowy

Równanie stanu systemu zredukowanego jest postaci

$$\dot{x}_1 = -g^*_{\varepsilon_1}(x_1),\ \dot{x} = -g^*_{\varepsilon_2}(x_2) \tag{18}$$

Na Rys. 15 przedstawiono strukturę przestrzeni stanu systemu (18) z zaznaczonymi obszarami przyciągania asymptotycznie stabilnych punktów równowagi.



Rys. 15. Asymptotycznie stabilnymi punktami równowagi systemu (18) są $x^*_{(1)}=(0,0)$, $x^*_{(2)}=(0,\delta_1)$, $x^*_{(3)}=(\delta_1,\delta_2)$, $x^*_{(4)}=(0,\delta_2)$, $x^*_{(5)}=(-\delta_1,\delta_2)$, $x^*_{(6)}=(-\delta_1,0)$, $x^*_{(7)}=(-\delta_1,-\delta_2)$, $x^*_{(8)}=(0,-\delta_2)$, $x^*_{(9)}=(\delta_1,-\delta_2)$

Aby zapewnić dla systemu przedstawionego na Rys. 14, o równaniu stanu (17) wymagań podanych w punktach A i B przyjmuje się następujące założenia odnośnie funkcji $F_1(\cdot)$, $F_2(\cdot)$ opisujących charakterystykę części bezinercyjnej:

1) $F_1(\cdot)$ i $F_2(\cdot)$ są klasy C^1 (to założenie zapewnia istnienie i jednoznaczność rozwiązań układu równań (17 [7], [11]), oraz

$$\lim_{x_1\to\infty} g^*_{\varepsilon_1}(x_1) - F_1(x_1, x_2^0) > 0, \quad \begin{matrix}\text{dla każdej}\\ \text{wartości } x_2^0,\end{matrix} \tag{19}$$

i

$$\lim_{x_2\to\infty} g^*_{\varepsilon_2}(x_2) - F_2(x_1^0, x_2) > 0, \quad \begin{matrix}\text{dla każdej}\\ \text{wartości } x_1^0;\end{matrix} \tag{20}$$

2) Dla każdego ustalonego x_2^0 równanie

$$g^*_{\varepsilon_1}(x_1) - F_1(x_1, x_2^0) = 0 \tag{21}$$

posiada dokładnie 5 różnych rozwiązań ze względu na x_1, oraz dla każdego ustalonego x_1^0 równanie

$$g^*_{\varepsilon_2}(x_2) - F_2(x_1^0, x_2) = 0 \tag{22}$$

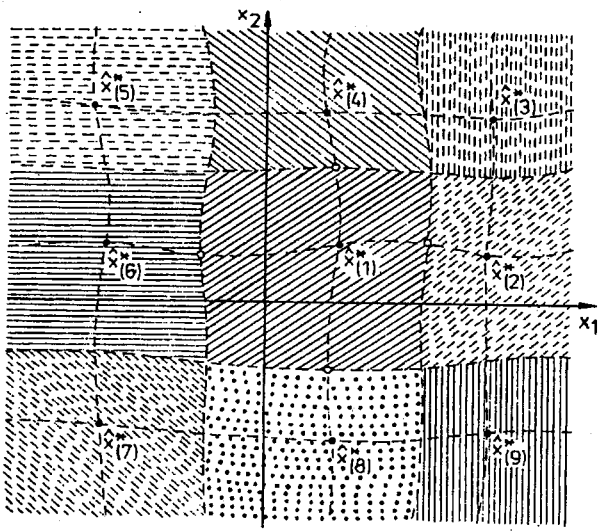
posiada dokładnie 5 różnych rozwiązań ze względu na x_2.

3. W każdym punkcie (x_1^*, x_2^*), w którym spełnione jest równanie (21) zachodzi

$$\left(\frac{\mathrm{d}g^*_{\varepsilon_1}}{\mathrm{d}x_1} - \frac{\partial F_1}{\partial x_1}\right) | (x_1^*, x_2^*) \neq 0, \tag{23}$$

oraz w każdym punkcie (x_1^*, x_2^*), w którym spełnione jest równanie (22) zachodzi

$$\left(\frac{\mathrm{d}g^*_{\varepsilon_2}}{\mathrm{d}x_2} - \frac{\partial F_2}{\partial x_2}\right) | (x_1^*, x_2^*) \neq 0. \tag{24}$$



Rys. 16. Przykład ilustrujący strukturę przestrzeni fazowej systemu (17), o schemacie jak na rys. 14, przy spełnieniu założeń wymienionych w punktach 1.–3. Punkty $\hat{x}^*_{(1)},\ldots,\hat{x}^*_{(9)}$ są asymptotycznie stabilnymi punktami równowagi systemu. Zaznaczono też obszary przyciągania poszczególnych punktów równowagi. Linie kreskowane są wyznaczone przez rozwiązania równań (21), (22)

Warunki wymienione w punktach 1. – 3. zapewniają posiadanie przez system (17), Rys. 14, cech właściwych dla neuralnych systemów komórkowych. Na Rys. 16 zilustrowano strukturę przestrzeni stanu systemu (17), z uwzględnieniem jej rozdziału na obszary asymptotycznej stabilności, przy założeniu spełnienia wymagań wymienionych w punktach 1. – 3. zapewniających podobieństwo obrazów fazowych dla systemu pełnego (17) i systemu zredukowanego (modelowwego) (16) w sensie istnienia odpowiedniego odwzorowania homeomorficznego $\Psi(\cdot)$ jak w punktach A. i B.

Uogólnienie rozważań dla dowolnej sieci komórkowej jak na Rys. 13, opisanej równaniem stanu (15), nie przedstawia trudności. Zakłada się tu autonomiczność

systemu, to znaczy $w(t) = const$. Przypadek systemu o dowolnym pobudzeniu zostanie rozważony w oddzielnym opracowaniu.

Następujące założenia zapewniają dla systemu autonomicznego (15), schemat z Rys. 13, spełnienie wymienionych w punktach A. B. wymagań odnośnie struktury jego przestrzeni fazowej (struktury portretu fazowego).

I. Funkcje $F_1(\cdot), \ldots, F_n(\cdot)$ są klasy C^1 oraz

$$\lim_{x_i \to \infty} g^*_{\varepsilon_i}(x_i) - F_i(x_1^0, \ldots, x_{i-1}^0, x_i, x_{i+1}^0, \ldots, x_n^0) > 0, \tag{25}$$

dla każdego $i = 1, 2, \ldots, n$ i każdego ustalonego ciągu x_j^0, wartości współrzędnych x_j, $j \neq i$.

II. Każde z równań

$$g^*_{\varepsilon_i}(x_i) - F_i(x_1^0, \ldots, x_{i-1}^0, x_i, x_{i+1}^0, \ldots, x_n^0) = 0, \tag{26}$$

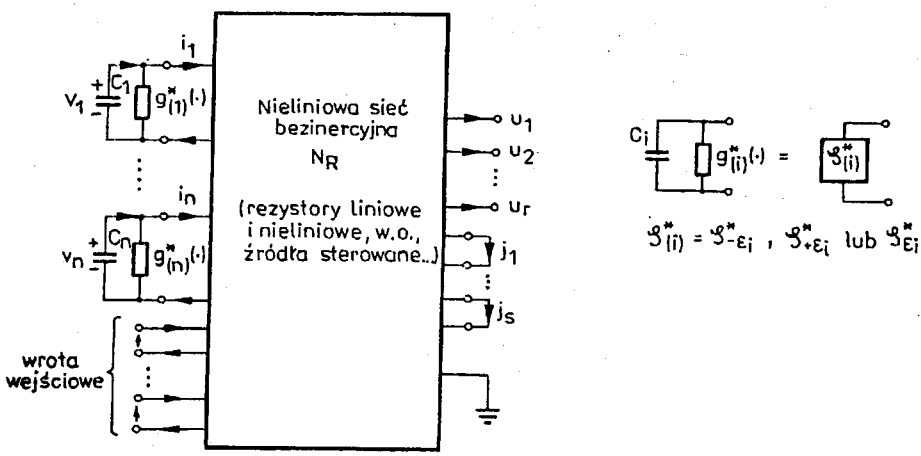
$i = 1, 2, \ldots, n$, posiada dla każdego ustalonego ciągu x_j^0 wartości współrzędnych x_j, $j \neq i$, dokładnie 5 różnych rozwiązań ze względu na x_i.

III. W każdym punkcie $(x_1^*, \ldots, x_n^*)$, w którym spełnione jest równanie (26), z ustalonym indeksem i, $i = 1, 2, \ldots, n$, zachodzi

$$\left(\frac{dg^*_{\varepsilon_i}}{dx_i} - \frac{\partial F_i}{\partial x_i}\right) | x^* \neq 0. \tag{27}$$

Rozwiązania każdego z równań (26), z ustalonym indeksem i oraz współrzędnych x_j, $j \neq i$, przyjmujących wszystkie możliwe wartości, wyznaczają 5 $(n-1)$-wymiarowych hiperpowierzchni w R^n. Każda z tych hiperpowierzchni jest wykresem ze względu na $(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$, to znaczy, jest parametryzowana przez $(n-1)$ zmiennych x_j, $j = 1, 2, \ldots, n$, $j \neq i$. $2n$ spośród wyznaczonych przez rozwiązania równań (26) hiperpowierzchni (w sumie otrzymuje się $5n$ hiperpowierzchni) wyznacza granice obszarów przyciągania asymptotycznie stabilnych punktów równowagi systemu (15). Lokalizację tych punktów równowagi wyznaczają punkty przecięcia pozostałych $3n$ hiperpowierzchni (por. ilustrację dla przypadku 2-wymiarowego przedstawioną na Rys. 16).

W powyższych rozważaniach przyjęto, że wszystkie funktory są typu obustronnego (symetrycznego). Warunki odnośnie sieci zawierającej dowolnego typu funktory otrzymuje się w analogiczny sposób. Na Rys. 17 przedstawiono schemat układu wielowrotnika stanowiącego realizację elektroniczną systemu komórkowego z Rys. 13. Elementy funktorów reprezentowne są przez odpowiednie dwójniki zgodnie z opisem podanym w p. A rozdziału 4, cz. I pracy (Rys. 19–26). Do wrót wejściowych dołączone są źródła niezależne, prądowe i napięciowe, stanowiące pobudzenie zewnętrzne dla sieci. Wektor wydajności źródeł pobudzających jest oznaczony symbolem w, $w = w(t)$. Po prawej stronie wielowrotnika zaznaczono wyróżnione węzły i gałęzie, których odpowiednie potencjały i prądy tworzą wektor zmiennych wyjściowych sieci.



Rys. 17. Realizacja elektroniczna systemu komórkowego

Przyjmując równanie wielowrotnika N_R w postaci

$$i = G(v, w), \tag{28}$$

gdzie $i = (i_1, \ldots, i_n), v = (v_1, \ldots, v_n), w = (w_1, \ldots, w_m)$ oraz $G(\cdot) = (G_1, \ldots, G_n)(\cdot)$, otrzymuje się równanie dynamiczne systemu

$$C\frac{dv}{dt} = -g^*(v) - G(v, w(t)), \tag{29}$$

($C = \mathrm{Diag}(C_1, \ldots, C_n)$, $g^*(v) = (g^*_{(1)}(v_1), \ldots, g^*_{(n)}(v_n))$).

Równanie stanu jest uzupełniane odpowiednim równaniem wyjścia określającym zależność zmiennych wyjściowych $u_1, \ldots, u_r, j_1, \ldots, j_s$ od zmiennej stanu v systemu i wektora pobudzeń w.

Równanie dynamiczne (29) sieci ma taką samą postać jak równanie stanu dla ogólnej sieci komórkowej z Rys. 13. W szczególności więc mają tu zastosowanie wszystkie podane uprzednio uwagi dotyczące portretu fazowego systemu ogólnego, w tym też podane w punktach I – III warunki zapewniające posiadanie przez rozważaną sieć cech charakterystycznych dla neuralnych systemów komórkowych. Można tu też rozważyć ogólniejszy typ sieci przyjmując, że charakterystyka N_R wielowrotnika bezinercyjnego jest pewną podrozmaitością różniczkowalną w $R^n \times R^n \times R^m$.

Z drugiej strony natomiast, jeżeli dla elementu $g^*(\cdot)$ zostanie przyjęty model odcinkowo–liniowy , jak np. na Rys. 22, rozdział 4, cz. I pracy, oraz jeżeli klasa elementów tworzących sieć bezinercyjną zostanie ograniczona do elementów o charakterystykach odcinkowo–liniowych (rezystory liniowe, elementy o charakterystykach sigmoidalnych aproksymowanych odcinkami etc.), to w badaniu zachowania się trajektorii systemu mogą być wykorzystane metody analizy odcinkowo–liniowej.

4. PRZYKŁAD SYSTEMU NEURALNEGO DLA REALIZACJI ZADANIA WYZNACZANIA WARTOŚCI MINIMALNEJ I MAKSYMALNEJ FUNKCJI WIELU ZMIENNYCH

W rozdziale tym zostanie przedstawiona konstrukcja systemu elektronicznego realizującego zadanie wyznaczania wartości minimalnej i maksymalnej funkcji wielu zmiennych. Podstawowy blok wchodzący w skład systemu posiada cechy właściwe rozważanym w tej pracy systemom neuralnym.

Z teoretycznego punktu widzenia, rozważany w tym punkcie problem należy do klasy zagadnień programowania nieliniowego [4], [9]. Dokładne sformułowanie jest następujące:

Dana jest funkcja $f(\cdot): R^n \ni x \to f(x) \in R$, klasy C^2, oraz przedział domknięty $\bar{P}$ w R^n,

$$\bar{P} = \{x_1, x_2, \ldots, x_n) \in R^n : a_1 \leqslant x_i \leqslant b_i\},\ a_i < b_i. \tag{30}$$

Należy:

a. wyznaczyć wszystkie punkty x* ekstremum lokalnego funkcji f(·) w obrębie obszaru $\bar{P}$. (x^* jest punktem minimum funkcji $f(\cdot)$ w obszarze $\bar{P}$, jeżeli dla pewnego otoczenia otwartego U punktu x^* w R^n zachodzi $f(x^*) \leqslant f(x)$, dla wszystkich $x \in P \cap U$. Jeżeli minima są izolowane, to $f(x^*) < f(x)$ dla każdego $x \in U \cap P$ i $x \neq x^*$, *gdzie* U jest pewnym otoczeniem otwartym punktu x^* w R^n. Analogiczna uwaga dotyczy punktów maksimum.

b. znaleźć wartość minimalną i wartość maksymalną funkcji $f(\cdot)$ w obszarze $\bar{P}$.

c. wyznaczyć punkt x_g^* minimum globalnego i punkt $\bar{x}_g^*$ maksymum globalnego funkcji $f(\cdot)$ w obszarze $\bar{P}$ (ewentualnie wszystkie te punkty, gdy jest ich więcej niż po jednym).

Z punktu widzenia systemowego realizacja zadań wymienionych w punktach a.–c. sprowadza się do przedstawienia konstrukcji odpowiedniego systemu dynamicznego, w tym w szczególności systemu zrealizowanego w technice układów elektronicznych, wykonującego automatycznie zadania wymienione w punktach a.–c.

A konkretnie:

1. dokonać syntezy układu dynamicznego (to znaczy systemu opisanego równaniem stanu postaci $\dot{x} = g(x)$, $g(\cdot): R^n \to R^n$ jest funkcją ciągłą) takiego. że dla każdego warunku początkowego $x^0 = x(t = 0)$, $x^0 \in \bar{P}$, dodatnia półtrajektoria $x(x^0, t)$, $t \geqslant 0$, systemu zmierza w granicy dla $t \to \infty$ do jednego z punktów x^* ekstremum lokalnego funkcji $f(\cdot)$ w obrębie obszaru $\bar{P}$. Ponadto, jeżeli ekstrema lokalne funkcji $f(\cdot)$ są izolowane, to każdy z punktów x^* jest asymptotycznie stabilnym punktem równowagi $\dot{x} = g(x)$. Odnośnie konstruowanego systemu $\dot{x} = g(x)$ zakłada się, że funkcja $g(\cdot)$ zapewnia jednoznaczność jego rozwiązań, oraz że dla dowolnie ustalonego warunku początkowego $x^0 \in R^n$, dodatnia półtrajektoria $x(x^0, t)$ jest ograniczona.

2. dokonać syntezy układu dynamicznego takiego, że dla każdego warunku początkowego $x^0 = x(t = 0)$, dla dodatniej półtrajektorii $x(x^0, t)$ tego systemu zachodzi:

granica $\lim_{t\to\infty} f(x(x^0, t))$ istnieje i równa się ona wartości minimalnej (odpowiednio, maksymalnej) funkcji $f(\cdot)$ w obrębie obszaru $\bar{P}$ (wartość graniczna nie zależy od wyboru punktu początkowego x^0).

3. Wykorzystując układy wykonujące zadania jak w punkatach 1. i 2. dokonać syntezy całości systemu wyznaczającego wszystkie punkty x^* ekstremum lokalnego funkcji $f(\cdot)$ w przedziale $\bar{P}$ (przy założeniu, że funkcja posiada ekstrema izolowane) i znajdującego lokalizację punktów $\underline{x}_g^*$ i $\bar{x}_g^*$.

Synteza układu dynamicznego realizującego zadanie 1 jest zagadnieniem dość szeroko i szczegółowo opracowanym, tak w zakresie matematycznym jak też w sensie syntezy odpowiedniego systemu elektronicznego [9], [10]. Dlatego też uwaga zostanie skoncentrowana na zadaniu przedstawionym w punkcie 2., znacznie mniej opracowanym, a mającym istotne znaczenie w realizacji zadania kompleksowego 3. Przedstawiona też zostanie struktura blokowa całości systemu.

Na wstępie przytoczone zostaną twierdzenia matematyczne oraz wynikające z nich wnioski mające zastosowanie w kontekście problemu syntezy przedstawionego w p. 2.

Twierdzenie 1 (Weierstrassa [8])

Dla funkcji ciągłej określonej na przestrzeni zwartej i o wartościach w R, zbiór jej wartości zawiera swoje kresy.

W szczególności teza twierdzenia zachodzi dla każdego podzbioru ograniczonego i domkniętego Z w R^n (np. przedziału domkniętego $\bar{P}$) i funkcji ciągłej $f(\cdot) : Z \to R$. W zbiorze Z istnieją takie punkty $\bar{x}^*$ i $\underline{x}^*$ w których

$$f(\bar{x}^*) = \max_{x\in Z} f(x) \tag{31}$$

i

$$f(\underline{x}^*) = \min_{x\in Z} f(x) \tag{32}$$

(ma to miejsce dla dowolnej przestrzeni zwartej Z).

Wniosek z Twierdzenia Hahna–Mazurkiewicza ([8], str. 251)

Podzbiór zwarty spójny i lokalnie spójny przestrzeni R^n jest obrazem ciągłym odcinka [0, 1]. To znaczy, dla każdego takiego podzbioru Z w R^n istnieje funkcja ciągła $\zeta_Z(\cdot) : [0, 1] \ni t \to \zeta_Z(t) \in Z$ przekształcająca odcinek [0, 1] na Z, Im $\zeta_Z(\cdot) = Z$.

Z wniosku wynika natychmiast, że dla zadanej funkcji ciągłej $f(\cdot)$ określonej na pewnym podzbiorze zwartym Z w R^n, spójnym i lokalnie spójnym $f_{min|Z} = \min_{x\in Z} f(x)$ i $f_{max|Z} = \max_{x\in Z} f(x)$ można wyznaczyć poszukując tych wartości dla funkcji (jednej zmiennej)

$$f \circ \zeta_Z(\cdot)$$

określonej na odcinku [0, 1], gdzie $\zeta_Z(\cdot)$ jest pewną funkcją ciągłą odwzorowującą przedział [0, 1] na Z;

$$f_{max|Z} = \max_{t\in[0,\ 1]} f \circ \zeta_Z(t) \tag{33}$$

i

$$f_{min|Z} = \min_{t\in[0,\ 1]} f \circ \zeta_Z(t). \tag{34}$$

Formuły (33), (34) nie są wykorzystywane w tej pracy, mogą one być jednak przydatne w praktycznym rozwiązywaniu rozważanego tu problemu.

W proponowanej konstrukcji systemu dynamicznego dla wyznaczania wartości $f_{min|Z}$ i $f_{max|Z}$ dla zadanej funkcji ciągłej $f(\cdot)$ na zbiorze zwartym Z wykorzystuje się pojęcie trajektorii wszędzie gęstej [1], [11] oraz możliwość generacji tego rodzaju trajektorii, jak też trajektorii dowolnie dokładnie je aproksymujących, w gładkich układach dynamicznych.

Niech Z będzie podzbiorem zwartym w R^n, dla którego istnieje funkcja $\tilde{x}(\cdot) : [0, \infty) \ni t \to Z$, klasy C^1, wyznaczająca w Z trajektorię $\tilde{x}(t)$ wszędzie gęstą w Z. To znaczy trajektorię przebiegającą w nieskończenie bliskim sąsiedztwie każdego z punktów ze zbioru Z, w pewnej chwili czasu $t \geqslant 0$ [1]. Podzbiorem Z jest w tym opracowaniu przedział domknięty $\bar{P}$.

Niech $f(\cdot)$ będzie funkcją rzeczywistą klasy C^1 na pewnym podzbiorze zwartym Z w R^n, oraz niech $\mu(\cdot) : R \to \{0, 1\}$ będzie funkcją schodkową daną przepisem:

$$\mu(y) = \begin{cases} 1, \text{ dla } y \leqslant 0, \\ 0, \text{ dla } y > 0. \end{cases} \tag{35}$$

Zachodzi następujące twierdzenie.

Twierdzenie 2.

Dla danego podzbioru zwartego Z w R^n, dla którego istnieje pewna trajektoria $\tilde{x}(t)$, $t \in [0, \infty)$, klasy C^1 wszędzie gęsta w Z ($\tilde{f}(t) \triangleq f(\tilde{x}(t))$ oraz $\dot{\tilde{f}}(t) \triangleq (\mathrm{d}/\mathrm{d}t\,\tilde{f}(t))$) wartość graniczna wyrażenia całkowego

$$\Phi(t) = \int_0^t \mu(\tilde{f}(\tau) - \Phi(\tau) \cdot \mu(\dot{\tilde{f}}(\tau)) \cdot \dot{\tilde{f}}(\tau)\mathrm{d}\tau + f(\tilde{x}(0)), \tag{36}$$

dla $t \to \infty$, jest dobrze określona oraz

$$\lim_{t\to\infty} \Phi(t) = f_{min|Z}. \qquad \square$$

Prosty dowód tego twierdzenia jest pominięty.

Zachodzi też twierdzenie dualne. Niech mianowicie $\nu(\cdot) : R \to \{0, 1\}$ będzie funkcją schodkową daną przepisem:

$$\nu(y) = \begin{cases} 0, \text{ dla } y < 0, \\ 1, \text{ dla } y \geqslant 0. \end{cases} \tag{37}$$

Twierdzenie 3.

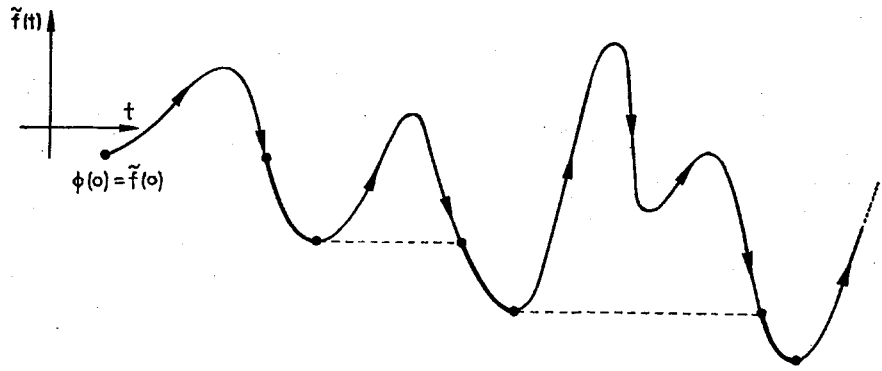
Dla danego podzbioru zwartego Z w R^n, dla którego istnieje pewna trajektoria $\tilde{x}(t)$, $t \in [0, \infty)$, klasy C^1 wszędzie gęsta w Z wartość graniczna całki

$$\Phi(t) = \int_0^t \nu(\tilde{f}(\tau) - \Phi(\tau)) \cdot \nu(\tilde{f}^*(\tau)) \cdot \tilde{f}^*(\tau) d\tau + f(\tilde{x}(0)), \tag{38}$$

dla $t \to \infty$, jest dobrze określona oraz

$$\lim_{t \to \infty} \Phi(t) = f_{max|Z}. \quad \square \tag{39}$$

Na Rys. 18 przedstawiono ilustrację odnośnie formuł podanych we wzorach (36), (38).



Rys. 18. Przebieg $\tilde{f}(t) = f(\tilde{x}(t))$ dla pewnej funkcji $f(\cdot)$ i trajektorii $\tilde{x}(t)$ (wszędzie gęstej w zadanym podzbiorze Z w R^n). Zaznaczono odcinki czasu, na których odpowiednie wyrażenia podcałkowe w (36) i (38) przyjmują wartość różną od zera (na tych odcinkach ma miejsce całkowanie)

Trajektoria wszędzie gęsta w danym obszarze Z nazywana jest dalej trajektorią przeszukującą (dla tego obszaru). Generacja takiej trajektorii dla obszaru w postaci przedziału $\bar{P}$ nie przedstawia z teoretycznego punktu widzenia trudności. Przykładem krzywej gładkiej w $\bar{P}$, której wykres jest podzbiorem wszędzie gęstym w $\bar{P}$ jest linia zakreślona w obszarze $\bar{P}$ przez punkt biorący udział jednocześnie w n ruchach harmonicznych opisanych (we współrzędnych $x_1, x_2, \ldots, x_n$ w R^n) równaniami

$$\tilde{x}_i(t) = A_i \cdot \sin(\omega_i \cdot t + \zeta_i) + B_i, \; i = 1, 2, \ldots, n, \tag{40}$$

$t \in [0, \infty)$, gdzie A_i i B_i są ustalone tak, że

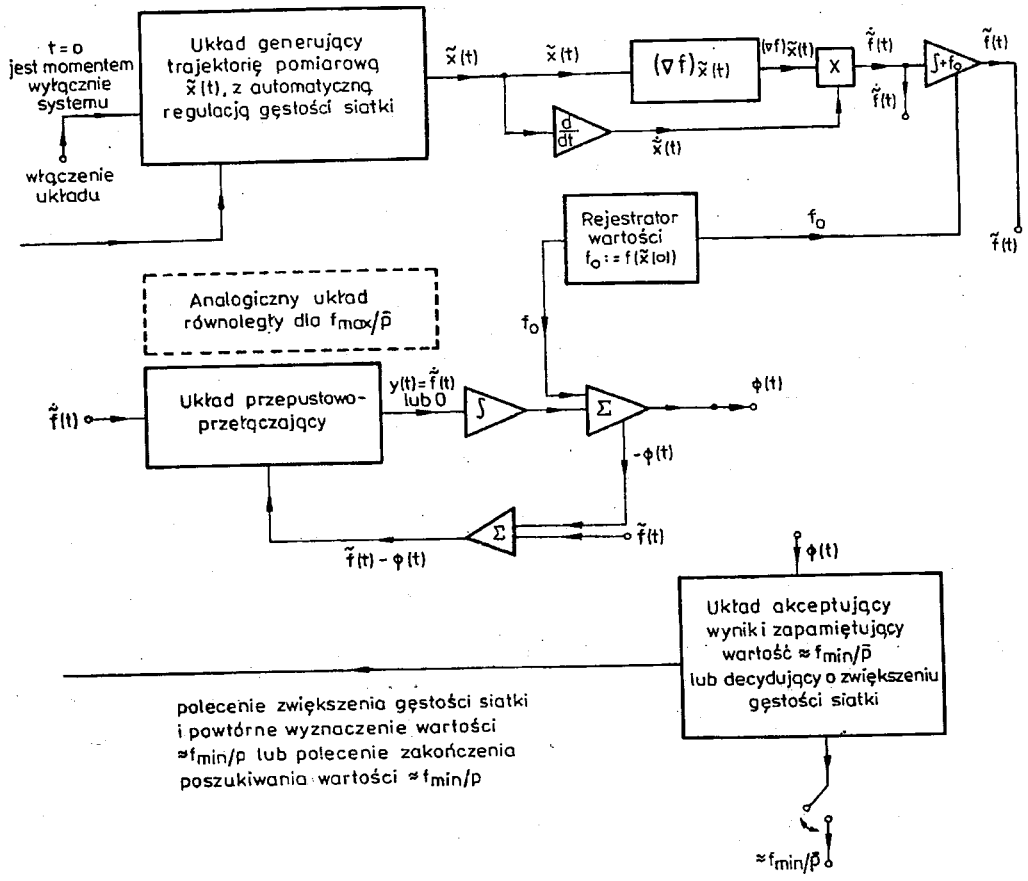
$$a_i = -|A_i| + B_i,$$
$$b_i = |A_i| + B_i,$$

(liczby a_i, b_i występują w opisie (30) przedziału $\bar{P}$, ζ_i jest fazą początkową ruchu $\tilde{x}_i(t)$, natomiast ω_i, $i = 1, 2, \ldots, n$, są pulsacjami dodatnimi tworzącymi układ nierezonansowy [1]; to znaczy, nie istnieje taki układ n liczb całkowitych $k_1, k_2, \ldots, k_n$ nie wszystkich równych 0, dla którego zachodziłoby

$$\sum_{i=1}^{n} k_i \cdot \omega_i = 0.$$

Jeżeli natomiast dla układu (niezerowych) pulsacji $\omega_1, \omega_2, \ldots, \omega_n$ stosunek $\frac{\omega_j}{\omega_k}$ dla każdych dwóch spośród nich jest liczbą wymierną, to ten układ pulsacji określa się jako rezonansowy. Dla układu rezonansowego pulsacji, wykres odpowiedniej trajektorii $\tilde{x}(t)$ opisanej równaniami (40) jest krzywą zamkniętą w R^n, i zawartą w obszarze P, o okresie danym przez najmniejszą wspólną wielokrotność okresów $T_i = 2\pi/\omega_i$ ruchów składowych. W przypadku $n = 2$ wykres trajektorii rezonansowej przedstawia krzywą Lissajou.

Generacja trajektorii przeszukującej opisanej równaniami (40) nie przedstawia z praktycznego punktu widzenia trudności. Zważywszy jednak, że przedział czasu dla wyznaczenia wartości $f_{min|\bar{P}}$ i $f_{max|\bar{P}}$ nie może być nieskończenie długi, nie można stawiać w układzie fizycznym (tutaj, układzie elektronicznym) warunku odnośnie generowanej



Rys. 19. Schemat blokowy układu elektronicznego realizującego zadanie wyznaczania wartości $f_{min|\bar{P}}$ i $f_{max|\bar{P}}$

trajektorii pomiarowej $\tilde{x}(t)$, aby była ona „z natury" wypełniającą wszędzie gęsto obszar $\bar{P}$. Wymaga się natomiast, aby trajektoria pomiarowa była okresowa i aby pokrywała (za czas równy jej okresowi) obszar $\bar{P}$ odpowiednio gęstą i możliwie równomiernie rozłożoną siatką. Wymagania te spełnia trajektoria (40) z rezonansowym układem pulsacji, ustalonych stosownie do wymagań odnośnie dokładności pomiaru.

Istnieją też inne proste konstrukcje trajektorii pomiarowych spełniających z zadaną dokładnością wymagania odnośnie gęstości i równomierności. Na Rys. 19 przedstawiono schemat blokowy układu elektronicznego realizującego zadanie wyznaczania wartości minimalnej $f_{min|\bar{P}}$ i wartości maksymalnej $f_{max|\bar{P}}$ pewnej funkcji $f(\cdot) : R^n \to R$ klasy C^1 na zadanym przedziale $\bar{P}$ w R^n.

System z Rys. 19 stanowi bezpośrednią realizację formuł podanych we wzorach (36) i (38), przy zastąpieniu trajektorii przeszukującej wszędzie gęstej w $\bar{P}$ trajektorią pomiarową o odpowiednio dużej gęstości siatki i zapewniającą możliwie jak najbardziej równomierne pokrycie obszaru $\bar{P}$.

W bloku $(\nabla f)_{\tilde{x}(t)}$ wyznaczana jest wartość gradientu funkcji $f(\cdot)$ w bieżącym punkcie $x = \tilde{x}(t)$ na trajektorii pomiarowej. Jest to jedyny blok w systemie, w którym zawarte są dane odnośnie funkcji $f(\cdot)$ (a dokładniej, zawarty jest w nim przepis określający wartość $(\nabla f)_x$ w dowolnym punkcie $x \in \bar{P}$). Na podstawie wyznaczonej wartości wektora $(\nabla f)_x$ w bieżącym punkcie $\tilde{x}(t)$ trajektorii pomiarowej oraz wartości pochodnej (prędkości) $\dot{\tilde{x}}(t) \triangleq \frac{d}{dt}\tilde{x}(t)$ wyznaczana jest wartość pochodnej funkcji $f(\cdot)$ wzdłuż trajektorii $\tilde{x}(t)$, w bieżącej chwili czasu t,

$$\dot{\tilde{f}}(t) = \frac{d}{dt} f(\tilde{x}(t)) = (\nabla f)_{\tilde{x}(t)} \cdot \dot{\tilde{x}}(t). \tag{41}$$

Układ całkujący wyznacza wartość funkcji $f(\cdot)$ w bieżącym punkcie $x = \tilde{x}(t)$ wzdłuż trajektorii $\tilde{x}(t)$,

$$f(\tilde{x}(t)) = \tilde{f}(t) = \int_0^t \dot{\tilde{f}}(\tau)d\tau + f(\tilde{x}(0)). \tag{42}$$

Wartość $\tilde{f}(t)$ jest potrzebna do sterowania w bloku układu przepustowo–przełączającego.

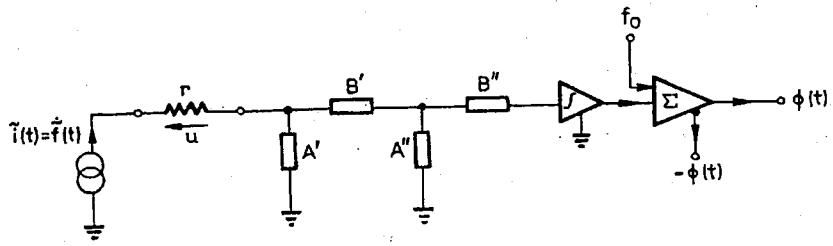
O jakości funkcjonowania całego systemu decyduje sprawność funkcjonowania bloku przepustowo–przełączającego. Na wejście tego bloku podawane są dwa sygnały: sygnał różnicowy $\tilde{f}(t) - \Phi(t)$ oraz sygnał $\dot{\tilde{f}}(t)$. Pierwszy z nich pełni rolę sygnału sterującego, natomiast drugi jest zarówno sygnałem sterującym jak i przetwarzanym. Transmitancja układu przepustowo–przełączającego dana jest wzorem:

$$y(t) = v(\tilde{f}(t) - \Phi(t)) \cdot v(\dot{\tilde{f}}(t)) \cdot \dot{\tilde{f}}(t), \tag{43}$$

w odniesieniu do części systemu wyznaczania $f_{max|\bar{P}}$, oraz

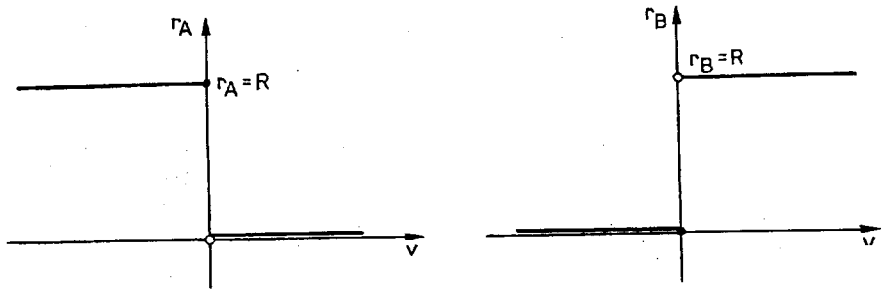
$$y(t) = \mu(\tilde{f}(t) - \Phi(t)) \cdot \mu(\dot{\tilde{f}}(t)) \cdot \dot{\tilde{f}}(t), \tag{44}$$

w odniesieniu do części systemu wyznaczania $f_{min|\bar{P}}$.



Rys. 20. Schemat konstrukcji bloku przepustowo-przełączającego

Zakłada się dostępność w układzie przebiegu prądowego $\tilde{i}(t)$ o wartości $\tilde{f}(t)$ podawanego ze źródła prądowego załączonego na wyjściu układu mnożącego, oraz zakłada się, że przebiegi $\tilde{f}(t)$ i $\Phi(t)$ są osiągalne jako przebiegi napięciowe podawane ze źródeł napięciowych. Na Rys. 20 podano schemat układu elektronicznego realizującego transmitancję (44) (układ realizujący transmitancję (43) otrzymuje się w postaci analogicznej). A i B są nieliniowymi rezystorami sterowanymi napięciowo o charakterystykach jak na Rys. 21.



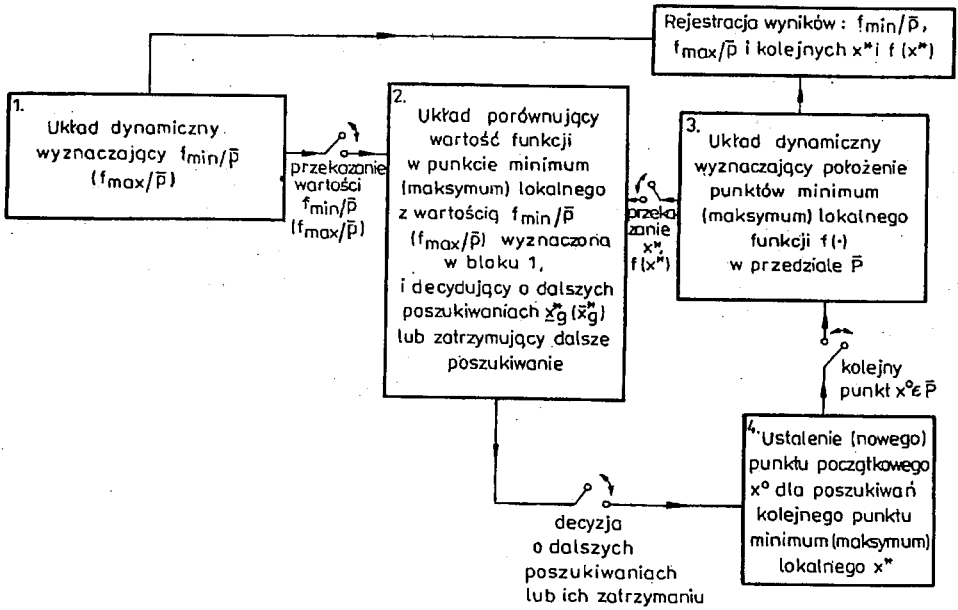
Rys. 21. Charakterystyki rezystorów A i B. A' i B' oraz A'' i B'' są parami identyczne. Dla A' i B' napięciem sterującym jest spadek napięcia u na rezystorze r, $u(t)=r\cdot\tilde{i}(t)$, gdzie $\tilde{i}(t)=\tilde{f}(t)$, natomiast napięciem sterującym dla rezystorów A'' i B'' jest napięcie różnicowe $\tilde{f}(t)-\Phi(t)$. Wartość R rezystancji jest na tyle duża, że można uznać stan $r_A=R$ lub $r_B=R$ za stan rozwarcia elementu r_A lub r_B, odpowiednio

Wartość rezystancji r powinna być dobrana tak, aby wartościom prądu $\tilde{i}(t) = \tilde{f}(t)$ w otoczeniu zera, gdzie zmieniające znak napięcie $u(t) = r \cdot \tilde{i}(t)$ ma powodować przełączanie rezystorów A', B', odpowiadała dostatecznie duża wartość $u(t)$.

Z tych samych względów również należy odpowiednio wzmocnić wartość napięcia różnicowego $\tilde{f}(t) - \Phi(t)$ sterującego rezystorami przełączanymi A'' i B'', w szczególności w otoczeniu zerowej wartości tego napięcia. Czas kluczowanego całkowania przebiegu $\tilde{f}(t)$, dla zadanej trajektorii pomiarowej $\tilde{x}(t)$ o okresie T jest równy T. Wyznaczane są przybliżone wartości $f_{min|\tilde{P}}$ i $f_{max|\tilde{P}}$, a następnie wyznaczane są te wartości dla trajektorii pomiarowej o dwukrotnie gęstszej siatce pokrycia obszaru $\tilde{P}$. W bloku decydującym o akceptacji wyników pomiaru porównywane są otrzymane wartości $f_{min_1|\tilde{P}}$ i $f_{min_2|\tilde{P}}$ oraz $f_{max_1|\tilde{P}}$ i $f_{max_2|\tilde{P}}$. Jeżeli różnice są odpowiednio małe, to

akceptowany jest wynik $f_{min_2|\bar{P}}$, $f_{max_2|\bar{P}}$. W przeciwnym razie gęstość siatki ulega zwiększeniu i pomiar jest powtarzany.

Po zaakceptowaniu wyniku jest on zapamiętywany i przekazywany do bloku 2 całego systemu (Rys. 22). Następuje inicjacja działania bloków 2, 3 i 4 mających za zadanie wyznaczenie punktów ekstremów lokalnych funkcji $f(\cdot)$ w obszarze $\bar{P}$, oraz mających za zadanie wyznaczenie punktów x_g^* i $\bar{x}_g^*$ minimum globalnego oraz maksymum globalnego.



Rys. 22. Schemat blokowy systemu dla wyznaczania położenia punktów i wyznaczania wartości ekstremalnych funkcji

Powracając do formuł ogólnych (36) i (38) odnotować należy, że błędne (to znaczy niezgodne z formułami (36), (38)) całkowanie może mieć miejsce w systemie fizycznym w przypadku, gdy w wyniku zakłóceń oraz naturalnych fluktuacji wartość sygnału sterującego $\tilde{f}(t) - \Phi(t)$ w tych przedziałach czasu, gdzie (dla zadanej trajektorii pomiarowej) $\tilde{f}(t) - \Phi(t) \cong 0$ jest interpretowana przez układ przełączający

$$\mu(\tilde{f}(t) - \Phi(t))$$

lub

$$\nu(\tilde{f}(t) - \Phi(t))$$

(transmitancje $\mu(\cdot)$ i $\nu(\cdot)$ dane są wzorami (35), (37)) jako wartość o znaku przeciwnym. To znaczy, ma miejsce całkowanie, pomimo że klucz ($\mu(\cdot)$ lub $\nu(\cdot)$) powinien być rozwarty lub odwrotnie. W tym kontekście można udoskonalić funkcjonowanie kluczy. Istotna jest też właściwość pewnej samokorekcji błędu przez system, co można zaobserwować rozważając przykładowy przebieg $\tilde{f}(t)$ na wykresie na Rys. 18, przy założeniu nieidealnego funkcjonowania kluczy $\mu(\cdot)$ i $\nu(\cdot)$.

ZAKOŃCZENIE

W pracy rozważono zagadnienie syntezy neuralnych sieci komórkowych. Zaproponowano w tym zakresie szereg oryginalnych rozwiązań. Ogólnie, strukturę systemu tworzy układ funktorów elementarnych połączonych z zewnętrzną siecią funkcyjną realizującą sprzężenia między funktorami, węzłami wejść i wyjść. Z punktu widzenia funkcjonalnego funktory są układami dynamicznymi selekcji – identyfikacji. Funktorem elementarnym sieci neuralnej Hopfielda jest 2-wymiarowy układ dynamiczny flip–flop.

Zaproponowany w pracy funktor elementarny jest układem dynamicznym 1-wymiarowym – w jego równaniu dynamicznym występuje jedna zmienna stanu. Taki wybór elementu podstawowego upraszcza w znacznym stopniu procedurę projektowania sieci, czyniąc ją bardziej uniwersalną i przejrzystą. Łatwiejsze jest też sterowanie rozmiarami obszarów selekcyjnych, prostszy sposób syntezy funktorów wielowymiarowych etc. Odnośnie przedstawionej realizacji elektronicznej dla funktora 1-wymiarowgo należy stwierdzić, że istnieje tu szereg innych rozwiązań, które można wykorzystać. Jest to bowiem typowy element dynamiczny bistabilny.

Pracę w zasadniczej części poświęcono syntezie systemów podejmowania decyzji (systemy selekcji – identyfikacji - decyzji). Systemy decyzyjne rozważono w sensie systemów mnogościowo–logicznych. W każdym z rozważanych w pracy zagadnień syntezy zadana jest rodzina obszarów decyzyjnych jak też funktor logiczny wnioskowania o decyzji. Tego typu ogólnie konstruowane systemy mogą być zastosowane w rozwiązywaniu różnych zadań szczegółowych, np. w syntezie systemów automatyki, w automatycznym rozpoznawaniu obrazów przetwarzaniu informacyjnym sygnałów etc. Istotnym elementem przedstawionych rozważań jest możliwość syntezy systemu przy dowolnie zadanych obszarach selekcyjno – decyzyjnych. W tym kontekście podano szereg alternatywnych rozwiązań.

Rozważane w pracy sieci należą do klasy dynamicznych systemów komórkowych, realizując w stanie ustalonym, dla $t \to \infty$, odwzorowanie continuum stanów początkowych w pewien dyskretny zbiór asymptotycznie stabilnych stanów równowagi, interpretowanych jako stany wyjściowe. Dla przedstawionego w rozdziale 3, cz. II pracy, ogólnego modelu neuralnej sieci komórkowej, z funktorami 1-wymiarowymi i dowolną nieliniową siecią funkcyjną sprzężeń, podano odpowiednie warunki zapewniające jej komórkowy charakter. Zaproponowana tu koncepcja neuralnej sieci komórkowej wydaje się być bardziej uniwersalna i prostsza w zastosowaniach niż koncepcja podana w pracach [5], [6].

W rozdziale 4, cz. II pracy, przedstawiono oryginalną koncepcję systemu automatycznego wyznaczania punktów i wartości ekstremalnych funkcji (systemu optymalizacyjnego). Podstawowy blok tego systemu ma cechy charakterystyczne sieci neuralnej.

Nie przedstawiono zastosowań proponowanych sieci w syntezie systemów analizy i przetwarzania obrazów, systemów uczących się i innych. Konstrukcja tego typu systemów jest ideowo taka sama jak w przypadku zastosowania sieci Hopfielda. Innym spośród zagadnień, które pozostają poza zakresem pracy są relacje wiążące

sieci neuralne pracujące w czasie dyskretyzowanym (systemy iterowane) i sieci pracujące w czasie ciągłym.

BIBLIOGRAFIA

1. W. I. Arnold: *Teoria równań różniczkowych.* Warszawa: PWN, 1983
2. L. O. Chua, D. N. Green: *A qualitative analysis of the behavior of dynamic nonlinear networks: stability of autonomous networks.* IEEE Trans. Circuits and Systems, 1976, vol. CAS–23, pp. 355–379.
3. L. O. Chua, D. N. Green: *A qualitative analysis of the behaviour of dynamic nonlinear networks: steady–state solutions of nonautonomous networks.* IEEE Trans. Circuits and Systems, 1976, vol. CAS–23, pp. 530–550
4. L. O. Chua, G. N. Lin: *Nonlinear optimization with constraints: a cookbook approach.* Int. J. Circuit Theory Appl., 1983, vol. 11, pp. 141–159
5. L. O. Chua, L. Yang: *Cellular neural networks: theory.* IEEE Trans. Circuits and Systems, 1988, vol. CAS–35, no 10, pp. 1257–1271
6. L. O. Chua, L. Yang: *Cellular neural networks: applications.* IEEE Trans. Circuits and Systems, 1988, vol. CAS–35, no 10, pp. 1273–1290
7. B. P. Demidowicz: *Matematyczna teoria stabilności.* Warszawa: WNT, 1972
8. R. Duda: *Wprowadzenie do topologii, część I. Topologia ogólna., część II. Topologia algebraiczna i topologia rozmaitości.* Warszawa: PWN, 1986
9. R. J. Duffin, E. L. Peterson, C. Zener: *Geometric Programming – Theory and Applications.* J. Wiley, New York, 1967
10. M. P. Kennedy, L. O. Chua: *Neural networks for nonlinear programming.* IEEE Trans. Circuits and Syst., 1988, vol CAS–35, no 5, pp. 554–562
11. V. V. Nemytskii, V. V. Stepanov: *Qualitative Theory of Differential Equations.* Princeton University Press, N. J., 1972
12. S. Osowski: *Projektowanie obwodów poprzez analizę rozszerzonych sieci neuronowych.* Mat. XII KKTOiUE, Rzeszów–Myczkowce, 17–20 X 1989, t.1, ss. 157–162
13. K. Preston, Jr., M. J. B. Duff: *Modern Cellular Automata: Theory and Applications.* Plenum, New York, 1984
14. A. Szatkowski: *Contribution to the qualitative theory of non–linear systems and networks.* Int. J. Cir. Theor. Appl., 1980, vol. 8, pp. 373–393
15. S. Wolfram (Eds.): *Theory and Applications of Cellular Automata.* World Scientific, New York, 1986

A. SZATKOWSKI

NEURAL CELLULAR SYSTEMS SYNTHESIS FOR DECISION MAKING PART II: SET – LOGIC SYSTEMS AND GENERAL CELLULAR NEURAL SYSTEMS

Summary

In the Part I a universal structure of cellular neural system for selection – identification tasks has been proposed, where 1-dimensional selectors are basic functors. The cellular set – logic systems in the Part II are the extension of the systems considered in the Part I. They are the system realization of set – logic functors. Both mathematical model as well as electrical circuit model of general cellular system with 1-dimensional selectors attached to non–linear coupling system (network) have been proposed.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1992, 38, z. 2–3, ss. 513–514

Knowledge base paradigm early stage of digital image processing

TOMASZ OSTROWSKI

Instytut Elektroniki i Informatyki, Politechnika Szczecińska

Received 1992.04.20

Authorized 1992.05.20

REPORT*

Early stages of digital image processing require suitable application of the preprocessing operators. This paper proposes knowledge base framework aimed to support not skilled users with the design and run-time application of image preprocessing procedures.

Subject terms: convolution, knowledge based image processing median, morphology, object oriented programing.

INTRODUCTION

Usually, the initial step in digital image analysis is so-called preprocessing, i.e., sequence of input image transformations performed to prepare the image for specific analysis. Those transformations often involve: gray level rescaling, shading correction, linear and nonlinear filtrations, histogram linearization, etc..

The proper design of image processing operators and tuning their parameters to perform tasks mentioned above is strongly dependent on particular image processing problem under consideration as well as experiment's conditions, e.g., lighting and/or specimen's preparation, if one has to cope with microscope image analysis problem.

Those premises lead to the necessity to work on inteligent software, dedicated to support the creation of image processing procedures with the aid of some common image processing primitives such as: convolution, median, and morphological operators.

2. HARDWARE DESCRIPTION

The hardware requirements for digital image processing system are: RISC technology computer equipped with 512 x 512 pixel frame grabber connected to CCD camera. There is also need for specialized VLSI architecture to perform real time convolution, median and morphology.

* This work, lead in Technical University of Szczecin, is financed by the Committee of Scientific Research, nr 108/E–330

3. KNOWLEDGE BASE AND REASONING

Domain knowledge is under elicitation from the team of domain experts, e.g., material science engineers for metallography.

As the knowledge representation framework, rule based object oriented system encoded in C^{++} seems to be most suitable.

Reasoning mechanism used to work out the processing strategy are: forward and backward chaining engines with fuzzy handling of certainty factors.

There is also promising to use neural network computing to recognize domain image under consideration belongs to (metallography, tissue analysis, etc.), for the switching between multiple knowledge bases.

4. IMAGE PROCESSING PRIMITIVES

As the image processing primitives we mean convolution filters and based on them operators (Laplace, Sobel, Laplacian of Gaussian [1], etc.), binary and gray level morphology [2], median and its combinations with convolution, e.g., linear-median hybrid edge detectors (LMHED) [3]. Results of application concerning few mentioned above primitives (Laplacian of Gaussian and morphology) can be found in [4].

REFERENCES

1. W. E. L. Grimson and E. C. Hildreth: Comments on "Digital step edges from zero crossings of second directional derivatives", IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI – 7, no. 1, pp. 121-127, Jan. 1985
2. R. M. Haralick, S. R. Sternberg and X. Zhuang: Image analysis using mathematical morphology, IEEE Trans. Pattern Anal. Machine Intell.,vol. PAMI-9, no. 7, pp. 532-550, July 1987
3. Y. Neuvo, P. Heinonen and I. Defee: Linear-median hybrid edge detectors, IEEE Trans. circuits Syst., vol. CAS-34, pp. 1337-1343, Nov. 1987
4. T. Ostrowski, P. Bajkowski, K. Wicyński and R. Musiał: Feature detection in digital pictures using Image Processing System VIPS (in Polish), Kwartalnik Elektroniki i Telekomunikacji nr 3-4/1991, t.37, ss.549-570

T. OSTROWSKI

Streszczenie

Wstępne etapy cyfrowego przetwarzania obrazów wymagają właściwego doboru operatorów przetwarzania. Proponowana jest organizacja bazy wiedzy wspierającej użytkownika w doborze procedur cyfrowego przetwarzania obrazów.

SPIS TREŚCI

CONTENTS

Cena zł 60 000,–

Kwartalnik Elektroniki i Telekomunikacji

Wpłaty na prenumeratę przyjmowane są na okresy kwartalne:

— na terenie kraju — jednostki kolportażowe „Ruch" i urzędy pocztowe właściwe dla miejsca zamieszkania lub siedziby prenumeratora,

— na zagranicę — „RUCH" S.A. Oddział Warszawa 00-958 Warszawa, konto PBK XIII Oddział Warszawa 370044-1195-139-11.

Prenumerata ze zleceniem dostawy za granicę jest o 100 % wyższa od krajowej.

Dostawa zamówionej prasy następuje:

— przez jednostki kolportażowe „Ruch" — w sposób uzgodniony z zamawiającym,

— przez urzędy pocztowe — pocztą zwykłą na wskazany adres, w ramach opłaconej prenumeraty z wyjątkiem zlecenia dostawy za granicę pocztą lotniczą do odbiorcy zagranicznego, której koszt w pełni pokrywa prenumerator.

Terminy przyjmowania wpłat na prenumeratę:

— krajową i zagraniczną — do 20.XI. na I kw. roku następnego
do 20.II. na II kw.
do 20.V. na III kw.
do 20.VIII. na IV kw.

Bieżące numery można nabyć w Księgarni Wydawnictwa Naukowego PWN Sp. z o.o. ul. Miodowa 10. Również można je nabyć, a także zamówić (przesyłka za zaliczeniem pocztowym) we Wzorcowni Ośrodka Rozpowszechniania Wydawnictw Naukowych PAN, Pałac Kultury i Nauki, 00-901 Warszawa.

Subscription orders for all the magazines published in Poland available through the local press distributors or directly through the

Foreign Trade Enterprise

ARS POLONA

00-068 Warszawa, Krakowskie Przedmieście 7, Poland

our bankers:

BANK HANDLOWY S.A. 201061-710-13100

POLSKA AKADEMIA NAUK
KOMITET ELEKTRONIKI I TELEKOMUNIKACJI

PL ISSN 0035-9386
Nr Indeksu 363189

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY

dawniej

ROZPRAWY ELEKTROTECHNICZNE

TOM XXXVIII — ZESZYT 4/1992

WYDAWNICTWO NAUKOWE PWN
WARSZAWA 1993

POLSKA AKADEMIA NAUK
KOMITET ELEKTRONIKI I TELEKOMUNIKACJI

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY

TOM XXXVIII — ZESZYT 4/1992

WYDAWNICTWO NAUKOWE PWN
WARSZAWA 1993

ADRES REDAKCJI

00-665 Warszawa, ul. Nowowiejska 15/19 Politechnika, pok. 470
Instytut Telekomunikacji, Gmach im. prof. JANUSZA GROSZKOWSKIEGO

Dyżury Redakcji: środy i piątki, godz. 14–16
tel. 628 89 81 21007-737

Telefony domowe: Redaktora Naczelnego: 12 17 65
Zast. Red. Naczelnego: 26 83 41
Sekretarza Odpowiedzialnego: 25 29 18

WYDAWNICTWO NAUKOWE PWN
Warszawa, ul. Miodowa 10

Ark. wyd. 16.25 Ark. druk. 13.5	Podpisano do druku w grudniu 1992 r.
Papier offsetowy kl. III 70 g. B-1	Druk ukończono w styczniu 1993 r.

Skład: Wyd. **GP-BIT** W-wa ul. Marymoncka 34. Druk i oprawa: Zakład Poligr., W-wa ul. Położowa 8

Szanowni Czytelnicy

Zwracamy się z propozycją prenumeraty naszego wydawnictwa pt.

KWARTALNIK
ELEKTRONIKI I TELEKOMUNIKACJI
ELECTRONICS AND TELECOMMUNICATIONS
QUARTERLY

Czasopismo to jest bezpośrednią kontynuacją wydawanego od 38 lat pod auspicjami Polskiej Akademii Nauk kwartalnika pt. „Rozprawy Elektrotechniczne"
KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI jest czasopismem Komitetu Elektroniki i Telekomunikacji Polskiej Akademii Nauk i w związku z tym na jego łamach publikowane są prace naukowe dotyczące podstaw teoretycznych i zastosowań z zakresu elektroniki, telekomunikacji, mikroelektroniki, optoelektroniki, radiotechniki i elektroniki medycznej. Zamieszczane są prace oryginalne, w tym przyczynkowe oraz zawierające analizy porównawcze metod lub systemów, syntetyczne ujęcia określonej dziedziny wiedzy, omówienia aktualnego stanu lub postępu danej gałęzi techniki oraz omówienia perspektyw rozwojowych.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI przeznaczony jest dla specjalistów oraz studentów zajmujących się wymienioną tematyką. W Kwartalniku są zamieszczane artykuły w języku polskim lub angielskim, wybranym przez Autora.

Warunki prenumeraty oraz informacje dla Autorów podane są w zakończeniu bieżącego numeru Kwartalnika.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1992, 38, z. 4, ss. 523-536

621.317.341
621.3.082

Zastosowanie metody różnic skończonych do analizy niesymetrycznych linii mikropaskowych

PAWEŁ KLONECKI, ZYGMUNT LANGOWSKI

Instytut Telekomunikacji i Akustyki, Politechnika Wrocławska

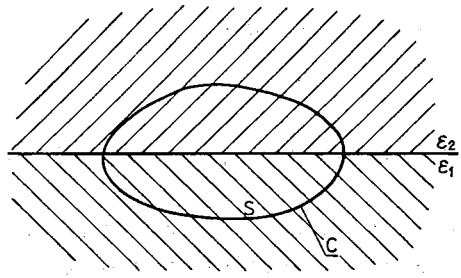
Otrzymano 1991.11.08

Autoryzowano do druku 1992.06.10

Artykuł prezentuje metodę różnic skończonych, która może być użyta do rozwiązania wielu nowoczesnych problemów w praktyce mikrofalowej, sformułowanych w postaci równań różniczkowych (także cząstkowych) lub całkowych ze skomplikowanymi warunkami brzegowymi.
Znaleziono równania różnic skończonych dla dowolnego położenia węzłów siatki, w ośrodku jednorodnym lub warstwowym. Prezentowany jest sposób obliczania błędu siatki i błędu obliczeniowego.
Metoda została zastosowana w analizie quasistatycznej niesymetrycznych linii mikropaskowych, z jedną warstwą dielektryczną, pojedynczą lub sprzężoną, z dodatkowymi paskami o potencjale równym zero.

1. WSTĘP

Wiele problemów inżynierskich w elektrotechnice może być sformułowanych w postaci cząstkowych równań różniczkowych (np. Laplace'a, Poissona, Gaussa). Rozwiązanie ich sprowadza się zazwyczaj do znalezienia pewnej wielkości skalarnej (najczęściej potencjału) spełniającej dane równanie przy określonych warunkach brzegowych. W praktyce często spotyka się ośrodki warstwowe, jak na rys. 1. Jeżeli



Rys. 1. Ośrodek warstwowy z zaznaczonym obszarem całkowania

poszczególne warstwy ośrodka są jednorodne, to wektor natężenia pola elektrycznego $\bar{E}$ oraz potencjał pola Φ są powiązane znanymi czterema równaniami:

$$\bar{E} = -\nabla \Phi \quad \text{– definicja potencjału} \tag{1}$$

$$\nabla^2 \Phi = -\rho/\varepsilon \quad \text{– prawo Poissona} \tag{2}$$

$$\nabla^2 \Phi = 0, \quad \text{dla} \quad \rho = 0 \quad \text{– prawo Laplace'a} \tag{3}$$

$$\oint_S \varepsilon \bar{E}\, d\bar{S} = -\oint_C \varepsilon \frac{\partial \Phi}{\partial \bar{n}}\, dC = \Sigma q \quad \text{– prawo Gaussa} \tag{4}$$

gdzie:

∇ – operator różniczkowy Hamiltona.
ρ – gęstość ładunku swobodnego w zamkniętym obszarze S
$\varepsilon = \varepsilon_r \varepsilon_0$ – przenikalność elektryczna ośrodka
ε_r – względna przenikalność dialektryka
ε_0 – przenikalność elektryczna swobodnej przestrzeni
Σq – suma ładunków swobodnych wewnątrz obszaru zamkniętego S, ograniczonego konturem C
$\bar{n}$ – wektor jednostkowy, normalny na konturze C

W zależności od rodzaju warunków brzegowych rozważa się trzy zagadnienia brzegowe:

$$\Phi(p) = f_1(p) \quad \text{– zagadnienia Dirichleta} \tag{5}$$

$$\left(\frac{\partial \Phi}{\partial n}\right)_p = f_2(p) \quad \text{– zagdnienia Neumanna} \tag{6}$$

$$\left(\frac{\partial \Phi}{\partial n}\right)_p = f_3(p)\,[\Phi(p) - f_4(p)] \quad \text{– zagadnienie Hankela} \tag{7}$$

gdzie:

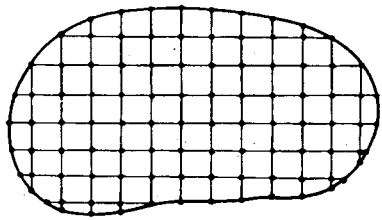
$f_1(p) - f_4(p)$ – dowolne, z góry zadane funkcje.

Dla rozpatrywanych w pracy obszarów wykorzystano dwa pierwsze zagadnienia. Do rozwiązywania przedstawionych równań (1–4) zastosowano jedną z metod różnicowych – metodę różnic skończonych.

2. ISTOTA METODY RÓŻNIC SKOŃCZONYCH

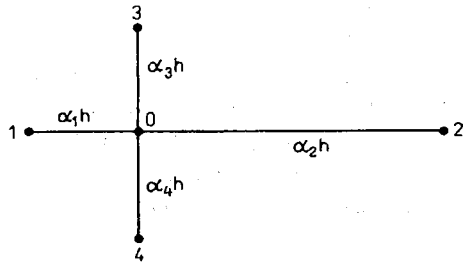
Metoda różnic skończonych polega na zastąpieniu pochodnych cząstkowych w równaniach różniczkowych przez odpowiednie ilorazy różnicowe [1,2,3]. Powstają w ten sposób tak zwane równania różnicowe, które wiążą ze sobą wartości

poszukiwanej funkcji w pojedynczych, odosobnionych punktach. Punkty te wybiera się na ogół tak, aby utworzyły węzły siatki regularnej, pokrywającej cały badany obszar (rys. 2), który może być skończony lub nieskończoy. Dla obszarów skończonych znane są warunki brzegowe, natomiast obszary nieskończone przybliża się obszarem skończonym lub stosuje się specjalne metody symulujące dopasowujące brzeg otwarty [4]. Równania różnicowe zapisane dla każdego węzła siatki są równaniami liniowymi. W ten sposób rozwiązanie zagadnienia, opisanego skomplikowanymi równaniami różniczkowymi, zostaje sprowadzone do rozwiązania układu równań algebraicznych.



Rys. 2. Przykładowy obszar pokryty siatką różnicową

Rodzaje stosowanej siatki zależą od geometrii badanego obszaru jak i przyjętego układu współrzędnych. Przeważnie stosuje się siatki trójkątne, kwadratowe, prostokątne, sześciokątne. W pracy [5] opisano siatki „samodostosowujące się" do geometrii układu, co znacznie redukuje liczbę powstających węzłów.
Rozważania ograniczono tylko do dwuwymiarowego i prostokątnego układu współrzędnych. Wobec tego zastosowano siatkę kwadratową lub prostokątną. Położenie sąsiednich punktów siatki dwuwymiarowej przedstawiono na rys. 3.



Rys. 3. Układ węzłów dla siatki o nierównych ramionach

Niech poszukiwana funkcja potencjału $\Phi(x,y)$ jest funkcją holomorficzną w rozpatrywanym obszarze. Rozwijając funkcję $\Phi(x,y)$ w szereg Taylora w punktach o współrzędnych (x_0,y_u) i (x_u,y_0) można różnicowe przybliżenie operatora Laplace'a w punkcie „0" (rys. 3) dla siatki o nierównych ramionach zapisać w następującej postaci [1,4]:

$$\nabla^2\Phi|_0 \approx \frac{2}{h^2}\left[\left(\frac{\Phi_1}{a_1}+\frac{\Phi_2}{a_2}\right)\frac{1}{a_1+a_2}+\left(\frac{\Phi_3}{a_3}+\frac{\Phi_4}{a_4}\right)\frac{1}{a_3+a_4}+ - \Phi_0\left(\frac{1}{a_1a_2}+\frac{1}{a_3a_4}\right)\right] \tag{8}$$

gdzie:

$$\nabla^2\Phi|_0 = \left(\frac{\partial^2\Phi}{\partial x^2}\right)_0 + \left(\frac{\partial^2\Phi}{\partial x^2}\right)_0$$

$a_j h$ dla $j = 1,2,3,4$ — odległości między węzłami siatki
$0 < a_j \leqslant 1$
$h > 0$ — rozmiar postawowej siatki kwadratowej
Φ_j — wartość funkcji $\Phi(x,y)$ w punkcie j.

Dla siatki kwadratowej ($a_j = 1$) równanie (8) redukuje się do następującej postaci:

$$\nabla^2\Phi|_0 \approx \frac{\Phi_1+\Phi_2+\Phi_3+\Phi_4-4\Phi_0}{h^2}. \tag{9}$$

Wykorzystując zależność (8) równanie Laplace'a dla punktu węzła siatki o numerze 0 można zapisać w następującej postaci:

$$\Phi_0 \approx \frac{a_1a_2a_3a_4}{a_1a_2+a_3a_4}\left[\left(\frac{\Phi_1}{a_1}+\frac{\Phi_2}{a_2}\right)\frac{1}{a_1+a_2}+\left(\frac{\Phi_3}{a_3}+\frac{\Phi_4}{a_4}\right)\frac{1}{a_3+a_4}\right]. \tag{10}$$

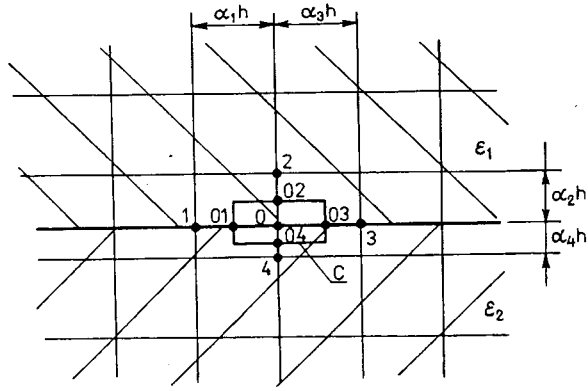
Zależność (10) dla warunku $a_j = 1$ (siatka o równych ramionach) redukuje się do wzoru:

$$\Phi_0 \approx \frac{\Phi_1+\Phi_2+\Phi_3+\Phi_4}{4}. \tag{11}$$

Równania (8) i (9) wyprowadzono stosując centralną metodę różnic. Metodę tą wybrano ze względu na jej zalety w stosunku do dwóch innych również stosowanych metod — metody „w przód" i metody „w tył" [1,4]. Należy zaznaczyć, że podane w pracy [1] wzory różnicowe są błędne i na ich podstawie nie można opracować poprawnego algorytmu do obliczeń numerycznych.

3. RÓWNANIA RÓŻNICOWE NA GRANICY DWÓCH OŚRODKÓW

Dla układu węzłów, jak na rys. 4, konieczne jest wyprowadzenie wzorów umożliwiających obliczenie wartości funkcji Φ na granicy obszarów (granica na rys. 4 oznaczona jest grubszą linią). Niech rozpatrywaną funkcją Φ będzie potencjał elektrostatyczny, a parametrem zmieniającym się skokowo na granicy przenikalność elektryczna, równa ε_1 i ε_2 odpowiednio nad i pod granicą.



Rys. 4. Układ węzłów przy granicy

Niech dla ogólności rozważań odległości sąsiadów węzła o numerze „0" od tegoż węzła „0" będą różne i równe odpowiednio $a_1 h$, $a_2 h$, $a_4 h$. Poprowadzono kontur C przechodzący przez cztery dodatkowe punkty (punkty te oznaczone są na rysunku 4 symbolami 01, 02, 03, 04) znajdujące się w połowie odległości między węzłem „0", a węzłami 1, 2, 3, 4. Zamieniając całkę liniową po konturze C z równania (4) na odpowiadającą jej sumę różnicową, a różniczkę potencjału na odpowiadające jej równanie różnicowe i zakładając brak ładunków swobodnych ($\Sigma q = 0$), można zapisać prawo Gaussa w następującej postaci [1,4]:

$$\varepsilon_1 \frac{\Phi_1 - \Phi_0}{a_1 h} \frac{a_2 h}{2} + \varepsilon_1 \frac{\Phi_2 - \Phi_0}{a_2 h} \frac{a_1 + a_3}{2} h + \varepsilon_1 \frac{\Phi_3 - \Phi_0}{a_3 h} \frac{a_2 h}{2} +$$

$$+ \varepsilon_2 \frac{\Phi_3 - \Phi_0}{a_3 h} \frac{a_4 h}{2} + \varepsilon_2 \frac{\Phi_4 - \Phi_0}{a_4 h} \frac{a_1 + a_3}{2} h + \varepsilon_2 \frac{\Phi_1 - \Phi_0}{a_1 h} \frac{a_4 h}{2} = 0. \qquad (12)$$

Jeśli odległości sąsiednich węzłów od węzła „0" są jednakowe, czyli gdy $a_1 = a_2 = a_3 = a_4 = 1$, co równoznaczne jest z tym, iż punkty leżące na granicy pokrywają się z węzłami siatki, wzór ten upraszcza się do następującej postaci:

$$\Phi_0 = \frac{\Phi_1 + \Phi_3}{4} + \frac{\Phi_2 \varepsilon_1 + \Phi_4 \varepsilon_2}{2(\varepsilon_1 + \varepsilon_2)}. \qquad (13)$$

Wzory (12) i (13) pozwalają obliczyć przybliżoną wartość funkcji Φ w punktach na granicy między dwoma ośrodkami, co umożliwia obliczenie laplasjanu w węzłach siatki leżących bezpośrednio nad i pod linią graniczną.

Wyznaczenie rozkładu potencjału przy określonych warunkach brzegowych sprowadza się do wykonania następującego algorytmu [4]:

– pokrycie badanego obszaru siatką różnicową;
– wyznaczenie zależności obowiązującej dla każdego węzła siatki (na podstawie wzorów 10, 11, 12 lub 13);

— rozwiązanie powstałego układu równań liniowych, mającego jednoznaczne rozwiązanie.

Ostatni etap — rozwiązanie układu równań liniowych sprawiać może dużo kłopotów, ze względu na możliwą dużą ilość niewiadomych, nawet wielu tysięcy. W opracowanym programie „RS" największa możliwa ilość niewiadomych wynosi ponad 33 tysiące. Konieczne jest więc stosowanie metod przybliżonych, wykorzystujących rzadkość i diagonalność powstającej macierzy [6,7,8]. W programie „RS" wykorzystano jedną z metod iteracyjnych — metodę kolejnych przybliżeń.

4. BŁĄD METODY I OBLICZEŃ NUMERYCZNYCH

Wynik — rozkład otrzymanej funkcji w węzłach siatki, otrzymany metodą różnic skończonych zawiera błąd, na który składają się dwa składniki. Pierwszy związany jest z zastąpieniem równania różniczkowego przez równanie różnicowe i jest funkcją długości siatki (błąd siatki); drugi wynika ze skończonej dokładności rozwiązania powstającego układu równań (błąd zaokrągleń). Wielkość tych błędów nie może być obliczona dokładnie, można jedynie oszacować ich przybliżoną wartość. Wykazuje się [4,9], że błąd wprowadzany przez zastosowanie siatki o równych ramionach jest mniejszy, niż wprowadzany przez siatkę o różnych ramionach. Dlatego pożądane jest stosowanie siatki prostokątnej tylko w sytuacjach, gdy jest to konieczne, na przykład przy brzegach lub granicach między ośrodkami nie pokrywających się z linią przechodzącą przez węzły siatki. Oszacowanie błędu siatki δ_s wykonuje się korzystając ze wzoru określającego błąd maksymalny, nieprzekraczalny w żadnym z punktów badanego obszaru [9]:

$$\delta_s = \frac{M_4 h^2 r^2}{24} \tag{14}$$

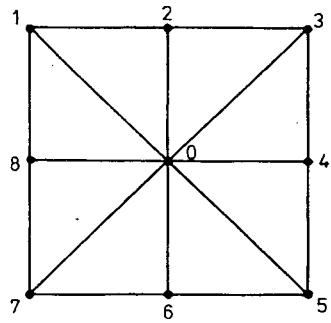
gdzie:

M_4 — maksymalna wartość bezwzględna czwartej pochodnej, liczonej w węzłach siatki, umieszczonych w kole o promieniu r.

Przydatność powyższego wzoru nie jest jednak duża, ze względu na konieczność znajomości czwartej pochodnej. Próby obliczenia jej wartości na podstawie wzoru różnicowego nie zawsze dają prawidłowe wyniki. Zastosowana metoda oceny błędu siatek polega na wykorzystaniu bardziej dokładnego równania różnic — dla dziewięciu węzłów, jak na rys. 5. Wtedy różnicowe przybliżenie laplasjanu wyraża się następującym wzorem:

$$\nabla^2 \Phi \approx 4\Phi_1 + \Phi_2 + 4\Phi_3 + \Phi_4 + 4\Phi_5 + \Phi_6 + 4\Phi_7 + \Phi_8 - 20\Phi_0 . \tag{15}$$

Błąd siatek jest wtedy liczony jako maksymalna wartość bezwzględna wyrażenia (15) w tych wszystkich węzłach, w których odległości od sąsiednich węzłów są jednakowe.



Rys. 5. Układ węzłów przy oszacowaniu błędu siatki

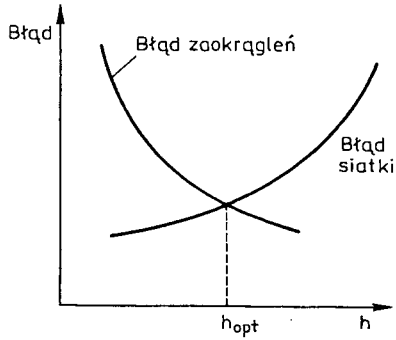
$$\delta_s = max|\nabla^2\Phi|. \tag{16}$$

Oszacowanie błędu wprowadzonego przez zaokrąglenia obliczeń w komputerze, przy iteracyjnym rozwiązaniu układu równań, można przeprowadzić w oparciu o następujący wzór [9]:

$$\delta_z = \frac{m_{res}\, r^2}{4h^2} \tag{17}$$

gdzie:
m_{res} to maksymalna bezwzględna odchyłka pomiędzy poprzednim, a właśnie wyliczonym przybliżeniem potencjału.

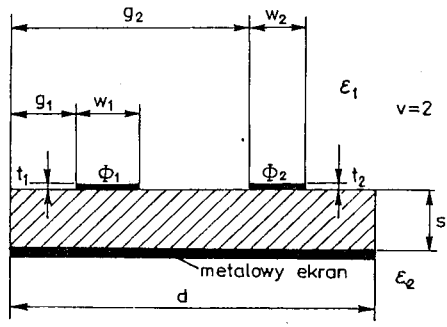
Jak widać błąd siatki i zaokrągleń zależy wprost od rozmiaru stosowanej siatki, co przedstawiono na rys. 6. Powinno się wybrać dostatecznie mały rozmiar siatki h, aby otrzymać wystarczającą dokładność metody. Z drugiej strony zbyt małe h daje duże błędy zaokrąglenia. Istnieją metody pozwalające na odpowiedni dobór rozmiaru siatki, w zależności od wielkości badanego obszaru i zmienności poszukiwanej funkcji [4,9].



Rys. 6. Zależność błędu metody od długości ramion siatki

5. ZASTOSOWANIE METODY DO ANALIZY LINII MIKROPASKOWYCH

Z rozważań teoretycznych [10,2] wynika, iż założenie rodzaju pola typu TEM w liniach paskowych jest wystarczająco ścisłe w zakresie częstotliwości, dla których rozmiary poprzeczne są małe w stosunku do długości fali. Ten sposób podejścia pozwala na sprowadzenie problemu analizy do rozwiązania dwuwymiarowego równania Laplace'a z funkcji potencjału w przekroju poprzecznym linii, przy oznaczeniach, jak na rys. 7. Niesymetryczna linia mikropaskowa



Rys. 7. Przekrój poprzeczny przez linię mikropaskową

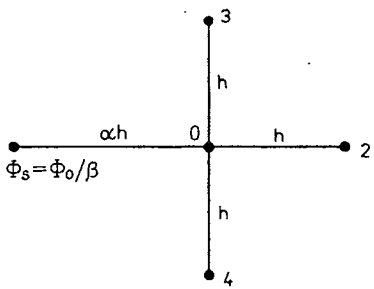
v — ilość pasków
Φ_i — potencjał poszczególnych pasków
$\mathbf{d}$ — szerokość struktury
s — wysokość warstwy dielektyka
t_i — wysokość kolejnych pasków
w_i — szerokość kolejnych pasków
g_i — odległość kolejnych pasków od lewego brzegu struktury
ε_1 — przenikalność elektryczna dielektryka $(\varepsilon_1 = \varepsilon_r \varepsilon_0)$
ε_2 — przenikalność elektryczna dielektryka $(\varepsilon_2 = \varepsilon_0)$

ekranowa jest tylko z jednej strony. Wynika stąd, że potencjał równany jest zeru dopiero w nieskończonej odległości od pasków. Tymczasem, aby zastosować metodę różnic skończonych, trzeba wyodrębnić obszar ograniczony, by móc pokryć go siatką obliczeniową. Jednym z możliwych rozwiązań tego problemu jest założenie istnienia hipotetycznego ekranu o zerowym potencjale, umieszczonego dostatecznie daleko. Powoduje to jednak nadmierne zwiększenie ilości węzłów w siatce.

Innym rozwiązaniem, zastosowanym w programie „RS", jest symulacja dodatkowych punktów leżących poza wyodrębnionym obszarem [4]. Sposób obliczenia wartości poszukiwanej funkcji opiera się na założeniu o jednostajności zmian potencjału — warunku brzegowym Neumanna. Na przykład dla węzła Φ_0 z rys. 8 jest to punkt Φ_s. Po przyjęciu, iż wartość potencjału w punkcie $\Phi_s = \Phi_0/\beta$ oraz, że odległość punktów Φ_0 i Φ_s wynosi ah, otrzymano następującą zależność obliczoną na podstawie wzoru (13):

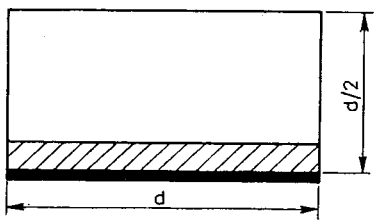
$$\Phi_0 := \frac{a}{a+1}\left[\left(\frac{\Phi_0}{a\beta} + \Phi_2\right)\frac{1}{a+1} + \frac{\Phi_3+\Phi_4}{2}\right]. \tag{18}$$

Powyższy wzór potraktowano w „sensie informatycznym" – jako przypisanie wielkości Φ_0 nowej wartości, obliczonej na podstawie starej jej wartości. Dla rozważanych struktur niesymetrycznych linii mikropaskowych można przyjąć, że $a=\beta=10$. Wynika to z przeprowadzonych przez autorów badań symulacyjnych. Jednak w opracowanym programie wartości tych wielkości można zmieniać dowolnie.



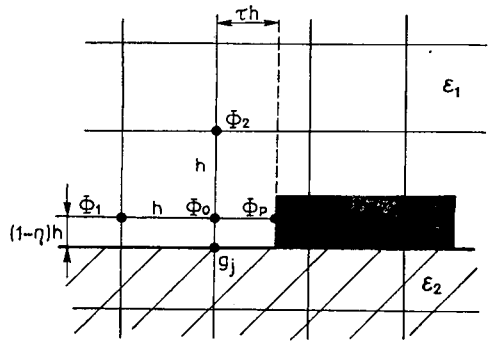
Rys. 8. Układ węzłów granicznych z symulowanym punktem o potencjale Φ_s

Aby zmniejszyć wpływ błędów wynikających z przyjęcia jednostajności zmian potencjału na granicach wyodrębnionego obszaru, przyjęto specjalny kształt brzegów siatki w postaci prostokąta o bokach d × d/2 (rys. 9). Na dolnym brzegu prostokąta



Rys. 9. Granica wyodrębnionego obszaru

potencjał opisany został zagadnieniem Ditrichleta ($\Phi=0$). Na pozostałych – w każdym ze skrajnych węzłów siatki założony został warunek brzegowy Neumanna. Takie zdefiniowanie brzegu niesie ze sobą pewne ograniczenia, co do rozmieszczenia pasków (np. nie mogą być one za wysokie i umieszczone za blisko brzegów struktury). Ze względu na skończone rozmiary siatki, wymagane są również odpowiednio duże rozmiary geometryczne pasków i odległości pomiędzy nimi [4]. Badany obszar pokryto kwadratową siatką różnicową, stosując prostokątne oczka tylko w okolicach granicy pomiędzy ośrodkami. Maksymalna ilość węzłów siatki zdefiniowana w programie „RS" wynosi (2^8+1) * (2^7+1) plus 257 węzłów na granicy ośrodków, co daje 33410. Rozwiązaniem powstałego układu równań jest rozkład potencjału w węzłach siatki.



Rys. 10. Jedna z możliwych sytuacji na granicy ośrodków

Algorytm postępowania w celu obliczenia pojemności jest następujący:
— obliczyć rozkład potencjału dla struktury rzeczywistej, zawierającej dwie warstwy dielektryczne;
— obliczyć rozkład potencjału dla hipotetycznej struktury, powstałej po usunięciu z linii warstwy dielektrycznej i wprowadzeniu w jego miejsce powietrza;
— na podstawie dwóch otrzymanych rozkładów potencjału i różnicowej postaci prawa Gaussa wyznaczyć można [(1,4]) ładunek swobodny Σq zgromadzony na pasku dla dwóch rodzajów linii. Stąd na podstawie znanej zależności $C = \Sigma q/U$, gdzie U jest jednostkowym napięciem pomiędzy paskiem a ekranem, można obliczyć pojemności dla linii o powietrznym i warstwowym ośrodku dialektycznym — odpowiednio C_0 i C_ε. Parametry charakterystyczne linii można obliczyć na podstawie następujących zależności [1,10], przy założeniu, że obecność dielektryka nie wpływa na indukcyjność struktury:
— impedancja charakterystyczna linii:

$$Z_0 = \frac{1}{c\sqrt{C_\varepsilon C_0}} \tag{19}$$

— prędkość propagacji fali elektromagnetycznej w linii:

$$v_p = c\sqrt{C_0}/\sqrt{C_\varepsilon} \tag{20}$$

— względna efektywna przenikalność elektryczna:

$$\varepsilon_{ef} = C_e/C_0 \tag{21}$$

— indukcyjność jednostkowa linii:

$$L_\varepsilon = L_0 = \frac{1}{C_0 c^2} \tag{22}$$

gdzie:

c — prędkość światła.

W programie „RS" określa się błąd obliczenia potencjału jako sumę błędów liczonych ze wzorów (17) i (14), po zastosowaniu różnicowego przybliżenia czwartej pochodnej. Następnie określa się błąd wyznaczenia pojemności liczonej z potencjału i na podstawie znanych zależności metrologicznych określa się błąd pozostałych wielkości, obliczanych na podstawie potencjału.

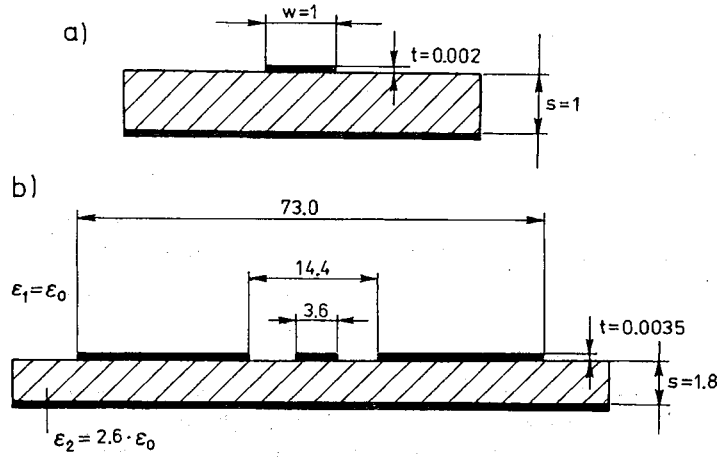
6. OPIS PROGRAMU „RS" I WYNIKI PRZYKŁADOWYCH OBLICZEŃ

Celem sprawdzenia poprawności podanych w p. 3, 4 i 5. zależności został opracowany program „RS" do obliczeń numerycznych. Proponuje się w nim cztery zdefiniowane rozmiary siatek, o odległościach pomiędzy węzłami równych d/32, d/64, d/128 i d/256, gdzie d jest szerokością rozpatrywanej struktury. Użytkownik programu może również zdefiniować własny rozmiar siatki, nie mniejszy jednak od d/256. Każdy punkt siatki kwalifikowany jest do odpowiedniej klasy, jednoznacznie określającej położenie względem czterech sąsiednich węzłów. Każdy typ z około 30 znalezionych, ma określone różnicowe przybliżenie odpowiedniego równania różniczkowego. Na przykład dla sytuacji na rys. 10 równanie różnicowe można zapisać w następujący sposób:

$$\Phi_0 = \frac{\tau(1-\eta)}{\tau-\eta+1}\left[\left(\frac{\Phi_p}{\tau}+\Phi_1\right)\frac{1}{1-\tau}+\left(\Phi_2+\frac{g_j}{1-\eta}\right)\frac{1}{2-\eta}\right] \tag{23}$$

gdzie:

τh – odległość punktu o potencjale Φ_0 od paska
$(1-\eta)h$ – odległość punktu o potencjale Φ_0 od granicy ośrodków
g_j – wartość potencjału w węźle granicznym
Φ_p – potencjał paska.



Rys. 11. Badane struktury i ich wymiary: a) niesymetryczna linia mikropaskowa. b) łatowy prostokątny element mikropaskowy, zasilany niesymetryczną linią mikropaskową w płaszczyźnie łaty „z wcięciem"

Tablica 1

Wyniki obliczeń impedancji charakterystycznej Z_c i względnej efektywnej przenikalności dielektrycznej ε_{ef} niesymetrycznej linii mikropaskowej o wymiarach podanych na rys. 11a

metoda obliczeń / parametry	metoda różnic skończonych obliczenia własne	metoda różnic skończonych [1]	metoda Hammerstad'a [12][1]	Bem, Katulski [11]
Z_0 [Ω]	49.4	51.62	49.74	49.7
δZ_0 [%]	0.6	–	0.8	–
ε_{ef}	7.02	–	6.45	6.49
$\delta\varepsilon_{ef}$ [%]	1	–	0.5	–

1) Obliczenia wykonano dla częstotliwości F = 1 GHz.

Tablica 2

Wyniki obliczeń impedancji charakterystycznej Z_c i względnej efektywnej przenikalności dielektrycznej ε_{ef} niesymetrycznej linii mikropaskowej o wymiarach podanych na rys. 11b

metoda obliczeń / parametry	metoda różnic skończonych obliczenia własne	metoda spektralna [10, 13]
Z_0 [Ω]	60.1	60.1
δZ_0 [%]	3.0	1.0
ε_{ef}	2.20	–
$\delta\varepsilon_{ef}$ [%]	1.5	0.5

Za pomocą programu „RS" obliczono podstawowe parametry elektryczne (Z_c, ε_{ef}) dwóch niesymetrycznych linii mikropaskowych, przedstawionych na rys. 11a i 11b. Wyniki obliczeń oraz sumaryczne dokładności wyznaczonych parametrów (δZ_c, $\delta\varepsilon_{ef}$) zestawiono w tablicach nr 1 i nr 2. W tablicach tych dla porównania podano wyniki obliczeń uzyskane w oparciu o wzory podane przez innych autorów [1,11,12,13]. Obliczenia własne programem „RS" przeprowadzono dla siatki o mającej 8 tysięcy węzłów. Wyniki obliczeń uzyskane za pomocą programu „RS" są zgodne z wynikami obliczeń innych autorów, można zatem uważać, że opracowane algorytmy są poprawne, a program jest w pełni przydatny do wyznaczania parametrów niesymetrycznych linii mikropaskowych i sprzężonych linii mikropaskowych o prawie dowolnej geometrii, w których rozchodzi się fala typu parzystego. Przykłady niesymetrycznych linii mikropaskowych, których parametry elektryczne nie mogą być obliczone programem „RS" podano w pracy [4].

Czas obliczeń (podany niżej dla komputera kompatybilnego z IBM XT 10 MHz z koprocesorem) zależy od ilości węzłów w siatce. Na przykład dla siatki o największej ilości węzłów (ponad 33000 węzłów) wynik uzyskano po 180 minutach, a dla siatki o około 500 węzłach po 0.7 minuty.
Program, napisany w języku C, jest typu „przyjaznego", obsługa jego jest prosta. Ma minimalne wymagania sprzętowe (IBM XT, 640 kB RAM, karta graficzna Hercules, pożądany jest koprocesor). W programie zawarta jest procedura rysująca rozkład potencjału w badanej linii. Dokładny opis programu znajduje się w [4].

WNIOSKI

1. Zaprezentowana w artykule metoda różnic skończonych oraz omówiony program „RS" umożliwiają obliczenia parametrów elektrycznych niesymetrycznych linii mikropaskowych o różnych kształtach przekroju przewodnika, np. prostokątnym, kołowym, trójkątnym, trapezowym. Dokładność obliczeń jest porównywalna z dokładnością uzyskiwaną w innych metodach, np. w metodzie spektralnej [10].

2. W artykule omówiono jedynie proste zastosowanie metody różnic skończonych dla obszaru dwuwymiarowego. Może stanowić on jednak istotny krok w kierunku rozwiązywania bardziej skomplikowanych zagadnień. Metody różnicowe są szczególnie przydatne w rozwiązywaniu problemów promieniowania anten i rozpraszania fal elektromagnetycznych [1]. W pracy [14] omówiono metodę różnic skończonych w zastosowaniu dla obszarów trójwymiarowych i w dziedzinie czasu. W pracy tej zaprezentowano wyniki obliczeń parametrów elektrycznych anteny mikropaskowej, nie zamieszczając jednak algorytmów, według których przeprowadzono obliczenia.

BIBLIOGRAFIA

1. M. Iskander i inni: *Computational Methods in Electromagnetics.* IEEE Transactions on Educations, vol. 31, No. 2, May 1988, pp. 101–114
2. T. Morawski, W. Gwarek: *Teoria pola elektromagnetycznego.* Warszawa: 1978 WNT
3. E. Kącki: *Równania różniczkowe cząstkowe w zagadnieniach fizyki i techniki.* Warszawa: 1989 WNT
4. P. Klonecki: *Zastosowanie metody różnic skończonych do analizy niesymetrycznych linii paskowych.* Praca dyplomowa. Instytut Telekomunikacji i Akustyki Politechniki Wrocławskiej. Wrocław 1990
5. M. Salazar-Palma i inni: *Finite element analysis of general inhomogeneous, anisotropic and multiconductor transmission lines – lines structures employing an efficient self adaptive mesh scheme.* Proc. 19th European Microwave Conference. Sept. 1989, Stokholm, pp. 52–54
6. A. Samarski, J. Nikołajew: *Metody rozwiązywania równań siatkowych.* Warszawa: 1988, PWN
7. M. Łukaniszyn: *Przegląd szybkich procedur różnic skończonych do rozwiązywania zagadnień stacjonarnych opisanych równaniem Poissona.* Rozprawy elektrotechniczne, 1988, t. 34, zeszyt 3, ss. 15–21
8. M. Dyja, J., M. Jankowscy: *Przegląd metod i algorytmów numerycznych,* cz. 2. Warszawa: 1982. WNT

9. P. Browne, C. Hannaford: *Electric and Magnetic Field Problems.* W. Saunders Comp. Philadelphia, London, Toronto, 1975
10. A. Sawicki, K. Sachse: *Uogólniona metoda spektralnej quasistatycznej analizy linii paskowych.* Rozprawy Elektrotechniczne, 1986, t. 32, z. 4, ss. 1139–1153
11. D. J. Bem, R. Katulski: *Anteny paskowe.* Rozprawy Elektrotechniczne. 1988, t. 34, z. 1 ss. 237–259
12. S. Rosłoniec: *Algorytmy projektowania wybranych liniowych układów mikrofalowych.* Warszawa: 1987. WKiŁ
13. W. J. Krzysztofik: *Notch feed microstrip antenna and arrays.* 9th Symposium on Electromagnetics Compatibility, Wrocław, June 1988, pp. 313–318
14. B. Jecko i inni: *Microstrip antennas analysing using FDTD method.* Proc. 19th European Microwave Conference. Sept. 1989, Stockholm, pp. 40–42

P. KLONECKI, Z. LANGOWSKI

FINITE DIFFERENCE METHOD APPLICATION TO MICROSTRIP TRANSMISSION LINES

Summary

This paper deals with the finite-difference method. Various problems in a modern microwave technology that are formulated by both differential equations (including partial differential equations) and by integral equations can be solved with this method. The method can be also useful if boundary conditions are sophisticated.

The equations for an arbitrary arrangement of grid nodes in a multilayer structure were found. An estimation of grid and computation errors is included.

An application method to the quasistatistic analysis of microstrip lines, single or coupled, nonsymetrical, printed on a uniform substrate, coupled with grounded adjacent lines, is investigated.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1992, 38, z. 4, ss. 537-547

681.32

Systolic deconvolvers based on Nahman-Guillaume regularization

TOMASZ ADAMSKI

Instytut Podstaw Elektroniki, Politechnika Warszawska

Otrzymano 1992.02.14

Autoryzowano do druku 1992.04.10

Fast deconvolving circuits implemented as systolic arrays are proposed in the paper. Regularization techniques (Nahman-Guillaume method) are applied to obtain stable and robust results of computations. Presented circuits can be designed as ASICs and may be very useful in digital correction of measurement and signal restoration systems. Applications to equivalent-time waveform recorders are shortly discussed.

1. INTRODUCTION

Many problems related to digital correction of measurement systems and digital signal processing lead finally to the deconvolution problem (i.e. solution of the convolution equation (1), where $t \in R$ and $f,gh \in L^1(R)$) or more general to solving Fredholm integral equation of

$$\int_R h(t-x)f(x)\,\mathrm{d}x = g(t), \tag{1}$$

the first kind (2) where $f,g \in L^2([a,b])$, $k \in L^2([a,b]) \times [a,b])$, $a,b \in \bar{R}$ and f is an unknown.

$$\int_{[a,b]} k(t,x)f(x)\,\mathrm{d}x = g(t), \tag{2}$$

These equations belong to the category of so called ill-posed problems. To explain what is it „ill-posed problem", we have to define at first the notion of well-posed problem. The concepts of well-posedness and ill-posedness were formalized at the

beginning of this century by J. Hadamard. In general, well-posed problem is an operator equation $Kf=g$ (where $K:X\to X$ is an operator, X is a metric function space, g a given function, f an unknown function and $f,g\in X$) for which three following conditions are satisfied:

1) K is a surjection, 2) K is an injection, 3) An inverse operator K^{-1} is continous.

If at least one of the three above conditions is not fulfilled the problem (i.e. equation $Kf=g$) is called ill-posed. In numerical practice small value changes of data (i.e. changes of g and K) are quite natural because of quantization and noise phenomena occuring in data aquisition systems. In ill-posed problems these small changes can lead to significantly different solutions. In other words solutions are very sensitive on input data.

Example 1. The Fredholm integral equation of the first kind is the well known classical example of ill-posed problem. Assume for instance, that we want to solve the equation (3), where the kernel $k:[0,\pi]\times[0,\pi]\ni(x,s)\to k(x,s)=x\cdot\sin(s)\in R$. The function $f:[0,\pi]\ni x\to f(x)=1/2$ is a solution of (3) but also each of the functions $f_n:[0,\pi]\ni x\to f_n(x)\sin(nx)+1/2$ for $n=1,2,\ldots$ is a solution. The problem (3) is ill-posed because the integral operator given by the formula (3) is not an injection.

$$\int_0^{\pi} k(x,s)f(s)\,\mathrm{d}s = x \tag{3}$$

Example 2. If the integral operator K from the equation (2) has a continous kernel k then $K(C([a,b])\subseteq C([a,b])$ and for example the equation (4) (where the kernel k is given in the example 1 and 1(x)

$$\int_0^{\pi} k(x,s)f(s)\,\mathrm{d}s = 1(x-\pi/2) \tag{4}$$

is a step function) has no continnous solution. The integral equation (4) is the ill-posed problem because the integral operator given by (4) is not a surjection.

Example 3. If an operator $K:X\to Y$ is a compact bijection and X,Y are ∞-dimensional B-spaces then the operator $K^{-1}:Y\to X$ is not continuous and the equation $Kf=g$, for $g\in Y$ is an ill-posed problem. The operator from Fredholm equation of the first kind is compact then this equation is an ill-posed problem.

Many other examples of ill-posed problems arising in practical situation are given in [2] (computer tomography), [4]–[7] (image and signal restoration), [1],[7],[13],[19] (correction of measurement systems) and [2],[7] (tutoral review of ill-posed problems).

Several numerical techniques called regularization have been recently developped to overcome the difficulties due to the ill-posedness (see for instance [1]–[7],[12],[13]). The essence of the regularization consists in replacing an ill-posed equation by another in some sense close to the first but well-posed and not so sensitive on input

data changes. For example the Fredholm equation of the first kind (problem ill-posed) is replaced by the Fredholm equation of the second kind (problem well posed). Ufortunatelly, all regularization techniques are very time consuming. For example in such a typical problem as digital correction of sampling oscilloscope by deconvolution, deconvolving algorithm takes (for 1024 samples of input signal) about ten minutes on IBM PC–AT (12 MHz) computer. On the other hand in many applications the time of deconvolution is crucial point. Then we need fast deconvolving circuits implementing deconvolution with regularization in real time. In the sequel such circuits (deconvolvers) are described thoroughly. The main part of the proposed deconvolvers is a systolic array for M point DFT. For simplicity reasons the Nahman-Guillaume's algorithm (see [1],[19]) was applied as regularization method but other regularization techniques as for instance Tikchonov or Phillips regularization can be also implemented as systolic deconvolvers in a similar way.

2. SYSTOLIC CIRCUITS FOR DFT COMPUTATION

A well known useful tool to obtain solutions of convolution equations (1) is the Fourier transform and particularly, in numerical computation case, the descrete Fourier transform i.e. DFT. A systolic digital array for M-points DFT computation for finite sequences of complex samples $\bar{x}=[x(1), x(2),\ldots,x(M)]$ is shown in fig. 1.

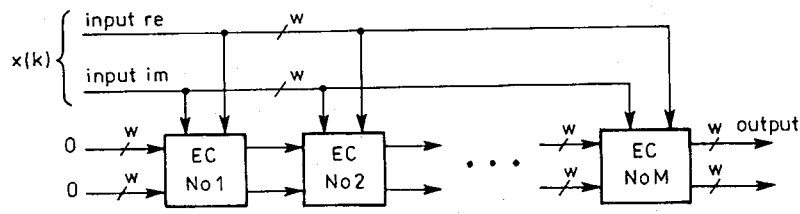


Fig. 1. Systolic circuit for M points DFT computation, EC denotes the elementary cell

The elementary cell of the proposed structure is shown in fig. 2. The circuit computes the M-dimensional complex vector $\mathrm{DFT}(\bar{x})=[\mathrm{DFT}(\bar{x}(1), \mathrm{DFT}(\bar{x})(2),\ldots$ $\ldots\mathrm{DFT}(\bar{x})(M)]=A,\bar{x}^T$, where $A=[W^{(k-1)(n-1)}]$ is a $M\times M$ matrix with $k,n=1,2,\ldots,M$ and $W=\exp(-j2\pi/M)$. The circuits can process data in parallel (in this case parameter w from fig. 1 is equal to the length of the used computer word) or in serial (more precisely with serial/parallel architecture for which $w=1$). The circuit is linear one dimensional systolic array composed from identical cascaded elementary cell EC_n. Elementary cells EC_n for $n=1,2,\ldots,M$ are identical but information (binary words) written in R3 registers (of each EC_n) specifically control each cell. For each $n=1,2,\ldots,M$ elementary cell EC_n has in R3 register written value W^n. Elementary cell EC_n is composed of the circuit generating consecutive values $W^{-n(n-1)n}.W^{(k-1)n}$ for $k=1,2,\ldots,M$, complex multiplier, complex adder and two registers R1, R2. In the

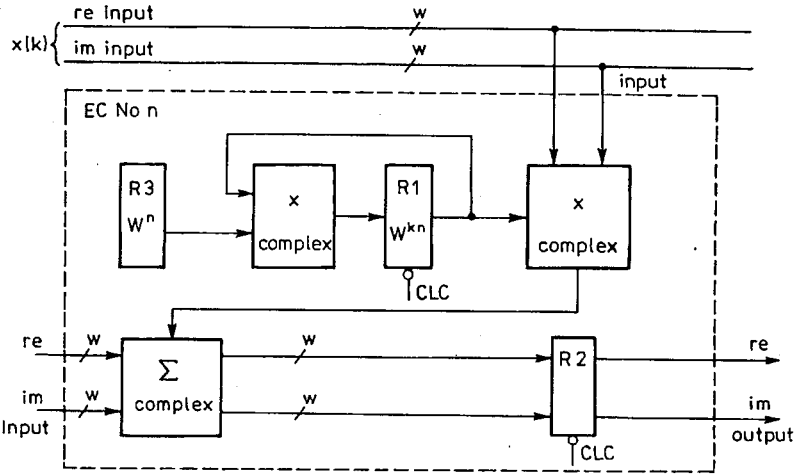


Fig. 2. The n-th elementary cell EC_n of the systolic circuit for DFT computation, k is a clock cycle number, $W^n = \exp(-j2\pi/M)$

initial state of the circuit in R1 register of each elementary cell EC^n are written for $n=1,2,\ldots,M$ values $W^{-n(n-1)}$. The last in the systolic line elementary cell EC_M may be simplified because W^{Mi} for arbitrary $i \in Z$ is equal to 1, then the generating part of the EC_M (R3, R1 and complex multiplier) is surplus. When the circuit is based on 2's complement notation the initial state must be set after every 2M clock cycles. In the case QRNS (Quadratic Residue Number System) notation, which is known as very useful in complex multiplication (see [11]), we do not need to return periodically to the initial state, but this special solution is not analyzed in the sequel. The initial states of R2 registers are arbitrary. The circuit works in 2M clock cycles during which we obtain on the output the vector DFT($\bar{x}$). More thoroughly in last M cycles we have consecutivly on the output DFT($\bar{x}$)(1), DFT($\bar{x}$)(2), DFT($\bar{x}$)(2),..., DFT($\bar{x}$)(M). The circuit computes these values directly from DFT definition creating at once M points of the DFT. Input data i.e. the complex vector $\bar{x}$ must occur on the input two times as a sequence: $\bar{y}=(\bar{x},\bar{x})=(x(1),x(2),\ldots,x(M),x(1),x(2),\ldots,x(M) \in C^{2M}$. Let $B=[b_{kn}]$ denotes the band matrix $M \times 2M$ with the band width M, defined by the formula

$$b_{kn} = \begin{cases} 0 & \text{for } 1 \leqslant n \leqslant k-1 \text{ or } n \geqslant k+M \\ W^{(k-1)n} & \text{for } k \leqslant n \leqslant k+M \end{cases} \tag{5}$$

and let $\bar{b}(k)$ for $k=1,2,\ldots,M$ denotes k-th column vector of the matrix B (i.e. $B=[\bar{b}(1),\bar{b}(2),\ldots,\bar{b}(M)]$ then of course $DFT(\bar{x})=B\cdot(\bar{y})^T$. The algorithm implemented by the circuit can be described recurrently in consecutive clock cycles in the following way $\bar{a}(0):=(0,0,\ldots,0) \in R^M$, $\bar{a}(k):=\bar{a}(k-1)+\bar{b}(k)\cdot x(k)$, $DFT(\bar{x})=\bar{a}(2M)$, where $k=1,2,\ldots,2M$, is a number of clock cycle. After M cycles on the output of the circuit we have $a(M)(1)=DFT(\bar{x})(1)$, after $M+1$ cycles $a(M+1)(2)=DFT(\bar{x})$, and so on during M consecutive clock cycles until the clock cycle number 2M. There are some

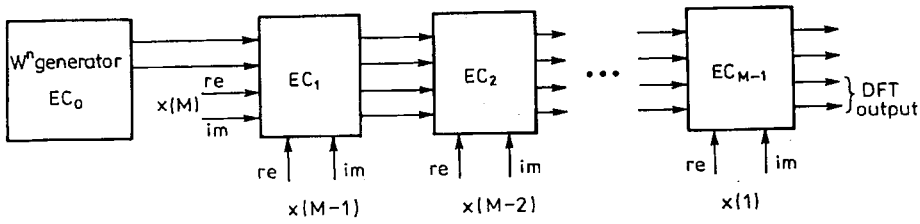


Fig. 3. Systolic circuit for M points DFT computation based on the Horner's algorithm, EC_0, EC_1, EC_2, ..., EC_M denotes elementary cells

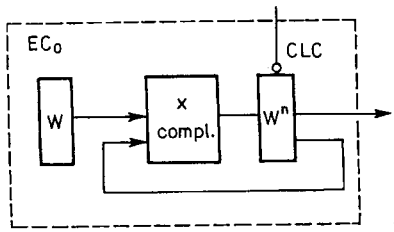


Fig. 4. The elementary cell EC_0 of the systolic circuit for DFT computation, based on the Horner's algorithm

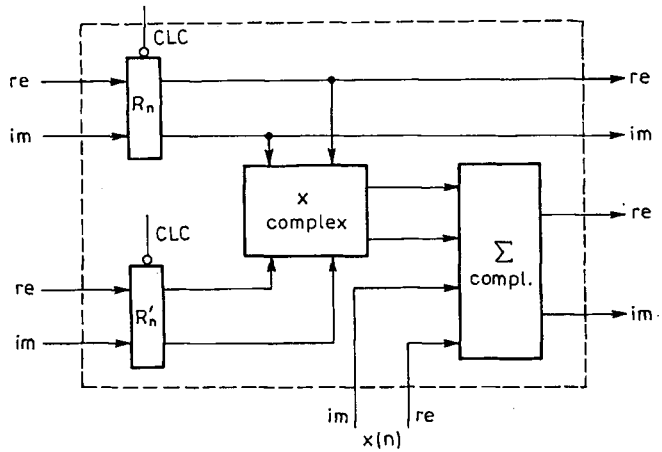


Fig. 5. The elementary cell EC_n(n = 1,2,...,M) of the systolic circuit for DFT computation, based on the Horner's algorithm

other concepts of fast systolic circuits which can be useful in DFT computation. For example we can find the DFT as a result of multiplication $A \cdot (\bar{x})^T$, where $A = [W^{(k-1)(n-1)}]$, with systolic circuit for matrix and vector multiplication (see Mead, Conway [9], Kung, Withehouse [10]). For economy reasons, for such a solution special circuits are needed to generate consecutive rows of the matrix A. Another useful concept makes use of the Horner's algorithm for fast computation of polynomial value in a given point. If $a_n t^n + a_{n-1} t^{n-1} + \ldots + a_1 t + a_0$ is a polynomial

with coefficients in a arbitrary field K and $F(t)$ denotes its value in a fixed point $t \in K$ then $F(t) = = (...(((a_n t + a_{n-1})t + a_{n-2})t + ...)t + a_0$. This equation is known as the Horner's algorithm. Systolic hardware implementation of the Horner's algorithm adopted to DFT computation is shown in fig. 3. Elementary cells EC_0 and EC_n for $n = 1,2,...,M$ are shown appropriatly in fig. 4 and fig. 5. Both 2's complement notation and QRNS approach can be used in these circuits.

Some concepts of fast hardware DFT circuits based on the FFT algorithm are discussed widely in [10],[16] and [17].

3. DECONVOLVING CIRCUITS

The proposed deconvolving circuit uses the Nahman-Guillaume method of regularization (see [1],[19]). The block scheme of the deconvolver is shown in Fig. 6. The records of the M samples are the input data and represent the function g from the formula (1). More exactly, input data is the vector $\bar{g} = [g(0),g(\triangle t),...$ $...,g((M-1)\triangle t)] \in R^M$, where $1/\triangle t$ is a frequency of sampling. On the output as a result of data processing we obtain the discrete version of f i.e. the solution of the equation (1). Let $\bar{h}$ denotes the M dimensional vector (M samples of h) describing the function h from (1). For each consecutive input data record the vector

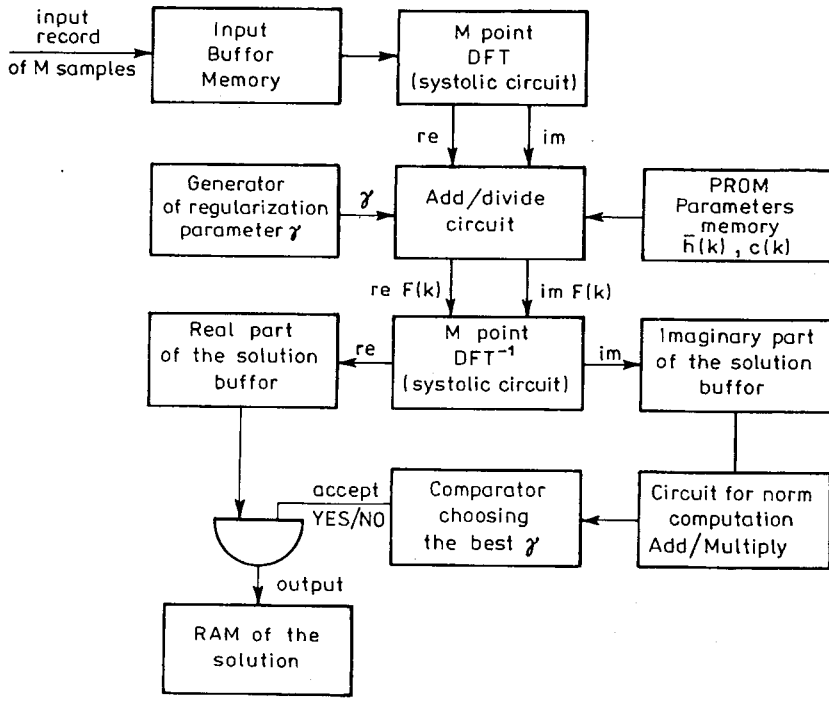


Fig. 6. Deconvolving circuit based on the Nahman-Guillaume regularization algorithm

$DFT(\bar{g})=[DFT(\bar{g})(1), DFT(\bar{g})(2),\ldots,DFT(\bar{g})(M)]$ is computed for choosen regularization parameter γ and then the sequence of values

$$F(k) = (DFT(\bar{g})(k) / DFT(\bar{h})(k) \cdot R(k)) \text{ for } k = 1,2,\ldots,M \tag{6}$$

where $R(k)=(DFT(\bar{h})(k))^2/((DFT(\bar{h})(k))^2+\gamma|C^2(k)|)$ and

$$|C^2(k)| = 6 - 8\cos(2\pi k/M) + 2\cos(4\pi k/M)$$

is computed in a special add/divide circuit. Then for obtained sequence $F(1),F(2),\ldots,F(M)$ the inverse DFT is computed along with the Eucliden norm (in R^M) of imaginary part of the inverse DFT. This last value can be treated as a quality coefficient of deconvolution. The above procedure is repeated $K+1$ times for $\gamma=\gamma_0,\gamma_1,\gamma_2,\ldots,\gamma_K$, where we assume $\gamma_k=\gamma_{k-1}+\Delta\gamma$ for $k=1,2,\ldots,K$ and fixed $\Delta\gamma>0$. As a descrete solution of (1) we admit the real part of the inverse DFT for this parameter γ, which gives the least value of the Eucliden norm of the immaginary part of DFT.

Note, the same circuit can be used for DFT and DFT^{-1} computation. Multiplexing or two separate similar circuits can be applied.

4. APPLICATIONS TO EQUIVALENT TIME WAVEFORM RECORDERS

Digital oscilloscopes with stroboscopic way of sampling (see [14],[15],[18]) or in other words equivalent time waveform recorders (ETWR) play important role in wide band (at present up to 20 GHz) measurements. In ETWR working in nanosecond region, there are some troublesome effects like:

- blowby
- nonlinearity of time axis
- sampling time jitter (i.e. random unstability of sampling)
- limited sampling head bandwidth

which significantly influences measured waveforms and decrease measurement accuracy of the whole system. Methods for the blowby and time axis nonlinearity corrections are shortly described in [18].

Assume, these corrections are done. Measurement distortions caused by sampling time jitter and sampling pulse of the head can be removed, provided the data (probability distributions and pulse shape) describing jitter and head are known i.e. were previously identified. More accurately, sampling time jitter and sampling head corrections can be implemented as appropriate independent deconvolutions (see equations (10),(11),(12)). The input record put to the deconvolver circuit is an arithmetic average of N_0 input records of the measured signal f. We assume in the sequel that principles of ETWR are well known to the reader then we do not explain basic facts concerning stroboscopic way of sampling.

Accepted in the paper equivalent linear model of sampling head is shown in Fig. 7. In ETWR the measured input voltage signal $f \in C(R)$ is repated many times (not

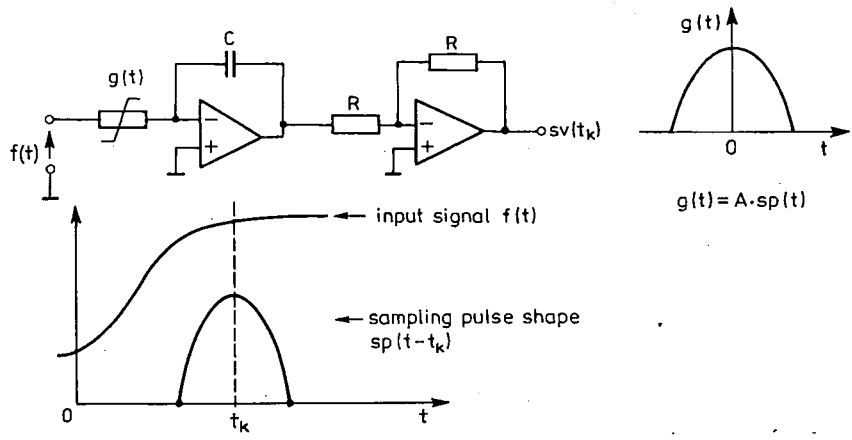


Fig. 7. Linear model of the sampling head in an equivalent time waveform recorder

obligatory in periodic way) and then sampled with stroboscopic method in deterministic (equivalent time) points t_k for $k=1,2,\ldots,M$. If a sample is taken in a time point t then the sample value sv(t) is equal to

$$sv(t) = \int_R \mathrm{sp}(x-t)f(x)\,\mathrm{d}x = (\tilde{s}\mathrm{p}*f)(t), \tag{7}$$

where $sp \in C(R)$ is the sampling pulse and $\tilde{sp}$ is defined as $\tilde{sp}(t)=sp(-t)$. In reality i.e. in the presence of sampling time jitter (STJ), true sampling time t'_k differs from assumed deterministic one t_k and for $k=1,2,\ldots,M$ we have $t'_k=t_k+T(t_k)$ instead of t_k, where $T(t)$ is a random variable describing the shift between the real sampling time t' and assumed one t. As a result of measurements corrupted by STJ, we obtain the finite random sequence of samples $sv(t_1+T(t_1))$, $sv(t_2+T(t_2))$, ..., $sv(t_M+T(t_M))$ insted of deterministic one $sv(t_1)$, $sv(t_2)$, ..., $sv(t_M)$. Taking into account conditions of measurements (i.e. the time which elapses between consecutive samplings), we can assume that the random variables $T(t_1)$, $T(t_2)$, ..., $T(t_M)$ are independent. Input signal measurements in equivalent time points t_1, t_2, ..., t_M are done N_0 times then all sampling process with jitter can be described by the following double sequence or random matrix

$$(sv(t_k+T_n(t_k)))_{(k,n)\in\langle 1,M\rangle\times\langle 1,N_0\rangle}, \tag{8}$$

where random variables $T_n(t_k)$ for $k\in\langle 1,M\rangle$, $n\in\langle 1,N_0\rangle$ are independent and for each $k\in\langle 1,M\rangle$, probability distribution of $T_k(t_k)$ is equal to the distribution of the random variable $T(t_k)$. All samples of the sequence $(sv(t_k+T_n(t_k)))_{n=1}^{N_0}$ for fixed $k\in\langle 1,M\rangle$ are displayed in the same time point on the ETWR screen or equivalently belong to the same class corresponding to the sampling time t_k. As a result the input signal image is scaterred or fuzzy as is shown in [18]. The simplest method to supress this troublesome effect is arithmetic averaging of samples as in the formula (9). But such a method, acceptable from image esthetic point of view, introduces significant

$$\mathrm{MN}(\mathrm{N}_0,t_k) = \left(\sum_{i=1}^{\mathrm{N}_0} sv(t_k+\mathrm{T}(t_k))\right)/\mathrm{N}_0 \tag{9}$$

errors to the input signal measurement. When applied averaging (9) can be considered as ideal (i.e. N_0 is sufficiently) large we can replace for fixed t_k the value $\mathrm{E}(sv(t_k+\mathrm{T}(t_k)) \stackrel{\mathrm{def}}{=} \overline{sv}(t_k)$ by $\mathrm{MN}((\mathrm{N}_0,t_k)$ and we obtain

$$\overline{sv}(t_k) = \int_{\Omega} sv(t_k+\mathrm{T}(t_k))\,\mathrm{dP} = \int_{\Omega}\left(\int_{\mathrm{R}} sp(x+\mathrm{T}(t_k))f(x)\,\mathrm{dx}\right)\mathrm{dP} = \tag{10}$$

$$= \int_{\mathrm{R}}\left(\int_{\mathrm{R}} sp(x-y)f(x)\,\mathrm{dx}\right)P_k(\mathrm{dy}) = \mathrm{MN}(\mathrm{N}_0,t_k),$$

where $(\Omega,\mathcal{M},\mathrm{P})$ is a probabilistic space for random variables $\mathrm{T}(t_1), \mathrm{T}(t_2), \ldots, \mathrm{T}(t_\mathrm{M})$ and P_k the probability measure induced by the random variable $t_k+\mathrm{T}(t_k)$. If the random variable $\mathrm{T}(t_k)$ has the density $g(\cdot,t_k)$ then

$$\overline{sv}(t_k) = \int_{\mathrm{R}} sv(t_k+x)\,g(x-t_k,t_k)\,l_1(\mathrm{dx}) = \mathrm{MN}(\mathrm{N}_0,t_k), \tag{11}$$

where l_1 is the Lebesgue measure on R. If $g(\cdot.t)$ do not depends on t then (11) can be rewritten as (12), where $\tilde{g}(x)=g(x,t)$

$$\overline{sv}(t_k) = \int_{\mathrm{R}}\left(\int_{\mathrm{R}} sp(x-y)f(x)\,\mathrm{dx}\right)\mathrm{P}_k(\mathrm{dy}) = \int_{\mathrm{R}}\left(\int_{\mathrm{R}} sp(x-y)f(x)\,\mathrm{dx})\,g(y-t_k,t_k\right)l_1(\mathrm{dy}) =$$

$$= (\tilde{sp}*f*\tilde{g})(t_k) = \mathrm{MN}(\mathrm{N}_0,t_k) \tag{12}$$

Formulas (10),(11) and (12) allow us to assess errors introduced by averaging of the jittered signal and are good start point to remove distortions caused by STJ and the sampling head. If we knew the STJ probability distributions $g(\cdot,t_k)$ and the pulse shape sp we could apply one of deconvolution methods to solve the equation (11) (Fredholm equation of the first kind) or the convolution equation (12). Then solving (11) or (12) we could find accurate shape of the measured signal i.e. the function f. In other words we can correct distortions caused in ETWR by sampling time jitter and the sampling head using appropriate DSP method.

Software implementation of the above described algorithm of deconvolution gives good results in sampling time jitter correction systems used in digital oscilloscopes.

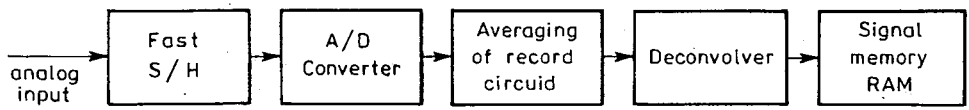


Fig. 8 Application of a deconvolver to sampling head and jitter correction in an equivalent time waveform recorder

But pure software solution, as we mentioned, is too slow to be fully useful in practice. Only fast hardware deconvolving circuits, like proposed in the section 3, creates possibility of real time correction. Computer simulation and experimental accuracy assessment prove that the deconvolution method reduced about ten times error of the input signal mesurement when compared with simple averaging method (see [19]).

CONCLUSIONS

1) Decribed in the paper systolic deconvolving circuits have high degree of parallelization then they are fast and well suited to the real time digital signal processing.
2) For presented hardware solutions, all three conditions characteristic for systolic arrays: modularity, locality and parallelism of data processing are fulfilled. As a consequence the proposed circuits can be easily implemented as one or several ASICs in VLSI technology.
3) Described applications of systolic deconvolvers to ETWR systems prove that this kind of circuits could be very useful in measurement and instrumentation.(Fig. 8)

LIST OF SYMBOLS

N – the set of natural numbers
Z – the set of integer numbers
R – the set of real numbers
$\bar{R}$ – the compactified set of real numbers $\bar{R}=R\cup(-\infty,+\infty)$
$L^p(R)$ – $p\geqslant 1$, the set of intergrable (with power p) functions
$C([a,b])$ – the set of all continous function on the interval $[a,b]$
$(\Omega,\mathcal{M},P)$ – the probabilistic space
l_1 – the Lebesgue measure on R

REFERENCES

1. W. L. Gans: *The Measurement and Deconvolution of Time Jitter in Equivalent Time Waveform Samplers.* IEEE Trans. on Instr. and Measur., March 1983, No 1
2. B. Hoffman: *Regularization for Applied Inverse and Ill-posed Problems.* Teubner-Texte zur Mathematik; Band 85, Leipzig 1986
3. C. W. Groetsch: *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind.* Pitman Research Notes in Mathematics; No 105, Boston, 1984
4. A. N. Tichonov, A. W. Gonczarski, W. W. Stiepanow, A. G. Jagoła: *Regularization Algorithms; (in Russian).* Moskwa: Nauka, 1983
5. A. B. Bakuszyński, A. W. Gonczarski: *Ill-posed Problems (in Russian);* Moscow University Publishers, Moskwa, 1989
6. A. B. Bakuszyński, A. W. Gonczarski: *Iterative Methods for Ill-posed Problems (in Russian);* Moskwa: Nauka, 1989
7. R. Morawski: *Metody odtwarzania sygnałów pomiarowych.* Metrologia i Systemy Pomiarowe; Politechnika Warszawska, 1989
8. B. White, M. Negahbon, C. Chen: *Design DSP into an ASIC.* Electronic Design, April 1990

9. C. Mead, L. Conway: *Introduction to VLSI Systems.* Addison-Wesley, 1982
10. S. Y. Kung, W. J. Whitehouse: *VLSI and Modern Signal Processing.* Prentice-Hall, 1985
11. A. Pieżyńska: *Arytmetyka resztowa liczb zespolonych i jej wykorzystanie w procesorze DFT.* Praca dyplomowa, Wydział Elektroniki Politechniki Warszawskiej, Warszawa 1988
12. A. Dyka: *Liniowa filtracja rozplotowa względem sygnałów o skończonym czasie trwania.* Zeszyty Naukowe Polit. Gdańskiej; Gdańsk 1989
13. A. Leśniewski: *Cyfrowa korekcja błędów dynamicznych powstających w układach wejściowych przyrządów pomiarowych.* Rozpr. doktorska Politechnika Warszawska, Warszawa 1980
14. J. Baranowski: *Synchroskopy stroboskopowe szerokopasmowe.* Warszawa: WNT, 1972
15. A. Riabinin: *Stroboscopic Oscillography (in Russian).* Moscow, 1972
16. H. Hikawa, V. K. Jain: *20 Milion Samples/s Wafer Procesor FFT Architecture*; In: Signal Processing Theories and Applications, L. Torres, E. Masgrau editors. Elservier Science Publishers, 1990
17. D. E. Kirk, J. G. Verly: *An Algorithm for Distributed Computation of FFT.* Comp. & Electr. Engineering 1990, Vol. 13, No 2
18. K. Zięcina: *Error Correction in High-Bandwidth Digital Sampling Oscilloscopes.* Proceedings of the 14-th KKTOiUE Waplewo, 1991
19. T. Adamski: *Cyfrowa korekcja chwilowej niestałości momentu próbkowania w rejestratorach szybkich impulsów.* Proceedings of the 12-th KKTOiUE Myczkowice, 1989

T. ADAMSKI

SYSTOLICZNE UKŁADY ROZPLATAJĄCE WYKORZYSTUJĄCE ALGORYTM NAHMANA-GUILLAUME'A

W pracy zaproponowano szybkie układy rozplatające zrealizowane w postaci sieci systolicznych. Ponieważ problem rozwiązania równania splotu jest tzw. problemem źle postawionym, konieczne jest wykorzystanie w algorytmie rozplotu metod regularyzacji. Jako metodę regularyzacji stosowaną w opisywanych sieciach systolicznych przyjęto algorytm Nahmana-Guillaume'a. Proponowane układy mogą być wykonane jako programowane układy typu ASIC stanowiąc użyteczny podsystem ułatwiający konstrukcję cyfrowych systemów pomiarowych. W pracy opisano również zastosowanie proponowanych układów do korekcji błędów wprowadzających przez chwilową niestałość momentu próbkowania w oscyloskopach cyfrowych.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1992, 38, z. 4, ss. 549-556

621.376.233

Some remarks on the phase detector circuit

TOMASZ ADAMSKI

Instytut Podstaw Elektroniki, Politechnika Warszawska

Otrzymano 1992.01.15

Autoryzowano do druku 1992.03.04

The paper deals with an output signal properties of the voltage multiplier circuit (MC) of the phase detector (PD), when input signals of the PD are non-synchronized i.e. periodic with different frequences. A mean value and spectral measure of the MC output signal are computed. These parameters are useful in PLL systems design.

1. INTRODUCTION

Phase locked loop systems (PLL) can be in locked (synchronized) or unlocked (unsynchronized) state. In the synchronized case, analysis of the PLL systems is relatively well elaborated both in deterministic or noised environment (see for example [3],[4],[5],[6] or [7]). In the unsynchronized state situation is much more complicated and leads, as we shall see, to ergodic theory or stochastic processes problems. The paper deals with an output signal properties of the multiplication circuit (MC) of the phase detector (PD) when input signals of PD are nonsynchronized i.e. periodic with different frequences. Such a situation is typical to the unlocked PLL systems. The block diagram of the analyzed PD is shown in the Fig. 1. The circuit consist of MC and low pass filter. The principles of this classical circuit when input signals are both rectangular or sinusoidal with the same frequency are well known. The MC output signal is a voltage signal denoted as $U_b(t)$ (see Fig. 1).

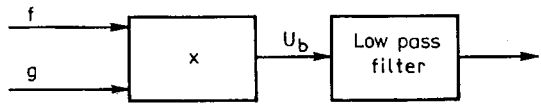


Fig. 1. The phase detector based on the voltage multiplier analyzed in the paper

If the input frequences are different, but both with the mean value zero, then the signal on the MC output has the mean value equal toro too. In the sequel theoretical motivations for this fact are presented. We assume in all paper that PD input signals $f:R\to R$ and $g:R\to R$ are Borel measurable, bounded and periodic with periods respectively T_1 and T_2. The above general formulation includes of course such particular signals f, g as rectangular or sinusoidal ones. In the section 2, we treat the input signals f, g as determinic and in the section 3 as random (Fig. 2).

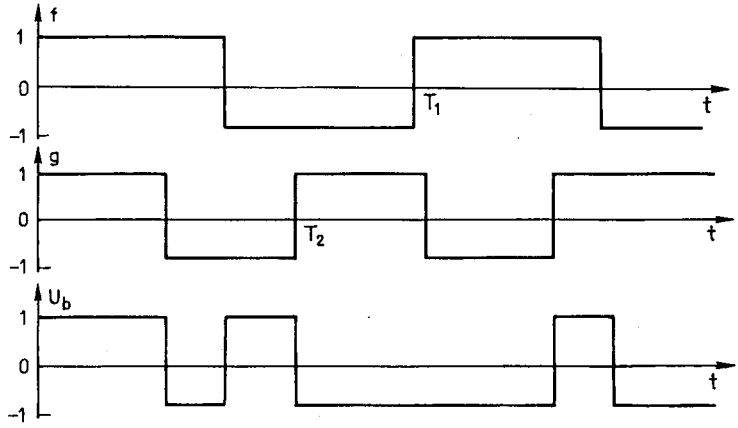


Fig. 2. Input signals f i g as rectangular waveforms

2. THE MEAN VALUE OF THE MC OUTPUT SIGNAL

Let $a\in R^+\setminus\{0\}$ and $x\in R$ and denote by $[x]_a$, x modulo a i.e. such a real number $0\leqslant y<a$ that there is an integer $k\in Z$ fulfilling the equation $x=k\cdot a+y$. Assume, we choose randomly the frequency $f_1=1/T_1$ of the input signal f from an interval $[f_{min}f_{max}]$ and the frequency $f_2=1/T_2$ of the input signal g from an interval $[f'_{min}f'_{max}]$. Let this choosing be described by independent (or not independent) random variables X, Y with continuous joint probability destribution given by a density function $h:[f_{min}f_{max}]\times[f'_{min}f'_{max}]\to R^+$. Hence the probability $P(X/Y=T_2/T_1\in R\setminus Q)$ is equal to 1. So, it is reasonable to admit that $T_2/T_1\in R\setminus Q$ and this assumption is valid in the sequel.

Let $S_1=[0,1)$ be the unit lenth circle, $\mathscr{L}_1$ a σ-field of Lebesgue measurable sets of S_1, l_1 Lebesgue measure and $G_a:S_1\to S_1$ for $a\in R\setminus Q$ a translation defined as $G_a(x)=[x+a]_1$. A dynamical system $(S_1, \mathscr{L}_1, l_1, G_a)$ is ergodic. For the same reason a dynamical systems $(S_{T_1}, \mathscr{L}_{T_1}, \mu, H_a)$ is ergodic, where $S_{T_1}=[0,T_1)$, $\mathscr{L}_{T_1}$ is a σ-field of Lebesgue measurable sets of S_{T_1}, $\mu=\frac{1}{T_1}l_1$ and $H_a:[0,T_1)\to[0,T_1)$ is given by the formula $H_a(x)=[x+a]_{T_1}$ for $a=T_2/T_1$. Thus from the Birhoff-Thinchin theorem (see suplement) applied to $T=H_a$ and the input signal f,we obtain for l_1 almost all $x\in R$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n-1}f([x+kT_2]_{T_1})=\frac{1}{T_1}\int_{[0,T_1]}f(x)\,l_1(\mathrm{dx})\,. \tag{2.1}$$

Because the functions f and g are measurable and bounded then for arbitrary $p\geqslant 1$ and $T>0$ we have $f,g\in L^p([0,T],\mathscr{L}_T,\nu)$ where ν is a finite measure on the interval [0,T] (in particular Lebesgue measure) and $\mathscr{L}_T$ is σ-field of Lebesgue measurable sets. Hence the function $t\to g(t)f(t)$ i.e. the MC output signal is l_1 integrable on $[0,T]$. We want to compute the mean value (2.2) of the MC output signal i.e. the mean value of the function $t\to g(t)f(t)$.

$$\lim_{T\to\infty}\frac{1}{T}\int_{[0,T]}f(x)\,g(x)\,l_1(\mathrm{dx})=\lim_{n\to\infty}\frac{1}{nT_2}\int_{[0,nT_2]}f(x)\,g(x)\,l_1(\mathrm{dx})=$$

$$=\lim_{n\to\infty}\frac{1}{nT_2}\sum_{k=1}^{n-1}\int_{[0,T_2]}f([x+kT_2]_{T_1})\,g(x)\,l_1(\mathrm{dx})\,. \tag{2.2}$$

Multiplying both sides of the equation (2.1) by a second input signal i.e. the function g, we obtain

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n-1}f([x+kT_2]_{T_1})\cdot g(x)=\frac{1}{T_1}g(x)\int_{[0,T_1]}f(x)\,l_1(\mathrm{dx})\,. \tag{2.3}$$

For each $n\in N\cup\{0\}$, $f([x+nT_2]_{T_1})\cdot g(x)\in L^1([0,T_2),\mathscr{L}_{T_2},l_1)$ then integrating both sides of the equation (2.3), dividing by T_2 and applying Lebesgue limit theorem we have

$$\frac{1}{T_2}\int_{[0,T_2]}g(x)\,\nu_2(\mathrm{dx})\,\frac{1}{T_1}\int_{[0,T_1]}f(x)\,l_1(\mathrm{dx})=$$

$$=\frac{1}{T_2}\int_{[0,T_2]}\left(\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n-1}f([x+kT_2]_{T_1})\cdot g(x)\right)l_1(\mathrm{dx})= \tag{2.4}$$

$$=\lim_{n\to\infty}\frac{1}{nT_2}\sum_{k=1}^{n-1}\int_{[0,T_2]}(f([x+kT_2]_{T_1})\cdot g(x))\,\nu_2(\mathrm{dx})=$$

$$=\lim_{n\to\infty}\frac{1}{nT_2}\int_{[0,nT_2]}(f([x]_{T_1})\cdot g(x))\,\nu_2(\mathrm{dx})=$$

$$= \lim_{n\to\infty} \frac{1}{nT_2} \int_{[0,nT_2]} f(x)\cdot g(x))\, l_1(\mathrm{dx})\,.$$

The equations (2.2) and (2.4) imply the following equation (2.5).

$$\lim_{T\to\infty} \frac{1}{T} \int_{[0,T]} f(x)\, g(x)\, l_1(\mathrm{dx}) = \frac{1}{T_2} \int_{[0,T_2]} g(x)\, l_1(\mathrm{dx}) \frac{1}{T_1} \int_{[0,T_1]} f(x)\, l_1(\mathrm{dx})\,. \quad (2.5)$$

Hence, if one of the input signals f,g (or both) has the mean value equal to zero then the mean value of the $U_b(t)$ is equal to zero too.

3. SPECTRAL MEASURE OF THE MC OUTPUT RANDOM SIGNAL

In this section we treat the MC input and output signals as random i.e. real stochastic processes. Specifically, we can identify the PD input signal f (which is periodic with a period T_1) with a trajectory of a stationary stochastic process $(X(t))_{t\in R}$, where $X(t)=A_1\cdot f(t+\varphi_1)$ for each $t\in R$ and A_1, φ_1 are random variables. Similarly we can identify the input signal g (periodic with a period T_2) with a trajectory of a stochastic process $(Y(t))_{t\in R}$, where $Y(t)=A_2\cdot g(t+\varphi_2)$ for each $t\in R$, and A_2, φ_2 are random variables. Assume, that A_1, A_2, φ_1, φ_2 are independent. Hence the stochastic processes $(X(t))_{t\in R}$ and

Fig. 3. Spectral measure of the PD output signal $U_b(t)$ when input signals are nonsynchronized sinusoidal ones with frequencies $f_1=1/T_1$ and $f_2=1/T_2$

$(Y(t))_{t\in R}$ are idependent too. Assume additionally that the random variable φ_1 has the uniform probability distribution on the interval $[0,T_1)$ and the random variable φ_2 has the uniform probability distribution on $[0,T_2)$. From the theorem 2 from the supplement, we obtain that $(X(t))_{t\in R}$ and $(Y(t))_{t\in R}$ are stationary. In the sequel we assume that the random variables A_1 and A_2 have second moments i.e. $A_1,A_2\in L^2(\Omega,\mathcal{M},P)$. If $A_1,A_2\in L^2(\Omega,\mathcal{M},P)$ then the processes $(X(t))_{t\in R}$ and $(Y(t))_{t\in R}$ are also stationary in wide sense. Therefore, the MC output process $(X(t)Y(t))_{t\in R}$ is also stationary in wide sense. An autocorrelation function R_1 of the stochastic process $(X(t))_{t\in R}$ is given by the following formula (3.1)

$$R_1(t) \overset{df}{=} E(X(t)X(0)) = E(A_1^2 f(t+\varphi_1)f(\varphi_1)) = E(A_1^2)E(f(t+\varphi_1)f(\varphi_1)) =$$

$$= \frac{E(A_1^2)}{T_1} \int_{[0,T_1]} f(t+x)f(x)\, l_1(\mathrm{dx}) = B_1\cdot(\tilde{f}*f)(-t) \quad (3.1)$$

where $B_1 = E(A_1^2)/T_1$, $*$ is the circle convolution on $[0,T_1)$ and $\tilde{f}(t) \stackrel{df}{=} f(-t)$ for each $t \in R$. For the sinusoidal function f i.e. when $f(t) = \sin\left(\frac{2\pi}{T_1} t\right)$, we have

$$R_1(t) = B_1 \int_{[0,T_1]} \sin\left(\frac{2\pi}{T_1}(t+x)\right) \sin\left(\frac{2\pi}{T_1} x\right)_1 l_1(\mathrm{dx}) = E(A_1^2)/2 \cdot \cos\left(\frac{2\pi}{T_1} t\right). \quad (3.2)$$

For a rectangular shape of the function f, i.e. when $f(t) = \mathrm{sgn}\left(\sin\left(\frac{2\pi}{T_1} t\right)\right)$ we have

$$R_1(t) = B_1 \int_{[0,T_1]} \mathrm{sgn}\left(\sin\left(\frac{2\pi}{T_1}(t+x)\right)\right) \cdot \mathrm{sgn}\left(\sin\left(\frac{2\pi}{T_1} x\right)\right) l_1(\mathrm{dx}) =$$

$$= \begin{cases} B_1 \cdot (T_1 - 4[t]_{T_1}) & \text{for } t \in \{t;\ kT_1 \leqslant t \leqslant kT_1 + T_1/2,\ k \in Z\} \\ B_1 \cdot (-3T_1 + 4[t]_{T_1}) & \text{for } t \in \{t;\ kT_1 + T_1/2 \leqslant t \leqslant (k+1)T_1,\ k \in Z\}. \end{cases} \quad (3.3)$$

Similarly, the correlation function R_2 of the stochastic process $(Y(t))_{t \in R}$ is equal to

$$R_2(t) = E(A_2^2) \frac{1}{T_2} \int_{[0,T_2]} g(t+x)\, g(x)\, l_1(\mathrm{dx}) = B_2 \cdot (\tilde{g} * g)(-t), \quad (3.4)$$

where $\tilde{g}(t) \stackrel{df}{=} g(-t)$ for each $t \in R$ and $*$ is the circle convolution on the interval $[0,T_2)$ and $B_2 = E(A_2^2)/T_2$.

Having the autocorrelation functions R_1 and R_2, we can compute the autocorrelation function R_3 of the MC output process $(X(t)Y(t))_{t \in R}$

$$R_3(t) \stackrel{df}{=} E(X(t)Y(t)X(0)Y(0)) = E(X(t)X(0))E(Y(t)Y(0)) = R_1(t) \cdot R_2(t). \quad (3.5)$$

The mean value of the process $(X(t)Y(t))_{t \in R}$ is given by the following formula (3.6).

$$E(X(t)Y(t)) = E(X(t))E(Y(t)) = \frac{E(A_1)E(A_2)}{T_1 T_2} \int_{[0,T_1]} f(x) l_1(\mathrm{dx}) \cdot \int_{[0,T_2]} g(x) l_1(\mathrm{dx}). \quad (3.6)$$

We would like to find a spectral measure of the stochastic process $(X(t)Y(t))_{t \in R}$ i.e. such a finite Borel measure μ_3 on the real axis R that its Fourier transform (3.7) is equal to R_3.

$$R_3(t) = \int_R e^{j\omega t}\, \mu_3(\mathrm{d}\omega). \quad (3.7)$$

For sinusoidal f and g we have $R_3(t) = \frac{1}{4} E(A_1^2)E(A_2^2) \cdot \cos\left(\frac{2\pi}{T_1} t\right) \cdot \cos\left(\frac{2\pi}{T_2} t\right) =$ $\frac{1}{8} E(A_1^2)E(A_2^2) \cdot \left(\cos\left(\left(\frac{2\pi}{T_1} - \frac{2\pi}{T_2}\right)t\right) + \cos\left(\left(\frac{2\pi}{T_1} + \frac{2\pi}{T_2}\right)t\right)\right)$. Hence, it is obvious that the

spectral measure in this case is the following $\mu_3 = B_3 \cdot (\delta_{\omega_1} + \delta_{-\omega_1} + \delta_{\omega_2} + \delta_{-\omega_2})$, where $B_3 = E(A_1^2)E(A_2^2)/8$, $\omega_1 = \frac{2\pi}{T_1} - \frac{2\pi}{T_2}$, $\omega_2 = \frac{2\pi}{T_1} + \frac{2\pi}{T_2}$ and $\delta_x : B(R) \to R^+$ is a Dirac measure in the point $x \in R$, defined for each $A \in B(R)$ as $\delta_x(A) = 1$ if $x \in A$ and $\delta_x(A) = 0$ if $x \notin A$. Therefore, μ_3 is a descrete measure concentrated in atoms ω_1, ω_2, ω_{-1}, ω_{-2}.
The above method can be used with similar results for nonsinusoidal functions f and g (for example rectangular functions or linear combinations of sinusoidal functions). Applying Fourier series expansions to periodic autocorrelation functions R_1 and R_2, we obtain atoms of the spectral measure μ_3. But, the obtained descrete spectral measure μ_3 can have in this general case infinit number of atoms.
From the formula (3.6) it follows that if the mean value of the function f or g is equal to zero then the mean value of the process $(X(t)Y(t))_{t \in R}$ is equal to zero too.

CONCLUSIONS

We proved in the paper, that if the mean value of one of the unsynchronized PD input signals f or g is equal to zero then the mean value of the MC output signal is equal to zero too. Suggested method for a spectral measure computation can be useful in PLL systems design.

LIST OF SYMBOLS

N	– set of natural numbers
R	– set of real numbers
R^+	– set of nonnegative real numbers
Q	– set of rational numbers
R^+	– set of nonnegative real numbers
$(\Omega, \mathcal{M}, P)$	– probabilistic space
$B(R)$	– σ-field of Borel sets of R
l_1	– Lebesgue measure on R
$L^p(M, \mathcal{M}, \mu)$	– for $p \geqslant 1$, set of integrable functions with the power p
$[x]_a$	– x modulo a i.e. $[x]_{2\pi} = \min\{x - k2\pi \geqslant 0;\ k \in Z\}$.

REFERENCES

1. S. W. Fomin, I. P. Kornfeld, J. G. Sinaj: *Ergodic Theory*; (in Russian); Moscow 1980
2. W. Szlenk: *Introduction to Smooth Dynamical Systems* (in Polish); Warsaw: PWN, 1982
3. F. M. Gardner: *Phaselock Techniques.* Wiley & Sons; New York, 1979
4. A. J. Viterbi: *Principles of Coherent Communication.* Mc Graw-Hill, New York, 1966
5. W. C. Lindsey: *Synchronization Systems in Communication and Control.* Prentice-Hall, New Jersey, 1972
6. J. Klapper, J. Frankle: *Phase-locked and Frequency Feedback Systems.* Academic Press, London 1972
7. W. Tichonov, M. Mironov: *Markov Chains* (in Russian); Moscow, 1977

SUPPLEMENT

Theorem 1 (Birkhoff-Thinchin ergodic theorem)

Assume $(M,\mathcal{M},\mu)$ is a probabilistic space with the complete measure μ, $f\in L^1(M,\mathcal{M},\mu)$ and $T:M\to M$ is an endomorphism of the space with measure $(M,\mathcal{M},\mu)$. If a dynamical system $(M,\mathcal{M},\mu,T)$ is ergodic then for μ almost each $x\in M$, we have

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n-1}f(T^k x)=\int_M f(x)\,\mu(\mathrm{d}x),$$

where $T^k=\underbrace{T\circ T\circ T\circ\cdots\circ T}_{k}$

The proof can be found in [1].

Theorem 2

If $f:[0,2\pi)\to R$ is a $(B([0,2\pi),B(R))$ measurable function, $(A,\psi):\Omega\to R^2$ two dimensional random variable on a probabilistic space $(\Omega,\mathcal{M},P)$ and the random variable φ is independent of (A,ψ) has the uniform probability distribution on the interval $[0,2\pi)$ then the stochastic process $(A\cdot f([\psi t+\varphi]_{2\pi}))_{t\in R}$ is stationary. If additionally the function f is bounded and the random variable $A\in L^2(\Omega,\mathcal{M},P)$ then the stochastic process $(A\cdot f([\psi t+\varphi]_{2\pi}))_{t\in R}$ is stationary in wide sense.

Proof. 1) For each $t_1<t_2<\ldots<t_{2n}$, $t_1,t_2,\ldots,t_n\in R$ a function

$$G:R^3\ni(x_1,x_2,x_3)\to(x_1 f([x_2t_1+x_3]_{2\pi}),\ x_1 f([x_2t_2+x_3]_{2\pi}),\ \ldots,\ x_1 f([x_2t_n+x_3]_{2\pi}))\in R^n$$

is $(B(R^3), B(R^n))$ measurable because for each $i=1,2,\ldots,n$, function $\pi_i\circ G$ is $(B(R^3), B(R))$ measurable, where $\pi_i:R^n\ni(x_1,x_2,\ldots,x_n)\to x_i\in R$ i.e. π_i is a projection.

2) For each $x\in R$ a random variable $[x+\varphi]_{2\pi}$ has the uniform probability distribution on $[0,2\pi]$. Similarly, a random variable $\tilde{\varphi}\overset{df}{=}[\psi t+\varphi]_{2\pi}$ for arbitrary fixed $t\in R$ has the uniform probability distribution on $[0,2\pi)$. It follows from the Fubini theorem and the assumption that random variables ψ i φ are independent. More precisely, let $B\in B(R)$, and denote $C=\{(x_1,x_2)\in R^2;\ [x_1t+x_2]_{2\pi}\in B\}$ then we obtain

$$P([\psi t+\varphi]_{2\pi}\in B)=E(\chi_{\{[\psi t+\varphi]_{2\pi}\in B\}})=\int_{R^2}\chi_C(x_1,x_2)\,P_{(\psi,\varphi)}(\mathrm{d}(x_1,x_2))=$$

$$=\int_{R^2}\chi_C((x_1,x_2)\,P_\psi\otimes P_\varphi(\mathrm{d}(x_1,x_2))=\int_R\left(\int_R\chi_C(x_1,x_2)\,P_\varphi(\mathrm{d}x_2)\right)P_\psi(\mathrm{d}x_1)=$$

$$=\int_R P_\varphi(C_{x_1})\,P_\psi(\mathrm{d}x_1)=\int_R P([x_1t+\varphi]_{2\pi}\in B)\,P_\psi(\mathrm{d}x_1)=\frac{l_1(B\cap[0,2\pi])}{2\pi}$$

where C_{x_1} denotes the section of the set C for fixed $x_1\in R$, i.e.

$$C_{x_1}=\{x_2\in R;\ [x_1t+x_1]_{2\pi}\in B\}.$$

3) For arbitrary $B_1\in B(R)$, denote $A_1\overset{df}{=}\{[\psi t+\varphi]_{2\pi}\in B_1\}\in\sigma([\psi t+\varphi]_{2\pi})\subset\mathcal{M}$ and $B_2=\{(x,y);\ [xt+y]_{2\pi}\in B_1\}$. The random (A,ψ) i φ are independent then we have

$$P([\psi t+\varphi]_{2\pi}\in B_1\,|\sigma(A,\psi))=P(A_1|\sigma(A,\psi))=E(\chi_{A_1}|\sigma(A,\psi))=E(\chi_{B_2}(\psi,\varphi)|\sigma(A,\psi))=$$

$$=E(\chi_{B_2}(x,\varphi))|_{x=\psi}=P([xt+\varphi]_{2\pi}\in B_1)|_{x=\psi}=\frac{l_1(B_1\cap[0,2\pi))}{2\pi}=P(A_1).$$

Hence, for each $A_1 \in \sigma([\psi t+1])$ and $A_2 \in \sigma(A,\psi)$ we obtain $P(A_1 \cap A_2) = S_{A_1} PCA_1 | \sigma(A,\psi) \mathrm{dP} = P(A_1)P(A_2)$ then σ-fields $\sigma(\tilde{\varphi})$ i $\sigma(a,\psi)$ are independent.

4) To prove stationarity of the stochastic process $(Af[\psi t+\varphi]_{2\pi}))_{t\in R}$ we have to prove, that for each $n \in N$, each $B \in B(R^n)$ and each sequence $t,t_1,t_2,\ldots,t_n \in R$, $t_1 < t_2 < \ldots < t_n$ the following formula is true

$$P((Af([\psi t_1+\varphi]_{2\pi}), Af([\psi t_2+\varphi]_{2\pi}), \ldots, Af([\psi t_n+\varphi]_{2\pi})) \in B) =$$
$$= P((Af(\psi(t_1+t)+\varphi]_{2\pi}), Af([\psi(t_2+t)+\varphi]_{2\pi}), \ldots, Af(\psi(t_n+t)+\varphi]_{2\pi})) \in B),$$

or equivalently we have $P_{(A,\psi,\varphi)}(G^{-1}(B)) = P_{(A,\psi,\tilde{\varphi})}(G^{-1}(B))$. (*)
Because the random variables (A,ψ) and $\tilde{\varphi}$ are independent then $P_{(A,\psi,\tilde{\varphi})} = P_{A,\psi)} \otimes P_{\tilde{\varphi}}$ and similarly, because (A,ψ) and φ are independent then we have $P_{(A,\psi,\varphi)} = P_{(A,\psi)} \otimes P_{\varphi}$. The random variables $\tilde{\varphi}$ i φ have (as we proved it in the point 2) the same probability distribution, then (*) is true ■

T. ADAMSKI

UWAGI DOTYCZĄCE UKŁADU DETEKTORA FAZY

Summary

Przeprowadzono analizę własności sygnału wyjściowego układu mnożącego (wchodzącego w skład detektora fazy) w sytuacji, gdy na wejście układu podawane są dwa przebiegi okresowe niezsynchronizowane. Obliczona jest wartość średnia i miara spektralna sygnału wyjściowego. Zaprezentowane są dwa podejścia do problemu. Ujęcie wykorzystujące teorię ergodyczną dyskretnych układów dynamicznych i ujęcie probabilistyczne bazujące na fakcie, że sygnał wyjściowy układu mnożącego można uważać za realizację procesu stochastycznego stacjonarnego w szerszym i węższym sensie. Przeprowadzona analiza jest użyteczna przy badaniu zachowania się układów z fazową pętlą sprzężenia zwrotnego (układy PLL) w stanie braku synchronizacji.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1992, 38, z. 4, ss. 557-582

621.384
535.215.4

Dyskusja teorii fotoemisji

JÓZEF WOJAS*

Profesor emerytowany Politechniki Warszawskiej

WŁODZIMIERZ, LESŁAW WOJAS

Katedra Matematyki, Akademia Rolnicza (SGGW), Warszawa

Otrzymano 1992.04.30

Autoryzowano do druku 1992.07.17

W pracy przedyskutowano kolejne trzy teorie zewnętrznego zjawiska fotoelektrycznego w półprzewodnikach. Teoria Spicera stworzyła trzy stopniowy model fotoemisji, który będzie analizowany. Kane opierając się na ogólnej teorii Spicera powiązał fotoemisję z przejściami prostymi i skośnymi elektronów. Dyskutowana jest widmowa zależność $Y(h\nu)$ dla różnych warunków pobudzenia i ucieczki elektronów. Autor * badając doświadczalnie zależność wydajności kwantowej od energii fotonów potwierdził słuszność teorii Kane'a. Najnowsza teoria Ballantyne'a dostarczyła model fotoemisji oparty na wynikach pracy Kane'a, lecz uwzględniający straty energii związane z rozproszeniem na fononach. Daje to możliwości wyznaczania z danych wydajności fotoemisji parametrów rozproszenia gorących elektronów.

1. TEORIA FOTOEMISJI SPICER'A

Pierwszą mającą podstawowe znaczenie teorią fotoemisji jest teoria Spicera [1]. Opracował on prosty, podstawowy model fotoemisji z ciał stałych. Ten model był oparty na założeniu, że fotoemisja jest o wiele badziej efektem objętościowym niż czysto powierzchniowym [2;3]. Co do optycznej absorbcji prowadzącej do fotoemisji przyjęto, że jest ona charakterystyczna dla wnętrza ciała i że proces fotoemisji jest podzielony na trzy następujące po sobie zdarzenia:

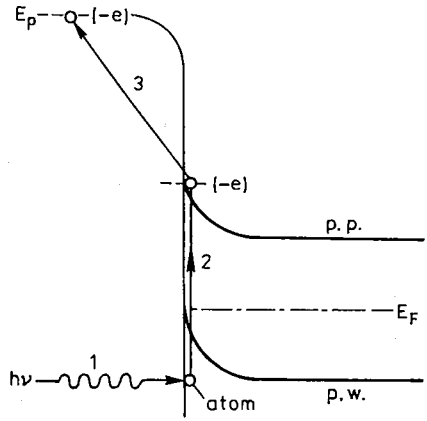
1) absorbcja optyczna,

2) transport pobudzonego elektronu do powierzchni (elektron może tracić energię podczas tego procesu) oraz

3) wyrzucenie elektronu przez powierzchnię do próżni.

Te etapy emisji są zilustrowane na rys. 1.

Ten model nie tylko dostarcza podstawy do zrozumienia praktycznego działania fotokatod, lecz prowadzi także do opracowania widzialnej i ultrafioletowej fotoemisji będących ważnym narzędziem badań podstawowych elektronowej struktury ciał stałych. Ostatnie teoretyczne badania fotoemisji wykazały, że ten trzy stopniowy model jest dobrym przybliżeniem, jak również, że założenie absorbcji optycznej we wnętrzu jest dobre.



Rys. 1. Współczesna interpretacja zjawiska fotoemisji zewnętrznej: proces 1 — absorpcja fotonu przez atom, 2 — przejście elektronu do pasma przewodnictwa przy powierzchni, 3 — przejście elektronu przez barierę powierzchniową na zewnątrz kryształu

W celu skonstruowania trzy stopniowego modelu fotoemisji na podstawach obliczeniowych muszą być napisane wyrażenia na prawdopodobieństwo w każdym z tych stopni (rys. 2). Iloczyn tych prawdopodobieństw będzie dawał całkowite prawdopodobieństwo fotoemisji. W celu napisania wyrażeń na te prawdopodobieństwa rozważmy warstwę materiału grubości dx w odległości x od powierzchni. Rozważmy najpierw zachodzące zdarzenie optycznego pobudzenia; należy zauważyć, że pewna część absorbcji optycznej może być zużyta na pobudzenie do stanów poniżej poziomu próżni, z czego nie jest możliwe otrzymanie fotoemisji (rys. 1 i 2). Zatem współczynnik absorbcji optycznej dla energii fotonów hν może być napisany jako

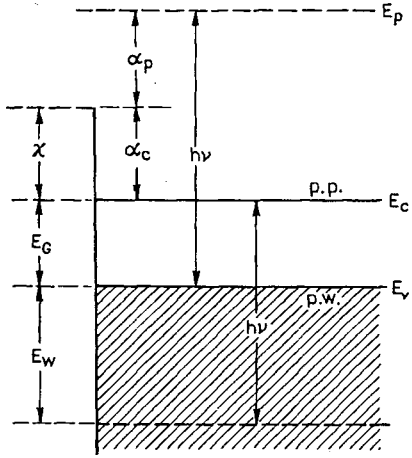
$$\alpha(h\nu) = \alpha_p(h\nu) + \alpha_c(h\nu) \tag{1}$$

gdzie α_p jest współczynnikiem absorbcji przy pobudzaniu do stanów powyżej poziomu próżni, a α_c jest współczynnikiem dla przejść poniżej poziomu próżni, natomiast prawdopodobieństwo absorbcji optycznej w warstwie jest:

$$dP_a(x, h\nu) = \alpha(h\nu)\, e^{-\alpha(h\nu)x}\, dx \tag{2}$$

a to może prowadzić do przedstawienia fotoemisji jako:

$$dP'_a(x, h\nu) = \alpha_p(h\nu)\, e^{-\alpha(h\nu)x}\, dx\,. \tag{3}$$



Rys. 2. Wykres pasm energetycznych fotoemitera: χ — powinowactwo elektronowe, E_G — przerwa energetyczna, $h\nu$ — energia padającego fotonu. Przejścia optyczne można podzielić na dwie grupy: a) do stanów powyżej poziomu próżni (stowarzyszone ze współczynnikiem absorpcji α_p b) i te do stanów poniżej poziomu próżni (stowarzyszone ze współczynnikiem absorpcji α_c, E_w jest najniższym poziomem energii w pasmie walencyjnym, z którego elektrony mogą być pobudzane do pasma przewodnictwa przez fotony o energii $h\nu$

Zostało wykazane, że prawodopodobieństwo osiągnięcia przez pobudzony elektron powierzchni z energią wyższą od energii poziomu próżni, czyli prawdopodobieństwo przejścia, jest dane przez:

$$P_p = A(h\nu)\, e^{-\frac{x}{L(h\nu)}} \tag{4}$$

gdzie $L(h\nu)$ jest charakterystyczną drogą (ucieczką) elektronu pobudzonego przez energię $h\nu$, $A(h\nu)$ zawiera się pomiędzy 1 a 1/2 zależnie od: 1) czy rozproszenie elektron-foton albo elektron-elektron określa wartość L, czy 2) energia pobudzonego elektronu jest określona względem poziomu próżni.

Jeśli dominuje rozproszenie elektron-elektron (przyjmując pobudzenie izotropowe co do kierunku) wtedy $A(h\nu) = 1/2$.

Jeśli dominuje rozproszenie elektron-foton i wszystkie elektrony są pobudzone do dostatecznie dużych energii, wtedy $A(h\nu)$ będzie w przybliżeniu równe 1.

Jeśli elektron osiągnie powierzchnię z energią powyżej poziomu próżni, nie ma jeszcze 100% prawdopodobieństwa, że zostanie on wyrzucony. Dla idealnej powierzchni prawdopodobieństwo będzie mniejsze od jedności, ponieważ:

1) tam jest skończone prawdopodbieństwo dla kwantowego odbicia przy powierzchni oraz

2) tylko składowa momentu pędu prostopadła do powierzchni uczestniczy w przekraczaniu powierzchniowej bariery potencjału.

Dla realnej powierzchni, gdzie mają miejsce rozproszenia elastyczne lub nieelastyczne — sytuacja jest nawet bardziej skomplikowana, stąd korzystnie jest bezpośred-

nio wprowadzić parametr prawdopodobieństwa ucieczki $P_u(h\nu)$, który może być określony doświadczalnie. Prawdopodobieństwo fotoemisji z warstwy położonej na głębokości x jest zatem równe:

$$dP(h\nu) = \alpha_p(h\nu)\, A(h\nu)\, e^{-[\alpha(h\nu)+1/L(h\nu)]}\, x_{P_u}(h\nu)\, dx\,. \tag{5}$$

Całkując wyrażenie (5) od 0 do ∞, otrzymujemy

$$P(h\nu) = Y_a(h\nu) = \frac{\dfrac{\alpha_p(h\nu)}{\alpha(h\nu)}}{1 + \dfrac{1}{L(h\nu)\,\alpha(h\nu)}}\, A(h\nu)\, P_u(h\nu)\,, \tag{6}$$

gdzie Y_a jest wydajnością na absorbowany foton. Często praktycznie jest napisać wydajność na padający foton:

$$Y_i(h\nu) = R(h\nu)\, Y_a(h\nu)\,. \tag{7}$$

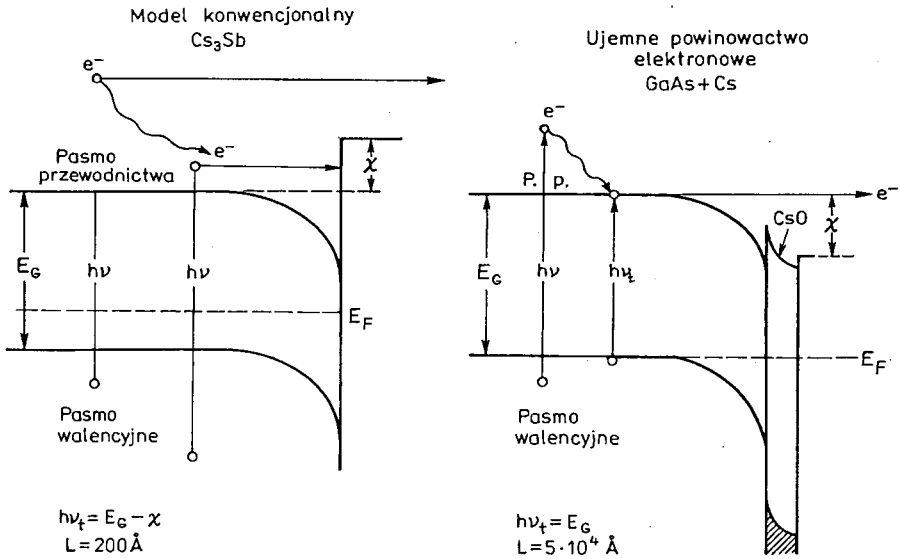
gdzie $R(h\nu)$ jest współczynnikiem odbicia od kryształu fotonów o energii $h\nu$.

Często przyjmuje się:

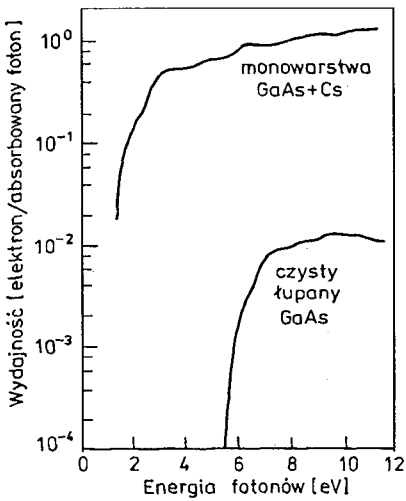
$$A(h\nu)\, P_u(h\nu) = B(h\nu) \tag{8}$$

a stąd

$$Y_i(h\nu) = \frac{\dfrac{\alpha_p(h\nu)}{\alpha(h\nu)}\, R(h\nu)\, B(h\nu)}{1 + \dfrac{1}{L(h\nu)\,\alpha(h\nu)}} \tag{9}$$



Rys. 3. Wykresy struktury pasmowej: model konwencjonalny Cs_3Sb i z ujemnym powinowactwem elektronowym fotoemitera GaAs+Cs



Rys. 4. Fotoemisja z GaAs z i bez jednoatomowej warstwy Cs na powierzchni. Jest to pewien przykład ważności (roli) znaczenia powinowactwa elektronowego κ w określeniu wydajności kwantowej półprzewodnika. Przyrost wydajności około dwa rzędy wielkości otrzymuje się dla energii fotonów powyżej 7 eV przez zredukowanie (zmniejszenie) κ z około 4 eV do 0 eV (wg W. E. Spicera)

Analiza wyrażenia (9) w ujęciu schematu pasmowego z rys. 2 i 3 będzie ważna dla czytelnika ze względu na minimalizację energii powinowactwa χ dającą powiększenia wydajności, na przykład czynnik $\frac{\alpha_p(h\nu)}{\alpha(h\nu)}$ będzie w przybliżeniu równy 1, gdy χ jest zredukowane do zera. Zauważmy też, że iloczyn $L(h\nu)$. $\alpha(h\nu)$ jest właśnie współczynnikiem absorbcji optycznej dla głębokości ucieczki $L(h\nu)$. Jeśli χ jest mniejsze, $L(h\nu)$ będzie rosło ponieważ (uśrednieniając) każdy elektron będzie mógł tracić większe ilości energii przy nieelastycznym zderzeniu, lecz będzie wyrzucany. Te efekty pokazuje rys. 4 [4 – 7].

2. TEORIA REDFIELDA

Należy wspomnieć o klasycznej teorii fotoemisji z obszaru ładunku przestrzennego półprzewodnika podanej przez Redfielda. Ściślej mówiąc Redfield [8] opracował teorię oceny wpływu obecności obszaru ładunku przestrzennego na fotoemisję z półprzewodników. Wykazał on, że taki wpływ w wielu wypadkach ma istotne znaczenie i prowadzi do następujących wniosków:

1) *„ogon fotoemisji"* powinien rozciągać się poza normalny obszar progu fotoemisji $h\nu_t$,

2) w ogólności energia krawędzi pasma walencyjnego nie może być wywnioskowana z obserwowanego progu,

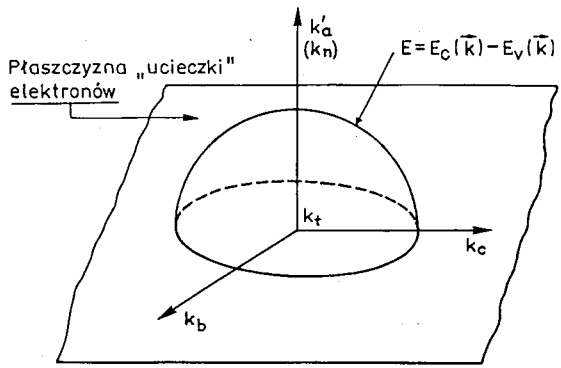
3) obszary ładunku przestrzennego dodatniego i ujemnego powodują efekty, które nie muszą być symetryczne; dopasowania wykładnicze do danych fotoemisyjnych w pobliżu progu są mało dokładne.

Redfield pokazał również (na bazie uzyskanych informacji), że niektóre uprzednio obserwowane anomalie mogą być wyjaśnione wpływem obszaru ładunku przestrzennego. Od dawna zdawano sobie sprawę, że efekty powierzniowe mogą mieć wpływ na zachowanie się zewnętrznej fotoemisji z półprzewodników. Taki wpływ może pochodzić z samych stanów powierzchniowych lub z sąsiadującego obszaru ładunku przestrzennego, w którym zaburzenie przypowierzchniowe zmienia niektóre z własności objętościowych materiału.

W swych rozważaniach Redfield przyjął za dominujące rozproszenie pobudzonego elektronu na fononach — w czasie przebywania przezeń (drogi) w przypowierzchniowym obszarze kryształu. Jego obliczenia dotyczące tego drugiego aspektu problemu (dowodzą) potwierdzają istnienie *„fotoelektrycznego efektu objętościowego"*; stąd uzyskane wnioski mają także wpływ na zagadnienie względnego stosunku fotoemisji objętościowej do fotoemisji powierzchniowej.

3. TEORIA FOTOEMISJI KANE'A

Kane przedyskutował ogólną teorię fotoemisji i podał w pracy [8] współczesną teorię fotoemisji zewnętrznej z półprzewodników uwzględniając przypadki przejść* prostych elektronów z możliwością rozproszeń elastycznych w obszarze ładunku przestrzennego i na powierzchni. Kane podał funkcje na wydajność fotoemisji w przypadku różnych mechanizmów pobudzenia i ucieczki elektronów z uwzględnieniem efektów powierzchniowych.



Rys. 5. Model fotoemisji w przestrzeni wektora k

Model powierzchni kryształu przedstawiony na rys. 5 zastosował Kane do przedstawienia fotoefektu zewnętrznego w przestrzeni wektora falowego $\vec{k}$. Powierzchnia „ucieczki elektronów" jest równoległa do rzeczywistej powierzchni półprzewod-

* Optyczne przejścia elektronów mogą być proste, zachodzące bez zmiany wektora falowego $\vec{k}$ i skośne, czyli ze zmianą wektora falowego elektronu.

nika. Elektrony o pędach położonych pod tą powierzchnią znajdują się w krysztale, a składowe normalne ich pędów oznaczono przez $\vec{k_n}$.

Punkty na powierzchni czaszy kulistej odpowiadają maksymalnym pędom elektronów emitowanych z półprzewodnika pod wpływem energii fotonów E. Ta czasza jest „powierzchnią energii optycznej" i opisana równaniem

$$E = E_c(\vec{k}) - E_v(\vec{k}), \tag{10}$$

gdzie E – bieżąca energia fotonów, E_c – energia emitowanych fotoelektronów względem poziomu próżni, E_v – energia elektronów na szczycie pasma walencyjnego.

Środek układu na rys. 5 odpowiada parametrom fotoelektrycznego progu: E_t – energia progowa ($h\nu_t$) fotonu; $\vec{k}_t$ – pęd fotoelektronów progowych.

Wyrażenie na energię elektronu przeniesionego do pasma przewodnictwa ($E_c > E_p$) przedstawiamy za pomocą następującej formuły:

$$E_c = \alpha(\vec{k_n})^2 + \alpha_1\vec{k_n} + \beta\vec{k_b^2} + \beta_1\vec{k_c^2} = E - E_t \tag{11}$$

gdzie $\alpha_1\vec{k_n}$ wprowadza element nieparaboliczności pasm przewodnictwa niektórych półprzewodników.

Aby związać energię E ze zmianami składowej $\vec{k_n}$, która decyduje o dokonaniu się emisji elektronu, można rozwinąć funkcję E na szereg Taylora w pobliżu punktu $\vec{k_t}$. Ograniczając szereg do dwu wyrazów:

$$E \simeq E_t + (\nabla_{\vec{k_n}}E)_{\vec{k_t}}(\vec{k_n} - \vec{k_{n0}}) \tag{12}$$

Z założenia $\vec{k_{n0}}$ równe jest zeru. Po wyznaczeniu stąd $\vec{k_n}$, następnie po podstawieniu do wzoru 11 i odpowiednich przekształceniach otrzymuje się

$$(E - E_t)\left[\frac{(\nabla_{\vec{k_n}}E)_{\vec{k_t}}^{\alpha_1}}{(\nabla_{\vec{k_n}}E)_{\vec{k_t}}}\right] = \alpha\vec{k_n^2} + \beta\vec{k_b^2} + \beta_1\vec{k_c^2}. \tag{13}$$

Zmieniając współczynniki przy zmiennych otrzymujemy równanie podane przez Kane [8] jako wstępne do ostatecznego wyprowadzenia funkcji wydajności $Y(E)$.

$$E - E_t = \beta'(\vec{k_n})^2 + \beta_b\vec{k_b^2} + \beta_c\vec{k_c^2} \tag{14}$$

gdzie β', β_b, β_c – stałe współczynniki.

Dalsze postępowanie w wyprowadzeniu funkcji $Y(E)$ sprowadza się głównie do przekształceń matematycznych. Zastosowano sferyczny układ współrzędnych i wyrażono składowe: $\vec{k_n}$, $\vec{k_b}$ i $\vec{k_c}$ oraz elementarny kąt bryłowy $d\Omega$ w funkcji kątów φ i Θ (rys. 5).

Po podstawieniu prostych zależności trygonometrycznych:

$$\beta' = \beta\cos\Theta$$
$$\beta_b = \beta\sin\Theta\cos\varphi$$
$$\beta_c = \beta\sin\Theta\sin\varphi$$

kwadrat modułu wektora $\vec{k}$ został wyprowadzony jako

$$(\vec{k})^2 = \frac{(E-E_t)}{\beta'\cos^2\Theta+\beta_b\sin^2\Theta\cos^2\varphi+\beta_c\sin^2\Theta\sin^2\varphi}. \tag{15}$$

Następnie przedstawiono elementarny przyrost wydajności dY „iloczynowo-ilorazowym" wyrażeniem trygonometrycznych funkcji φ i Θ, oraz współczynników β', β_b i β_c pomnożonym przez $\vec{k}^2$ dΩ. Z kolei, z całkowania dY otrzymuje się

$$Y = K(E - E_t), \tag{16}$$

gdzie K w funkcji energii E jest prawie stałym współczynnikiem i wynosi

$$K = \frac{const}{|\nabla_{\vec{k}_n} E|_{\vec{k}_t}(\beta_b\beta_c)^{1/2}\int_{sB}^{S}\delta[E_c(\vec{k}) - E_v(\vec{k}) - E]\,d\vec{k}}. \tag{17}$$

Otrzymano więc końcowy wzór Kane'a (17), który wskazuje na liniową zależność wydajności kwantowej od energii fotonów w przypadku objętościowych przejść prostych bez uwzględnienia rozproszeń elastycznych. Dalsze, bardziej złożone rozważania pozwoliły mu ustalić wykładnik potęgi funkcji Y(hν) jako 5/2 dla objętościowych przejść skośnych.

Kane podał więc widmową zależność Y(hν) dla różnych warunków pobudzenia i ucieczki elektronów.

W ogólności Kane [8;9] ustalił, że

$$Y(h\nu) = k(h\nu - h\nu_t)^n \tag{18}$$

gdzie n jest liczbą całkowitą i może przyjmować wartości od 2 do 7.

Wyniki teorii Kane były weryfikowane przez doświadczalne badania spektralnych rozkładów wydajności fotoemisji z kilku półprzewodników z grupy III – V [3;10].

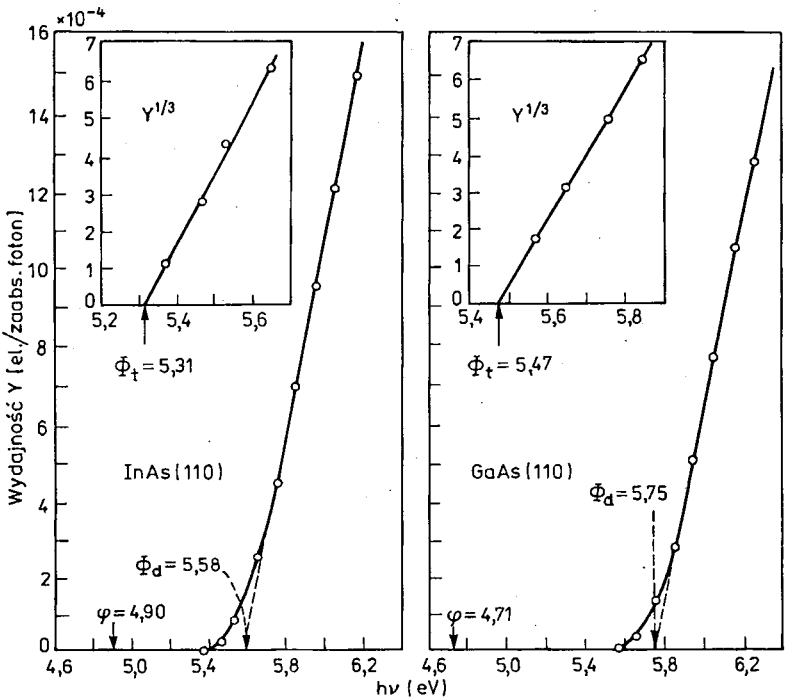
Przez dopasowywanie wykładników potęg do eksperymentalnie otrzymanych krzywych Y(hν) ustalono, że wydajność kwantową dla najniższych energii fotonów można zawsze opisać funkcją

$$Y = k(h\nu - h\nu_t)^3. \tag{19}$$

Potęgę 3, jako bliską wartości wykładnika 5/2 otrzymanego przez Kane, autorzy wspomnianych prac uzasadniają dominacją objętościowych przejść skośnych w przyprogowych obszarach wydajności. Przeprowadzone eksperymenty [5] wykazały, że trzecia potęga najlepiej pasuje do danych i że wydajność w pobliżu krawędzi (ponad progiem fotoemisji) spełnia równanie:

$$Y \propto (h\nu - h\nu_t)^{3+\frac{1}{2}} \tag{20}$$

gdzie właśnie hν_t jest zdefiniowane jako ekstrapolowana wartość progu wyznaczona poprzez prawo sześcienne (wykładnik – trzy), tak jak jest pokazane na rys. 6 we wstawkach u góry.



Rys. 6. Fotoelektryczna wydajność kwantowa w funkcji energii fotonów dla łupanych próbek InAs i GaAs (wg Gobeliego i Allena) ($h\nu_d$ — wartość krawędzi przejść prostych). Wstawiono eksploatacje progów prawem sześciennym

To prawo sześcianów wydaje się być dokładne do wielkości $\pm 0{,}5$ w wykładniku. Należy zauważyć, że położenie progu $h\nu_t$ otrzymanego z ektrapolacji zależy nieco od założonego prawa potęgowego, na przykład na rys. 7 wartość progu zmienia się od wielkości 5,56 eV dla $Y^{1/2}$, aż do 5,475 eV dla ekstrapolacji $Y^{1/4}$, to jednak przedstawia błąd jedynie $\pm 0{,}05$ eV z ekstrapolacji „sześciennej" i wobec tego mieści się w granicach błędu eksperymentalnego.

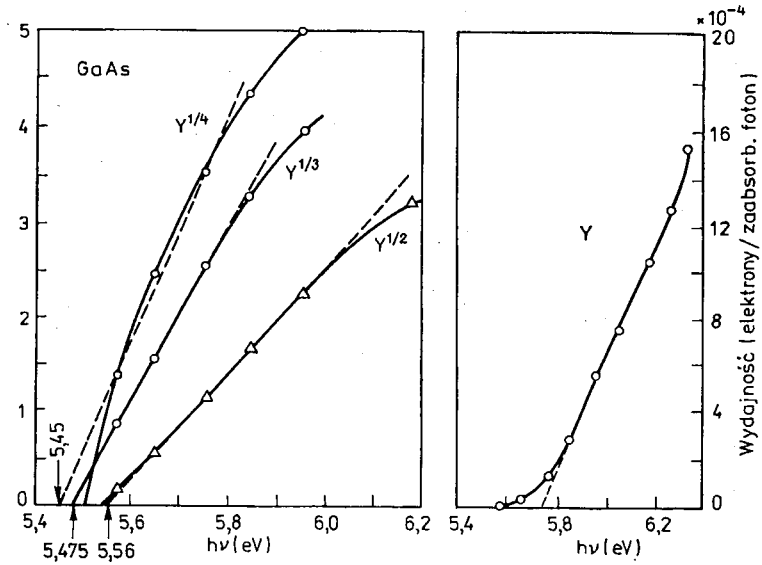
Teoria Kane'a zjawiska fotoelektrycznego [8] przewiduje różne postaci potęgowe dla widm wydajności w zależności od mechanizmu odpowiedzialnego za fotoemisję. Dla wydajności w pobliżu progu niżej wymienione procesy i odpowiadające im prawa potęgowe mogłyby odpowiadać za emisję:

1) Pobudzenie objętościowe, optyczne skośne:

$$Y \propto (h\nu - h\nu_t)^{5/2}$$

2) Stany objętościowe (powierzchnia jako absorber pędu):
   a) $Y \propto (h\nu - h\nu_t)^{3/2}$ dla powierzchni doskonałej
   b) $Y \propto (h\nu - h\nu_t)^{5/2}$ dla powierzchni rozpraszającej
3) Wzbudzenie optyczne proste (stany pasma powierzchniowego):

$$Y \propto (h\nu - h\nu_t)^{3/2}$$



Rys. 7. Prawo potęgowe przystosowane do fotoelektrycznej wydajności w pobliżu progu dla GaAs

4) Stany powierzchniowe związane z zakłóceniami, rozłożone w przedziale energii:

$$Y \propto (h\nu - h\nu_t)^2 .$$

Jest całkiem możliwe, że więcej niż jeden z tych procesów może brać udział w fotoemisji w pobliżu progu i w tym przypadku zmieszanie się procesów w różnych częściach z różnymi prawami potęgowymi spowoduje powstanie prawa potęgowego będącego średnią z indywidualnych procesów.

Z wymienionych i niewymienionych powodów nie jest łatwo określić mechanizm fizyczny odpowiedzialny za krzywe wydajności fotoemisji w pobliżu progu z zależności potęgowych. Jednak ponieważ próg fotoemisji $h\nu_t$ wyznaczany z ekstrapolacji prawa sześciennego pasuje do mierzonych najmniejszych wartości wydajności, można przyjąć, że powinien on dawać różnicę energii między energią poziomu próżni E_p i najwyższymi stanami energetycznymi zapełnionymi elektronami (o rozsądnej gęstości), czyli $h\nu_t$ z dokładnością $\pm 0{,}05$ eV.

Składnik niskiej wydajności scharakteryzowany przez wartość progową $h\nu_t$ odpowiada najwyższym zapełnionym stanom elektronowym półprzewodnika. W zasadzie, zapełnione stany energetyczne leżące między wierzchołkami pasma walencyjnego i poziomem Fermiego mogą być źródłem fotoelektronów. W ten sposób $h\nu_t$ musi odpowiadać albo krawędzi fotoelektronu Φ^*, albo wartości między Φ i φ, która zaznacza poziom takich zapełnionych stanów powierzchniowych.

* Wielkość Φ jest różnicą energii elektronu w najwyższym stanie pasma walencyjnego na powierzchni i na poziomie E_p

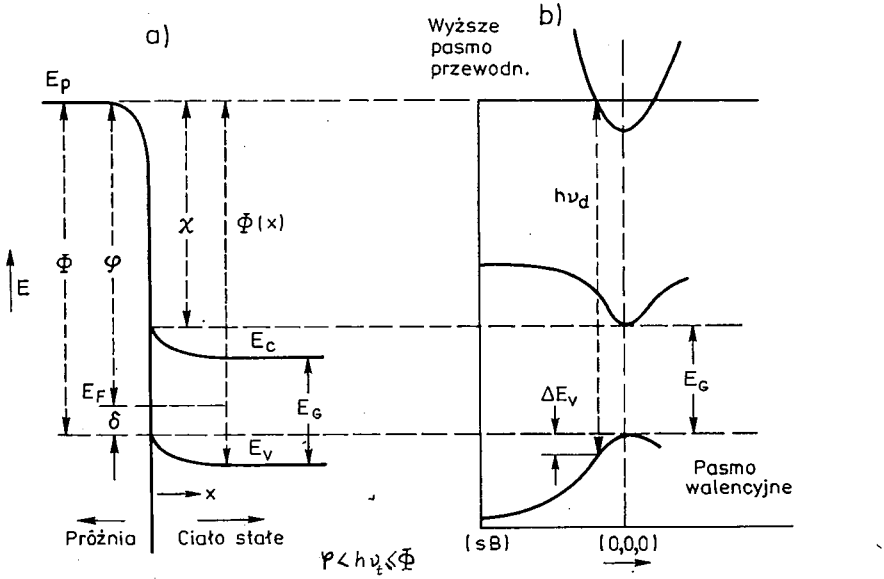
Jeżeli nie bierze się pod uwagę zapełnionych stanów powyżej szczytu pasma walencyjnego, lub jeśli one nie występują, można zapisać, że

$$h\nu_t = E_p - E_v = \Phi \tag{21}$$

W świetle obecnego stanu teoretycznej wiedzy o fotoemisji Φ powinno mieć nieco wyższą wartość, jeśli równość $h\nu_t$ i Φ naruszona jest przez zajęte stany powierzchniowe właśnie ponad szczytem pasma walencyjnego, tj. Φ powinno mieć wartość pośrednią spełniającą warunek:

$$\Phi_t \leqslant \Phi \leqslant \Phi_d \tag{22}$$

gdzie Φ_d (lub $h\nu_\alpha$) jest stałą górną granicą Φ, ponieważ odpowiada ona przejściom pochodzącym nieco poniżej maksimum pasma walencyjnego (rys. 8).



Rys. 8. Model zewnętrznego zjawiska fotoelektrycznego w połączeniu z pasmowym modelem półprzewodnika (sB — strefa Brillouina)

Innymi słowy można powiedzieć, że w każdym przypadku ekstrapolacja (liniowej) części krzywej wydajności dostarcza nam wartości krawędzi $h\nu_d$ lub (Φ_d), która jest górną granicą prawdziwej wartości Φ. Zauważyć można, że $h\nu_d$ będzie ściśle równa wielkości Φ w specjalnym przypadku, kiedy przejście zachodzi dla k=(000).

Powyższe argumenty pozwalają na ścisłe zdefiniowanie granic pomiędzy którymi zawiera się prawdziwa wartość Φ, a mianowicie:

$$h\nu_d \geqslant \Phi \geqslant h\nu_t \tag{23}$$

Gdy zarówno Φ, jak i praca wyjścia φ są znane, można określić energetyczny przedział między szczytem pasma walencyjnego i poziomem Fermiego na powierzchni; przedział ten wynosi:

$$\delta \equiv (E_F - E_p) S = \Phi - \varphi \tag{24}$$

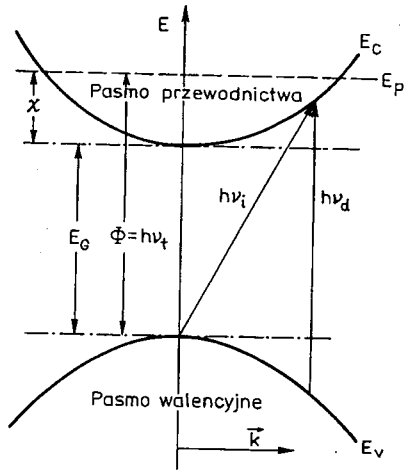
jest to charakterystyczna wielkość dla fotoemisji z półprzewodników.

Najwyżej zaznaczona na rys. 8 energia jest nazywana energią próżniową E_p. Poziom E_p odpowiada zeru energii kinetycznej swobodnego elektronu, lecz poza tym jest traktowany jako jeden z poziomów w pasmach przewodnictwa. Zatem próg fotoelektryczny, jak już wiemy, jest to minimalna energia fotonu $h\nu_t$ potrzebna do „podniesienia" elektronu w ciele stałym do poziomu próżni E_p, gdzie energetycznie możliwa jest ucieczka do próżni.

Wprowadzona już wielkość δ nazwana jest parametrem fotoelektrycznym i, jak mówiliśmy już, określa energetyczną odległość poziomu Fermiego od najwyżej położonych emitujących poziomów identyfikowanych ze szczytem pasma walencyjnego. Wartość δ może posłużyć do wyznaczania szerokości przerwy energetycznej E_G, jeśli znane jest położenie poziomu Fermiego E_F względem dna pasma przewodnictwa [9]

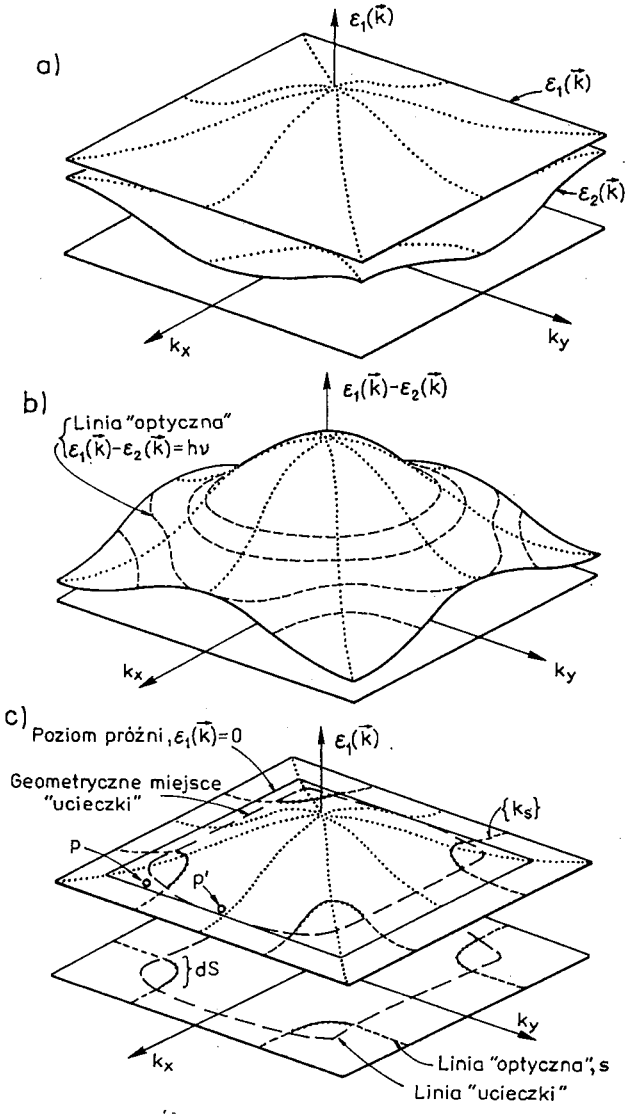
$$E_G = |E_p| + \delta \tag{25}$$

Aby przybliżyć dwuwymiarowy model pasmowy wzbudzenia fotoemisji przez przejścia proste i uzmysłowić sobie „poszerzanie się" trójwymiarowego pasma podajemy rys. 9 i rys. 10 [12–14].

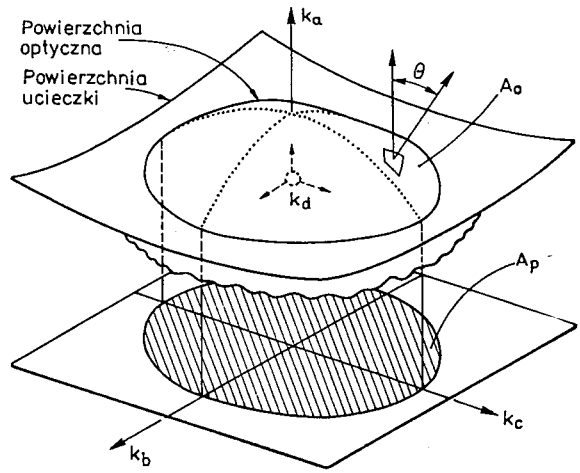


Rys. 9. Proste ($h\nu_i$) i skośne $\bar{h}\nu_i$ przejścia międzypasmowe w półprzewodniku

Na rys. 11 powierzchnie: ucieczki i optyczna są narysowane powyżej progu dla trójwymiarowych pasm. Środek układu k_d odpowiada progowej energii fotoefektu. Te powierzchnie są teraz „rozprężane" (według szeregów Taylora) powyżej punktu progowego k_d, który jest zdefiniowany jako ten punkt w przestrzeni k, gdzie pierwsza powierzchnia optyczna staje się styczna do powierzchni ucieczki, kiedy $h\nu$ jest zwiększone. Energia fotonu korespondująca do tej sytuacji jest $h\nu_d$ rys 9, która jest



Rys. 10. Model dwuwymiarowego pasma dla emisji fotoelektrycznej: a) Hipotetyczna energia pasm dla dwu-wymiarowego kryształu. Pasmo „przewodnictwa" jest oznaczone przez $\delta_1(\vec{k})$ i jest przypuszczalnie puste. Pasmo „walencyjne" to $\delta_2(\vec{k})$ i przypuszczalnie jest pełne. b) Struktura łączących się pasm $\delta_1(\vec{k})-\delta_2(\vec{k})$ odpowiada pasmom na rys. 1a). Pokazano kontury stałej różnicy energii („optyczne" linie); są to kontury, które dają miejsca dla „pionowych" przejść. c) Położenie końcowych stanów $\{k_s\}$ fotopobudzonych elektronów ilustrujące warunki ucieczki. Elektrony w końcowych stanach pokazane przez „paciorkowe" (kropki – kreski) linie leżą powyżej miejsc ucieczki i będą emitowane w nieobecności rozpraszania, jeśli będą miały grupową prędkość skierowaną do powierzchni. Miejsce ucieczki jest narysowane dla powierzchni prostopadłej wzdłuż $(\vec{k_x})$



Rys. 11. Przecięcie powierzchni ucieczki i optycznej dla pasm trójwymiarowych. A_0 jest polem optycznej powierzchni położonym ponad powierzchnią ucieczki. A_p jest rzutem A_0 na płaszczyznę k_b, k_c

fotoelektrycznym progiem dla przejścia prostego. Kształt „rozprężenia" wyrażają (obrazują) równania Kane'a:

$$E_c(\vec{k}) - E_v(\vec{k}) \approx h\nu_d + \alpha k_a + \mathrm{quad}(k_b,k_c)$$
$$E_c(\vec{k}) - \frac{\hbar^2 k_t^2}{2m} \approx \alpha_1 k_a + \mathrm{quad}_1(k_b,k_c) \tag{26}$$

gdzie: $E_c(\vec{k}) \geqslant \frac{\hbar^2 k_t^2}{2m}$ jest energią pasma przewodnictwa, a
$E_v(\vec{k})$ energią pasma walencyjnego;
k_t — składowa styczna do powierzchni;
k_a — składowa protopadła do powierzchni;
$\mathrm{quad}(k_b,k_c)$ i $\mathrm{quad}_1(k_b,k_c)$ są kwadratowymi funkcjami k_b i k_c [2]
a

$$\alpha = |\vec{\nabla}_k [E_c(\vec{k}) - E_v(k)]|k_d$$

i

$$\alpha_1 = \left|\vec{\nabla}_k \left[E_c(\vec{k}) - \frac{\hbar^2 k_t^2}{2m}\right]\right|_{k_d}.$$

4. TEORIA FOTOEMISJI BALLANTYNE'A

Nowszą teorią, opierającą się częściowo na teorii Kane'a, jest teoria Ballantyne'a. Ballantyne w pracy [15] p.t. *„Zjawisko strat energii do fotonów"* szczegółowo omówił swoją obszerną teorię fotoemisji. Powszechna metoda określenia progów fotoelek-

trycznych polega, jak wiemy, na przystosowaniu prawa potęgowego $Y = C(h\nu - h\nu_t)^n$ do mierzonej wydajności i znalezieniu energii dla której ekstrapolowana funkcja przecina oś energii dla zerowej wydajności Y. Kane rozwinął teeoretyczny model pasmowy fotoemisji, który przewidział różne prawa potęgowe z potęgami 1 do $\frac{5}{2}$ zależnymi od procesów fizycznych operujących blisko progu. Bardzo blisko progu fotoelektrycznego, powszechnym, eksperymentalnie obserwowanym prawem potęgowym w obu zewnętrznej i wewnętrznej strukturze MOS fotoemisyjnych wydajnościach jest prawo sześcienne. Zazwyczaj to prawo przystosowuje się do wydajności w przedziale 0,1 – 0,3 eV; wyżej-energetyczne części wydajności były czasem przystosowywane z różnymi prawami potęgowymi. Według Ballantyne'a z powodu rozmaitości przystosowań, które mogą być wykonane do poszczególnych odcinków danych wydajności i, z powodu ograniczonego zakresu prawdziwości każdego, wydaje się nieco wątpliwe przyjmowanie ogólnej formuły Kane za prawdziwą, a więc aktualna praktyka zawiera próby przystosowania wszystkich praw potęgowych, używając tego które daje najlepsze przystosowanie.

Ballantyne rozwija model emisji fotoelektrycznej oparty na pracy Kane, lecz włączający straty energetyczne związane z rozproszeniem na fononach. Ten model daje dobre przystosowania do wydajności Y i do krzywych EDC dla różnorodności materiałów w zakresach energii fotonów rzędu 1 eV i w rezultacie usuwa wiele niejasności związanych poprzednio z przystosowaniami potęgowymi. Model można stosować blisko progu, jeśli produkcja par lub rozproszenie elektron-elektron jest małe i jeśli głębokość absorpcji optycznej jest porównywalna lub większa od drogi rozproszenia fononowego gorących elektronów. Taki układ był poprzednio badany tylko metodami numerycznymi [16]. Przed Ballantyne nie podejmowano prób połączenia teorii pasmowej z uproszczonym modelem rozproszeń na gruncie teorii analitycznej, aby otrzymać przybliżone wyrażenia na krzywe EDC i wydajności Y w przypadku gdy dominuje rozproszenie fononowe.

Ballantyne rozważa proces wzbudzenia fotoemisji na drodze doświadczalnej jako zawierający trzy etapy:

a) fotowzbudzenie w ramach teorii pasmowej pojedyńczego elektronu,
b) transport fotowzbudzonych elektronów do powierzchni,
c) ucieczka elektronów do próżni.

Dalej Ballantyne rozkłada drugi etap i rozważa oddzielenie zjawiska zmiany rozkładu w przestrzeni k i strat energii, chociaż oba występują jednocześnie. W całej pracy pojęcie „rozproszenie fononowe" jest użyte ogólnie do opisu wszystkich rozproszeń elastycznych i quasielastycznych włączając rozproszenia na zanieczyszczeniach i defektach oraz rozproszenia cieplne w materiałach amorficznych. Ballantyne rozważa sytuacje gdy poziom próżni leży powyżej minimum pasma przewodnictwa. Wpływ zagięć pasm na pobudzenie i transport elektronów i towarzyszące danemu zagięciu silne pola blisko powierzchni nie są włączone do jego modelu, choć mają one znaczny wpływ gdy są obecne [17] [18]. Rozważa jednak efekty, gdy etap ucieczki z powodu obniżenia bariery przez przyłożone napięcie ma wpływ na wydajność i rozkład EDC podyktowany i przewidziany przez model pasmowy.

Ballantyne pokazuje przybliżone podejście do zagadnienia strat energii, w rozproszeniach na fononach, pobudzonych elektronów przed ich emisją z kryształu. Przedmiotem jest tu rozwinięcie prostych analitycznych formuł dla rozkładów energetycznych i wydajności, które wprowadzają efekt strat energetycznych. Metody są fenomenologiczne, ale dają rezultaty które ściśle przybliżają do wyników obliczeń numerycznych.

Rozkład energetyczny fotowzbudnych elektronów z głębokości x poniżej powierzchni, określono przez $\beta(h\nu,E,x)$:

$$\beta(h\nu, E, x) = \alpha(h\nu)\, e^{-\alpha x} N(h\nu,E), \tag{27}$$

gdzie $\alpha(h\nu)$ jest współczynnikiem absorpcji optycznej, który zależy od głębokości, a $N(h\nu,E)$ jest energetycznym rozkładem elektronów pobudzonych przez światło monochromatyczne.

Zastosowano straty energii zależne od rozproszeń na fononach w najprostszy sposób, przyjmując jednowymiarowy model w którym wszystkie elektrony mające początkowo energię E i głębokość x osiągają powierzchnię ze stratami energii $\frac{xE_{fon}}{l_{fon}(E)}$; gdzie E_{fon} jest średnią energią strat na jedno rozproszenie fononowe, a l_{fon} jest średnią drogą swobodną na rozproszenie fononowe. Później odnosi się l_{fon} do średniej liczby zderzeń rozproszeniowych w taki sposób, jak to prawidłowo policzono dla strat energii. Przyczyną zastosowania raczej liniowej niż kwadratowej zależności strat energii od x jest fakt, że to daje lepsze przybliżenie do aktualnej (ścieżki losowej) macierzy dla krzemu.

Prawdziwą zależność uśrednionych strat energii podawaną w skali liniowej lub kwadratowej z długością, zależną od wartości parametrów ustalił Kane [19], a Ballantyne, jak zobaczymy dalej przedyskutował te rozważania bardziej szczegółowo.

Jeśli energia elektronów w warstwach przypowierzchniowych jest określona przez E_1, wtedy część ich $f(E,E_1,x)$ pobudzonych przy głębokości x do energii E, która osiąga powierzchnię z energią E_1 jest równa:

$$\varphi(E, E_1, x) = \delta\left(E - \left[\frac{xE_{fon}}{l_{fon}(E)}\right] - E_1\right) \quad \text{dla} \quad E_1 \leqslant E \tag{28}$$

Równanie 28 nie jest tak *„nieobrobionym"* przybliżeniem jak to się mogło wydawać. Jeśli elektrony znajdujące się na głębokości x są pobudzone do energii E i w rozkładzie funkcji δ pojawiają się przy powierzchni ze średnią energią $\bar{E}_1(x)$ w aktualnym rozkładzie $f'(\bar{E})$ — obróbka zastosowana w równaniu (28) wymienia $f'(E_1)$ przez funkcję δ przy $\bar{E}_1(x)$, gdzie $\bar{E}_1(x) = E - \frac{xE_{fon}}{l_{fon}(E)}$. To powinno być prawidłową procedurą tak głęboko jak sięgają straty energii, jeśli liczba przypadków rozproszeń jest duża $\left[\frac{1}{\alpha(h\nu)} \gg l_{fon}(E)\right]$.

Duckett [20] wykonał dokładne badanie klasycznego, trójwymiarowego, (losowej ścieżki) problemu rozproszeń fononowych dla fotoemisji i wykazał, że te wyniki były bardzo dokładnie przybliżone przez jednowymiarowy model pokazany przez Kana [19], który został ostatnio skontrolowany (sprawdzony) na gruncie teoretycznym [18].

Przez staranną interpretację parametrów l_{fon} i E_{fon} model Ballantyne'a może być równoważny modelowi Kane'a, w tym stopniu w jakim dotyczy średniej energii strat. Kane dał uśrednioną energię strat dla elektronów pobudzonych przez światło o częstotliwości ν jako $\bar{n}E_{\text{fon}}$, gdzie E_{fon} jest średnią energią strat na jedno *„zdarzenie zderzenia"* uśrednioną zarówno na absorpcję jak i emisję fononu [15].

$$E_{\text{fon}} = <\hbar w_{\text{fon}} : [n_{\text{fon}}(\omega) + 1]> \tag{29}$$

gdzie $\hbar\omega_{\text{fon}}$ jest energią fononu, $n_{\text{fon}}(\omega)$ jest liczbą pobudzonych fononów i uśrednioną na wszystkie rodzaje fononów. Średnia liczba zderzeń z fononami dla elektronów pobudzonych przez światło o częstotliwości ν jest dana przez $\bar{n}$ gdzie w granicy, a » b. [14]

$$\bar{n} \simeq \frac{1}{2}\left(-\frac{a}{b}\right)^{1/2} \frac{1 + P}{(1-P) + (1+P)(b/a)^{1/2}} + \frac{a}{\alpha(h\nu) + (ab)^{1/2}} \tag{30}$$

gdzie $a = \sqrt{3}\,(<l_{3\text{fon}}(E)>)^{-1}$;

$b = \sqrt{3}\,(<l$ trójwymiarowa dla par $(E)>)^{-1}$

a P jest prawdopodobieństwem, że elektron jest zawrócony do kryształu po osiągnięciu powierzchni.

Uśrednienia dla prawdziwej trójwymiarowej długości na rozproszenie fononowe $l_{3\text{fon}}(E)$ i drogi przy kreacji par $l_{3p}(E)$ są wzięte w całym zakresie energii E pobudzonych fotoelektronów. Współczynnik $\sqrt{3}$ jest konieczny do prawidłowej relacji między jedno i trójwymiarowymi modelami. Model Ballantyne wprowadza dwie właściwości nie uwzględnione w pracy Kane [9]. On uwzględnia informacje o tych elektronach które poprzez rozproszenie fononowe tracą część energii i spadają poniżej poziomu próżni; *„włącza"* on również w wartość średniej drogi zależność długości rozproszenia fononowego od energii elektronów.

Porównanie równań (30) i (28) pokazuje, że

$$<\frac{x}{l_{\text{fon}}(E)}> = \bar{n} \tag{31}$$

gdzie uśrednienie lewej strony tego równania jest wzięte zarówno na x jak i na E. Przyjęto, że przedziały x i E w których uśredniamy zależą od hν. Zakres energii E jest dany przez równanie rozkładu wewnętrznie wzbudzonych elektronów:

$N(h\nu, E)\,dE = C(h\nu)\,dE$, $\quad 0 \leqslant E \leqslant E_{\max}$, gdzie C ma dość złożoną postać, a $<x>$

jest właśnie równe głębokości absorpcji $\alpha^{-1}(h\nu)$. Stąd explicite czyni się tę zależność przez założenie:

$$l_{\text{fon}}(\text{h}\nu) = \langle l_{\text{fon}}(E)\rangle \tag{32}$$

skąd:
$$l_{\text{fon}}(\text{h}\nu) = [\alpha(\text{h}\nu)\bar{n}]^{-1}. \tag{33}$$

Przez użycie równań (30) (32) (33) prosty, jednowymiarowy model Ballantyne'a może być odniesiony do rygorystycznego, trójwymiarowego modelu Ducketta.

Liczba $M(\text{h}\nu,E,E_1)$ elektronów pobudzonych w objętości do energii E, które osiągają powierzchnię z energią E_1, wynosi:

$$M(\text{h}\nu, E, E_1) = \int_0^\infty f(E, E_1, \text{x})\,\beta(\text{h}\nu, E, \text{x})\,\text{dx} =$$

$$= wN(\text{h}\nu, E)\,\text{e}^{-w(E-E_1)}; \quad E_1 \leqslant E \tag{34}$$

oraz
$$M(\text{h}\nu, E, E_1) = 0 \quad \text{dla} \quad E_1 > E$$

gdzie
$$w \equiv \frac{1}{E_{\text{fon}}}\,\alpha(\text{h}\nu)\,l_{\text{fon}}(E).$$

[Przyjmuje się, że fotoemiter jest grubszy niż $\alpha^{-1}(\text{h}\nu)$]

Grupa elektronów $Q(\text{h}\nu,E,E')$ początkowo pobudzonych do E, które uciekają przy uzyskaniu energii E' równej E_1 jest macierzą ucieczki:

$$Q(\text{h}\nu, E, E') = \frac{P'(E_1)\,M(\text{h}\nu, E, E_1)}{N(\text{h}\nu, E)}, \tag{35}$$

gdzie $P'(E_1)$ jest prawdopodobieństwem że elektron o energii E_1 poniżej powierzchni może uciec i jest dane przez równanie

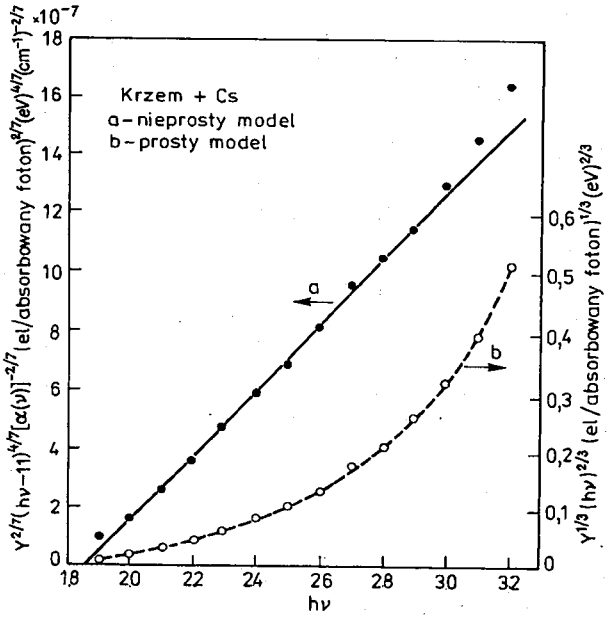
$$P'(E) = A(E)\,E, \quad \text{gdzie} \quad A(E) = \pi 2m\langle\alpha_1^{-1}\rangle:$$

$$\hbar \int_{SB} \delta[\mathscr{E}_1(\vec{k}) - E]\,\text{dk};$$

$\mathscr{E}_1$ — stała energia powierzchni.

Ballantyne dyskutując przejścia proste w objętości podał dość złożone wzory dla wydajności z przejściami prostymi wprowadzając efekty rozproszeń fononowych.

Średnia liczba przypadków rozproszeń przed emisją jest dana przez $\bar{n}$ z równania 30 i zależy od: odbicia elektronów od powierzchni, względnego prawdopodobieństwa produkcji par i zderzeń z fononami. Na przykład gorący elektron o energii 5 eV w krzemie doznaje rozproszeń fononowych około osiem razy częściej niż produkcji par. Jeśli nawet głębokość absorpcji optycznej w krzemie przy $\text{h}\nu = 5$ eV wynosi tylko 60 Å, co jest równe długości jednego rozproszenia fononowego $l_{3\text{fon}}$, równanie 30 pokazuje, że (uśredniony) przeciętny elektron doznaje 2,6 przypadków rozproszeń fononowych przed emisją, jeśli odbicie od powierzchni jest równe zero. Jeśli współczynnik odbicia $R > 0$ średnia liczba zderzeń fononowych rośnie. Te wyniki otrzymano przez obliczenia metodą Monte-Carlo które wykazały, że duża liczba



Rys. 12. Wydajność fotoelektryczna łupanego w próżni monokryształu Si o powierzchni cezowanej. Punkty doswiadczalne są takie same na wykresach a i b. Rozkład energetyczny dla przejść nieprostych jest dany przez ciągłą linię wykresu a. Rozkład teoretyczny dla przejść prostych jest dany przez prostą pokrywającą się z rzędną wykresu b

rozproszeń fononowych pojawia się nawet gdy $\frac{1}{a(\mathrm{h}\nu)} \simeq l_{3\mathrm{fon}}$, jeśli *„stożek ucieszki"* jest mały (co jest blisko progu). Uśrednione 2,6 rozproszeń (zderzeń) fononowych powinno wpływać na znaczną zmianę rozkładu pobudzonych elektronów w przestrzeni k.

Znane nam już prawo sześcienne, obserwowane dla wydajności blisko progu, wyjaśnił Ballantyne na podstawie objętościowych przejść prostych ze stratami energii zależnymi od rozproszeń fononowych.

Na rys. 12 pokazano porównanie między teorią i eksperymentem dla przejść skośnych w objętości kryształu. Rozkład wydajności dany jest dla łupanego w wysokiej próżni i pokrywanego cezem krzemu w zakresie energii fotonów między 1,9 a 3,0 eV; w tym przedziale przejścia są skośne, a wydajność jest dobrze dopasowana jak pokazano na rys. 12. Te same dane nie mogą być dopasowane do przejść prostych o czym świadczy duże odchylenie krzywej b od linii prostej a na tym rysunku. Model przejść skośnych dobrze dopasowuje się do wydajności w zakresie 1 eV. Przy wyższych energiach dopasowanego zakresu, krzywa doświadczalna Y(hν) rośnie powyżej niż przewiduje odpowiednie równanie, ponieważ fotoelektrony pochodzące z przejść prostych zaczynają uczestniczyć w fotoemisji [21].

Dane Edena [22] i obliczenia pasmowe wykonane przez niego wskazują, że próg przejść prostych dla krzemu znajduje się w pobliżu 3 eV. Przy dolnym końcu

dopasowanego zakresu $h\nu < 2$ eV w Si, wydajność eksperymentalna jest znowu większa niż była przewidziana, co jest spowodowane obecnością stanów powierzchniowych lub problemami eksperymentalnymi, takimi jak: niejednolita praca wyjścia emitera, fotoemisja z innych części aparatury doświadczalnej związana z rozproszeniem światła i gromadzenie elektronów powodowane przez przyrządy pomiarowe i pompy jonowe. Ta emisja ze *„stanów powierzchniowych"* jest problemem zupełnie ogólnym, dotyczy bowiem wydajności poniżej 10^{-5} lub 10^{-6} el./foton i jest obserwowana dla bardzo wielu materiałów (w tym związków grupy $A^{III}B^{V}$). Wyprowadzone przez Ballantyne'a równania [14]:

$$Y_{pr} = \frac{B\alpha(h\nu)}{(h\nu)^2 \mathscr{E}(h\nu)} \left(E_{max}^3 - \frac{1}{4} \bar{g} E_{max}^4 + \ldots \right) \tag{36}$$

i

$$Y_{skoś} = \frac{E\alpha(h\nu)(h\nu - h\nu_i)^{7/2}}{(h\nu - \Delta\mathscr{E})^2} \left[1 - \frac{4}{9} \bar{g}(h\nu - h\nu_i) + \ldots \right] \tag{37}$$

gdzie $\mathscr{E}$ jest energią kinetyczną wyemitowanych elektronów w próżnię; B jest stałą; $\bar{g} = \alpha(h\nu) l_{fon}(h\nu)/E_{fon}$ dają dobre dopasowania w środkowym i wyższym zakresie wydajności i pozwalają rozróżnić z większą wiarygodnością między wydajnością objętościową i wydajnościową zależną od powierzchni lub innych procesów. Zakres 1 eV dopasowany do równania Ballantyne jest także zgodny z przedziałem działania zjawiska *„rozszerzania się"* pasm energetycznych w Si przy poziomie próżni 1,9 eV. Kontrola kształtu pasm [23] wykazuje, że rozszerzenia (rozwinięcia) takie jak w równaniach (38)

$$\begin{aligned} h\nu &= E_1(\vec{k}) - E_2(\vec{k}) \approx h\nu_d + \alpha k_a + \mathrm{kwad}(k_b, k_c), \\ E_1(\vec{k}) - h^2 k_t^2 / 2m &\approx \alpha_1 k_a + \mathrm{kwad}_1(k_b, k_c), \end{aligned} \tag{38}$$

powinny być prawdziwe dla obszaru około 1 eV poniżej wierzchołka wyższych pasm walencyjnych i powyżej poziomu 1,9 eV (poziom próżni) drugiego – najniższego pasma przewodnictwa. W przypadku tego drugiego pasma *„rozwinięcie"* powinno być prawdziwe dla conajmniej 0,6 eV powyżej poziomu próżni (1,9 eV). Ostatnia sugestia oparta na fotomodulowanej wydajności [24], że rozszerzenia obszaru słuszności Kane'a są prawdziwe tylko do 0,1 eV powyżej progu dla czystego Si była dyskutowana przez Ballantyne'a; według niego, taki wniosek nie byłby sugerowany przez kształty (kontury) pasm dla wartości położenia progu przy 5,0 eV. Struktura obserwowana podczas eksperymentu modulacji była interpretowana na podstawie przejść do stanów powierzchniowych, które to przejścia ogólnie dają niskie wydajności Y, oraz są nazewnątrz zakresu teorii Ballantyne'a.

Ballantyne opracował model fotoemisji zależny od prostych, skośnych i nieprostych przejść objętościowych, a który jest oparty na przybliżeniu rozszerzenia (rozwinięcia) pasm wg Kane'a. Wyniki dla zależności wydajności od energii w róż-

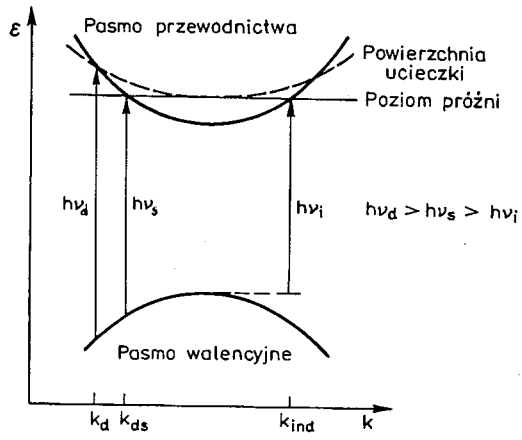
Tablica I

Zależność fotoelektronowej wydajności od energii fotonów w pobliżu progu dla różnych typów objętościowego pobudzenia

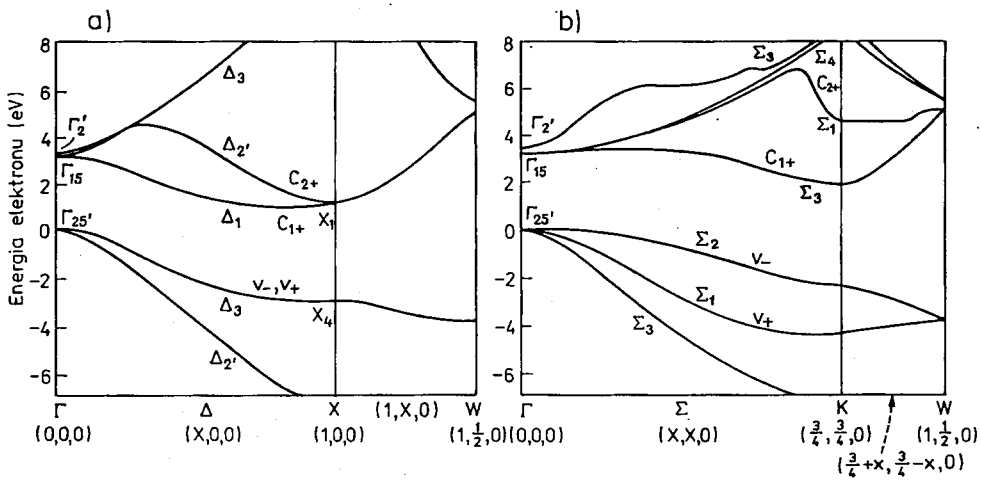
Rodzaj przejść	Zależność Y od $h\nu^+$	Uwagi
1(a) Przejście proste z quasielastycznym rozproszeniem i stratami energii, *„stany końcowe odstępu progowego"* nie wypadają przy ekstremum pasma przewodnictwa	$\frac{\alpha(h\nu)}{(h\nu)^2 E_2'(h\nu)}(h\nu - h\nu_s)^3$	
1(b) Nieproste przejścia, które wzbudzają prostokątny rozkład, ze stratami energii quasielastycznego rozproszenia	$\frac{(h\nu)}{(h\nu)^2 E_2'(h\nu)}(h\nu - h\nu_i)^3$	
1(c) Większość półprzewodników	taka sama jak 1(a) lub 1(b)	
2(a) Przejścia proste z quasielastycznym rozproszeniem i stratami energii. *„stany końcowe odstępu progowego"* nie wypadają przy ekstremum pasma przewodnictwa	$\frac{1}{(h\nu)^2 E_2' h\nu}(h\nu - h\nu_s)^2$	
2(b) Przejścia nieproste które wzbudzają rozkład prostokątny bez strat energii quasielastycznego rozproszenia	$\frac{1}{(h\nu)^2 E_2' h\nu}(h\nu - h\nu_i)^2$	
2(c) Większość metali	taka sama jak 2(a) lub 2(b)	
3 Tylko przejścia skośne, z quasi-elastycznym rozproszeniem i stratami energii, stany *„końcowe odstępu progowego"* nie wypadają przy ekstremum pasma przewodnictwa	$\frac{\alpha(h\nu)}{(h\nu - \Delta\mathscr{E})^2}(h\nu - h\nu_i)^{7/2}$	Tam gdzie różnica między $h\nu_{ds}$ i $h\nu_i$ jest duża np. w Si $\Delta\mathscr{E}$ — przerwa pasmowa dla przejść skośnych
4(a) Przejścia proste, bez quasi-elastycznego rozproszenia	$\frac{1}{(h\nu)^2 E_2'(h\nu)}(h\nu - h\nu_d)$	

$^+h\nu_d$, $h\nu_s$, $h\nu_i$ — energie fotoelektrycznych progów dla różnych rodzajów przejść

nych przypadkach zostały podsumowane w Tablicy I. W tablicy tej uczyniono odnośnik do trzech różnych typów energii progowych które są zilustrowane schematem na rys. 13. $h\nu_d$ jest progiem związanym z przejściami prostymi bez istnienia rozproszeń; i jak pokazano dalej w połączeniu z rys. 10 i 14 ten próg będzie pojawiał się przy przejściach do stanów wyraźnie powyżej poziomu próżni. Z tego powodu występuje tam niższy próg $h\nu_s$ dla przejść prostych z rozproszeniem, który zazwyczaj będzie mierzony w praktyce. Stany końcowe od tego progu będą zazwyczaj leżały



Rys. 13. Wykres $\mathscr{E}$ (k) ilustrujący fotoemisyjne energie progowe dla skośnych lub nieprostych przejść ($h\nu_i$), przejść prostych ($h\nu_d$), oraz prostych + rozproszenia elastyczne ($h\nu_s$)



Rys. 14. Struktura pasmowa Si według Kane'a [22] ilustruje regiony ważności „rozszerzań" pasma z powodu przejść do drugiego najniższego pasma przewodnictwa C_{2+}. Na tym rysunku zero energii jest szczytem pasma walencyjnego w przeciwieństwie do reszty opracowania Ballantyne'a, gdzie zero energii jest poziomem próżni. Wykresy (13a) i (13b) uwidaczniają strukturę wzdłuż kierunków symetrii. Dyskutowane poziomy próżni dla 1,8; 4,6 i 4,9 eV są pokazane linią kreskową (przerywaną). Punkty progowe k_d, k_{ds} i k_{niepr} dla tych poziomów próżni są wykazane przez skrzyżowanie

przy wartości punktu wektora k_{ds}, którego energia znajduje się przy poziomie próżni, aczkolwiek jest możliwym dla k_{ds} osiągnąć końcowe stany znacznie powyżej poziomu próżni co pokazano dla Si w połączeniu z rys. 14. Ponieważ natężenia intensywności przejść przy wartościach wektora k_{ds} i k_d na rys. 13 są porównywalne ze sobą, $h\nu_d$ nie może prawdopodobnie być obserwowane w obecności $h\nu_s$. W pewnych specjalnych przypadkach, jak wynika w połączeniu z rys. 14, wektory k_d i k_{ds}, a stąd $h\nu_d$ i $h\nu_s$ są identyczne.

Próg $h\nu_i$ dla skośnych i nieprostych przejść zawsze pojawia się w punkcie $k_{skoś}$ któremu odpowiada energia poziomu próżni; jednakże $h\nu_i$ może być mniejsze niż $h\nu_s$, jeśli k_{ds} odpowiada punktowi poniżej wierzchołka pasma walencyjnego lub punktowi powyżej poziomu próżni.

Za wyjątkiem nietypowych przypadków, (na przykład Si gdzie różnica między $h\nu_s$ i $h\nu_i$ jest bardzo duża), prawdopodobnie będzie trudno zaobserwować $h\nu_i$ ($h\nu_s$ – to przejścia proste z elastycznymi zderzeniami) (rys. 13).

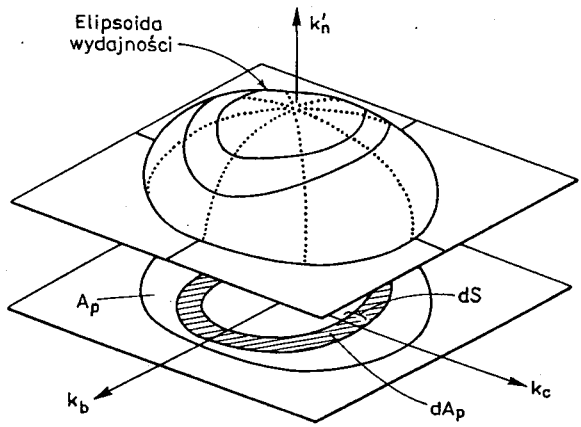
Dla dużej ilości materiałów (a w szczególności dla wysoko położonych poziomów próżni) struktura pasmowa jest dostatecznie skomplikowana, tak że różnica $h\nu_i - h\nu_s$ jest mała i większe natężenie przejść prostych będzie szyko *„zaciemniać"* zależność wydajności od $h\nu_i$. Wydaje się, że taki przypadek jest spełniony dla wszystkich eksperymentalnych danych które były analizowane (za wyjątkiem Si).

Gdy włączy się wpływy (na fotoemisję) strat energetycznych związanych z rozproszeniem na fononach, model przejść prostych przewiduje powszechnie obserwowane prawo trzeciej potęgi dla wydajności blisko progu. Jeśli uwzględni się zależność energetyczną współczynników (Tablica I) w wyrażeniu na wydajność, dobre dopasowania funkcji wydajności mogą być otrzymane w dużym przedziale dając znaczną wiarygodność tak określonych (otrzymanych) energii progowych. Dobre dopasowania są ogólnie otrzymywane przy wydajnościach większych niż 10^{-4} el./fot. wskazujących na procesy objętościowe. Wydajności poniżej 10^{-4} el./fot., najogólniej biorąc, rosną powyżej wartości przewidzianych dla pojedyńczych (wyodrębnionych) progów objętościowych. Ten wzrost może być związany z pewną liczbą procesów, takich jak emisja ze stanów powierzchniowych oraz ze stanów w przerwie energetycznej, niejednorodność prac wyjścia emiterów, fotoemisja z innych części aparatury związana ze światłem rozproszonym o gromadzenie elektronów generowanych przez inne części aparatury takie jak głowice próżniomierzy i pompy jonowe. Podane przez Ballantyne'a techniki identyfikacji i dopasowania wydajności mogą pomóc w odróżnieniu tych efektów od procesów objętościowych.

Gdy współczynniki zależne od energii fotonów są pomijalne, procedura przy dopasowaniu danych staje się bardziej niejasna z powodu tego, że kilka różnych zależności potęgowych można przystosować do niewielkiej porcji danych, a energia progowa tak określona jest mniej dokładna, ponieważ przystosowanie wykonuje się do nielicznych punktów eksperymentalnych. Wyniki Ballantyne'a pokazują, że przybliżenie pasmowego rozszerzenia-rozwinięcia może być prawdziwe dla $h\nu$ rzędu elektronowoltów lub więcej. Jeśli wprowadzimy rozróżnienia w sposobie oceny procesów powierzchniowych i objętościowych, to zazwyczaj byłoby możliwe osiągnąć dobrą zgodność doświadczenia z wyrażeniami Ballantyne'a poprzez kilka rzędów wzrostu wydajności. Jednakże są tam pewne efekty, które nie są włączone w jego model i które mogłyby spowodować znaczne odchylenie. Dwoma efektami, które są

napewno ważne, są: zagięcie pasm (zostało pokazane, że wpływa ono na eksponencjalne funkcje wydajności), oraz przypadek, gdy dno (lub silne minimum) pasma przewodnictwa jest powyżej poziomu próżni.

Fakt, że Ballantyne osiągnął doskonałe przystosowanie dla różnych materiałów nadaje nową wiarygodność przybliżeniu rozszerzenia pasmowego, gdy jest użyte ze środkami ostrożności. Stąd mogłoby się przydać do badania rozkładu kątowego fotoelektronów w przypadkach gdy rozproszenia są małe. Taki pomiar elipsoid emisji (rys. 15) dałby bezpośrednią informację o krzywiznach pasm.



Rys. 15. Elipsoida wydajności dla trój-wymiarowych pasm, ilustrująca konstrukcję obliczonych krzywych energetycznego rozkładu (EDC). Eliptyczny obszar dA_p jest rzutem na płaszczyznę $k_b k_c$ elementu elipsoidy wydajności wyciętego przez dwie koncentryczne kule które są powierzchniami stałych energii elektronów w próżni. Przyjęto, że normalna do powierzchni $\hat{k}'_n$ leży wzdłuż $\hat{k}_a$

Pewne nowe eksperymenty przeprowadzone przez Wootena i współpracowników [25] już poprzednio wykazały anizotropową emisję elektronów. Chociaż wyniki Ballantyne'a były wyprowadzone przy użyciu modeli objętościowych przejść prostych i skośnych, taka sama analiza rozproszeń może być użyta dla przejść nieprostych. W tych ostatnich przypadkach, gdy początkowy rozkład pobudzonych elektronów wewnątrz materiału może być z uzasadnieniem przybliżony przez rozkład prostokątny, przebieg Y i EDC jest podyktowany przez model przejść prostych.

Związek modelu Ballantyne'a z fononowymi stratami energii dostarcza nowej metody wyznaczania z danych wydajności parametrów rozproszenia gorących elektronów. Zastosowanie równań Ballantyne'a dotyczących EDC do danych eksperymentalnych będzie stanowić doskonałą technikę badań rozproszeń na fononach, jeśli wpływ eksperymentalnej rozdzielczości będzie mógł być usunięty [26].

Zestawienie wzorów z teorii Kane'a i Ballantyne'a

Przejścia objętościowe według Kane'a:

przejścia proste	$Y = k(h\nu - h\nu_t)^1$
przejścia skośne	$Y = k(h\nu - h\nu_t)^{5/2}$
przejścia proste z rozproszeniami quasielastycznymi	$Y = k(h\nu - h\nu_t)$

Przejścia objętościowe według Ballantyne'a:

przejścia proste z rozproszeniami quasielastycznymi i stratami energii

$$Y \sim \frac{\alpha(h\nu)}{h\nu^2 E_2(h\nu)}(h\nu - h\nu_d)^3$$

przejścia nieproste, rozkład EDC jest prostokątny, są straty energii przy quasielastycznych rozproszeniach

$$Y \sim \frac{\alpha(h\nu)}{(h\nu)^2 E_2(h\nu)}(h\nu - h\nu_i)^3$$

(Żeby porównać tę wydajność z (wydajnością) doświadczalną należy wyznaczyć k w odpowiednich jednostkach, w przedziale gdzie zaczynają dominować przejścia proste. k wyznaczamy z porównania Y_{teor} i $Y_{dośw}$)

przejścia skośne

$$Y \sim C\frac{\alpha(h\nu)}{h\nu}(h\nu - h\nu_i)^{7/2}$$

BIBLIOGRAFIA

1. W. E. Spicer: Phys. Rev. 112, 114 (1958); J. Appl. Phys. 31, 2077 (1960)
2. L. Apker, E. A. Taft, J. E. Dickey: Phys. Rev. 74, 1462 (1948)
3. G. W. Gobeli, F. G. Allen: Phys. Rev. 127, 141 (1962); 137, A245 (1965)
4. F. Bassani i inni: *Electronic States and Optical Transitions in Solids*, Oxford 1975
5. C. B. Berglund, W. E. Spicer: Phys. Rev. 136, A1030 (1964)
6. J. R. Chelikowsky, K. L. Cohen: Phys. Rev. B14, 556 (1976)
7. T. Grandke, L. Ley, M. Cardona: Solid State Commun., 23, 897 (1977)
8. D. Redfield: Phys. Rev., 129, 1809 (1961)
9. E. O. Kane: Phys. Rev. 127, 131 (1962)
10. E. O. Kane: Phys. Rev. 159, 624 (1967)
11. T. E. Fischer: *Surface Science*, 13, 30 (1969)
12. T. E. Fischer: Crit. Rev. Solid State Sci., 6, 401 (1976)

13. W. E. Spicer: J. Vac. Sci. Technol. 14, 372 (1977)
14. S. F. Lin, W. E. Spicer, R. S. Bauer: Phys. Rev. B14, 4551 (1976)
15. J. Ballantyne: Phys. Rev. B6, 1436 (1972)
16. R. N. Stuart, F. Wooten: Phys. Rev. 156, 364 (1967)
17. N. B. Kindig: J. Appl. Phys. 38, 3285 (1967)
18. B. F. Williams, R. E. Simon: Phys. Rev. Letters, 18, 485 (1967)
19. E. O, Kane: Phys. Rev. 147, 335 (1966)
20. S. W. Duckett: Phys. Rev. 166, 302 (1968)
21. D. C. Langreth: Phys. Rev. B3, 3120 (1971)
22. R. C. Eden: *Ph. D. thesis*, Stanford University, 1967; *Proceedings X International Conf. of Physics Semiconductors*, Cambridge, Mass., s. 221 (1970)
23. E. O. Kane: Phys. Rev. 175, 1039 (1968); 146, 558 (1966)
24. J. Rowe: *Procedings X International Conf. Physics Semiconductors*, Cambridge, Mass., s. 217 (1970)
25. F. Wooten, W. M. Breen, R. N. Stuart: Phys. Rev., 165, 703 (1968); F. Wooten, R. N. Stuart: Phys. Rev., 186, 592 (1969)
26. J. Kunz: *Topics Aplied Physics, Photoemission in Solid* II, 27, 299 (1979)

J. WOJAS, W. L. WOJAS

DISCUSSION OF THE PHOTOEMISSION THEORIES

Summary

This work consists the detailed discussion of succeding three theories of external photoelectric effects in semiconductors. Spicer's theory created photoemission model of three stages which will be analyse. Kane grounding on general Spicer's theory bound together photoemission with straight and through electron transitions. There is discussion of spectral dependence $Y(h\nu)$ on different conditions of electron excitation and escape. Author senior, investigating experimentally dependence quantum yield on photon energy confirmed rightness of Kane's theory. The latest Ballantyne's theory showed photoemission model grounding on results of Kane's work but taking into consideration losses of energy bound with phonon scattering. It gives possibility of calculation of parameters of hot electrons scattering on the ground of photoemission yield.

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI, 1992, 38, z. 4, ss. 583-615

681.32

Analiza własności statystycznych wielowartościowych sekwencji pseudolosowych generowanych przez k równolegle pracujące rejestry liniowe

ANTONI ZABŁUDOWSKI, BOŻYDAR DUBALSKI

Instytut Telekomunikacji, Akademia Techniczno-Rolnicza, Bydgoszcz

Otrzymano 1992.01.09

Autoryzowano do druku 1992.04.27

W pracy przedstawiono analizę własności wielowartościowych sekwencji pseudolosowych (MPS) generowanych nad ciałem GF(2^k). Sekwencje wielowartościowe są tworzone z k sekwencji binarnych otrzymywanych z rejestrów liniowych pracujących równolegle, przy czym każdej z sekwencji jest przypisana waga 2^i, gdzie $i=0,1,\ldots,k-1$. Analiza dotyczy następujących parametrów sekwencji wielowartościowej:

- rozkładu liczb pojawiających się w sekwencji;
- funkcji autokorelacji;
- rozkładu serii liczb w całej sekwencji.

WSTĘP

W różnych dziedzinach techniki do rozwiązywania problemów, zarówno o charakterze teoretycznym, jak i praktycznym, konieczne jest korzystanie z generatorów wielowartościowych sekwencji losowych (multilevel random sequences — MRS). Dziedzina zastosowania losowych ciągów wielowartościowych jest bardzo rozległa i obejmuje takie obszary, jak rozwiązywanie zadań numerycznych metodą Monte Carlo, badania statystyczne, symulowanie zjawisk i procesów technicznych, stochastyczne maszyny cyfrowe, urządzenia testująco-kontrolne, modelowanie systemów transmisji informacji w obecności szumu, czy też zagadnienia związane z identyfikacją systemów liniowych.

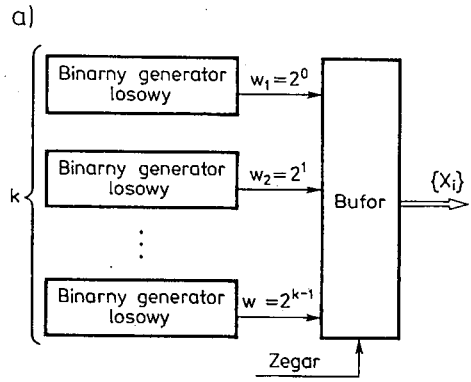
Jeszcze do niedawna dla otrzymywania sekwencji losowych stosowano generatory fizyczne, których zasada działania polegała na wykorzystaniu niektórych zjawisk fizycznych o charakterze stochastycznym. Obecnie te drogie, skomplikowane i trudne

do sterowania urządzenia, zostały wyparte przez zrealizowane sprzętowo lub programowo generatory wytwarzające deterministyczne sekwencje posiadające cechy ciągów losowych. Ciągi te noszą nazwę sekwencji pseudolosowych (multilevel pseudorandom sequences – MPS). W praktyce do realizacji sprzętowej generatorów wielowartościowych sekwencji pseudolosowych o rozkładzie równomiernym używa się, ze względu na prostotę układową, rejestrów liniowych działających nad ciałem Galois (p), gdzie p jest pewną liczbą pierwszą. Najczęściej wykorzystywane są w takim przypadku generatory wytwarzające sekwencje binarne [1, 3, 4, 6, 7, 9, 10, 11, 12, 13], co wynika z binarnego charakteru zasady działania układów stosowanych w większości urządzeń elektronicznych. Generatory binarne wykorzystywane są do konstrukcji generatorów sekwencji wielowartościowych utworzonych nad ciałem Galois (2^k) [9,11,13], chociaż w wielu publikacjach analizowane są również własności ciągów generowanych nad ciałem GF(p), gdzie $p>2$ [2, 4, 5, 8, 9, 12, 15, 16].

W niektórych pracach (np. [3]) pojawia się stwierdzenie, że najlepszym, w sensie probabilistycznym, źródłem MRS o rozkładzie równomiernym generującym sekwencje nad ciałem Galois (2^k), jest użycie k autonomicznych, niezależnie pracujących generatorów wytwarzających binarne (zero-jedynkowe) ciągi losowe. Ponieważ każdej binarnej sekwencji składowej przypisywana jest waga $w_i=2^i$, i$\in\{0,1,\ldots,k-1\}$, stąd każdemu wektorowi $<x_0,x_1,\ldots,x_{k-1}>$, pojawiającemu się na wyjściu tak zbudowanego generatora, jest przypisana liczba $\{X_i\}$ należąca do zbioru $\{0,1,\ldots,(2^k-1\}$.

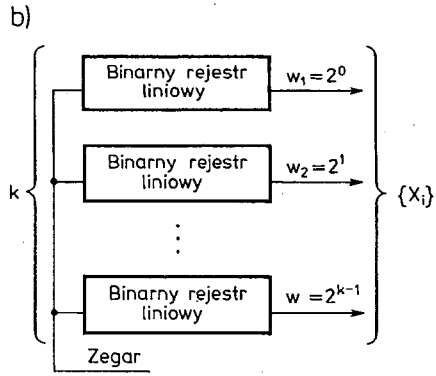
$$X_r = \sum_{i=0}^{k-1} x_{r_i} 2^i, \qquad x_{r_i} \in \{0,1\}. \tag{1.a}$$

Schemat blokowy takiego generatora pokazano na rys. 1.a.



Rys. 1a. Schemat blokowy generatora losowej sekwencji wielowartościowej

Często, dla uniknięcia wpływu wartości średniej na wielkość parametrów statystycznych generowanych ciągów, autorzy [3,6,9] w swych pracach przekształcają ciągi binarne zero-jedynkowe w homomorficzne sekwencje, których elementy należą do



Rys. 1b. Schemat blokowy generatora pseudolosowej sekwencji wielowartościowej

zbioru $\{-1,1\}$. Wówczas każdemu wygenerowanemu wektorowi sekwencji wielowartościowej przypisywana jest liczba $\{X_r'\}$, która należy do zbioru $\{-(2^k-1)\ (2^k-3),\ldots, (2^k-3),\ (2^k-1)\}$.

$$X_r' = \sum_{i=0}^{k-1} x_{r_i} 2^i, \qquad x_{r_i} \in \{-1,1\}\,. \tag{1.b}$$

Stwierdzenie, że generator pokazany na rys. 1.a posiada najlepsze własności losowe, stało się podstawą konstrukcji generatora wielowartościowych sekwencji pseudolosowych, którego schemat blokowy pokazano na rys. 1.b. Generator ten składa się z k rejestrów liniowych, czyli rejestrów przesuwających z odpowiednio dobranym sprzężeniem zwrotnym, będących źródłami pseudolosowych ciągów binarnych. Generator taki uznawany był dotychczas za najlepszy pod względem własności losowych, jednakże stosowanie tego typu generatorów, ze względu na zbytnią złożoność realizacyjną, nie uznawano za uzasadnione. Obecnie, gdy mogą być one zbudowane z elementów średniej i wielkiej skali integracji, rozwiązanie polegające na wykorzystaniu generatorów sekwencji binarnych pracujących równolegle staje się zarówno aktualne, jak i zasadne. Zatem niezbędne staje się również poznanie własności losowych sekwencji otrzymywanych z tak zrealizowanych generatorów.

W prezentowanej pracy dokonano oceny parametrów losowych wielowartościowych ciągów pseudolosowych otrzymywanych z sekwencji binarnych generowanych przez pracujące równolegle rejestry liniowe. Ocena generatorów pseudolosowych sekwencji wielowartościowych będzie polegać na określeniu na ile własności losowe generowanych sekwencji odbiegają od warunków losowości, które zdefiniowane zostały przez Golomba i Daviesa [3,6] dla ciągów binarnych, a które rozszerzono w niniejszej pracy na sekwencje wielowartościowe.

R1. Generowany ciąg zachowuje równomierność rozkładu liczb należących do zbioru $\{0,1,\ldots,(2^k-1)\}$ lub do zbioru $\{-(2^k-1),\ -(2^k-3),\ldots,-1,1,\ldots,(2^k-3),\ (2^k-1)\}$. tzn. każda z liczb pojawia się z jednakowym prawdopodobieństwem równym $1/2^k$.

R2. Znormalizowana funkcja autokorelacji generowanej sekwencji liczb posiada następujący przebieg:

$$\frac{Rx(n)}{Rx(0)} = \begin{cases} 1 & \text{gdy } n = iL, \ i = 0,1,\ldots \\ A & \text{gdy } n \neq iL, \end{cases}$$

gdzie L — długość generowanego ciągu, zaś wartość $A = 0$

R3. Generowany ciąg liczbowy spełnia prawo seryjności określone w następujący sposób:

liczba serii określonej długości jest jednakowa dla każdej z liczb ze zbioru $\{0,1,\ldots,(2^k-1)\}$ występujących w generowanej sekwencji, a liczba wszystkich serii o długości $(l-1)$ jest 2^k-krotnie większa od liczby serii o długości l.

Przed przystąpieniem do analizy własności pseudolosowych sekwencji wielowartościowych podane zostanie wyrażenie pozwalające wyznaczyć funkcję autokorelacji otrzymanych ciągów:

$$R'_x(n) = 1/L \sum_{l=1}^{L} x_l x_{(l+n)} \tag{2.a}$$

gdzie x_l i $x_{(l+n)}$ są liczbami należącymi do generowanego ciągu $\{X_l\}$.

Bardziej uniwersalnym od funkcji autokorelacji parametrem statystycznym do porównywania przebiegu funkcji autokorelacji ciągów liczb należących do różnych zbiorów jest funkcja autokowariancji [3,4,11] określona zależnością:

$$R_x(n) = 1/L \sum_{l=l}^{L} x_l x_{(n+l)} - E(X)^2 = R'_x(n) - E(X)^2 \tag{2.b}$$

gdzie $E(X)^2$ jest wartością oczekiwaną generowanego ciągu.

Ponieważ na przebieg funkcji autokowariancyjnej zasadniczy wpływ posiada funkcja autokorelacji, stąd w dalszych rozważaniach analizowany będzie jedynie przebieg funkcji autokorelacji generowanego ciągu liczbowego.

Wyznaczenie przebiegu funkcji autokorelacji polegać będzie na określeniu zależności tej funkcji od funkcji autokorelacji binarnych sekwencji składowych oraz funkcji korelacji skośnej pomiędzy tymi sekwencjami. Ponieważ liczby x_l i x_{l+n} są obserowane na wyjściach równolegle pracujących generatorów sekwencji binarnych, zatem korzystając z wyrażeń (1.a) lub (1.b) można je przedstawić w następującej postaci:

$$x_l = \sum_{i=0}^{k-1} x_{1_i} 2^i,$$

$$x_{(l+n)} = \sum_{i=0}^{k-1} x_{(1+n)_i} 2^i \tag{3}$$

gdzie x_{1_i} i $x_{(1+n)_i}$ należą do zbioru $\{0,1\}$ lub do zbioru $\{-1,1\}$.

Podstawiając wyrażenie (3) do formuły (2a) otrzymamy:

$$R'_x(n) = 1/L \sum_{l=1}^{L} \left(\sum_{i=0}^{k-1} x_{1_i} 2^i \sum_{j=0}^{k-1} x_{(l+n)_j} 2^j \right) =$$

$$= \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \left(1/L \sum_{l=1}^{L} x_{l_i} x_{(l+n)_j} \right) 2^{i+j}. \tag{4}$$

Ostecznie wyrażenie określające funkcje autokorelacji sekwencji wielowartościowych generowanych przez k równolegle pracujących rejestrów liniowych może być przedstawione w następujący sposób:

$$R'_x(n) = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} 2^{i+j} R'_{ij(n)}, \tag{5}$$

gdzie $R'_{ij(n)}$ określa:

dla $i \neq j$ — wartości funkcji skrośnej (dla przesunięcia n) między ciągiem binarnym o numerze i oraz ciągiem binarnym o numerze j;

dla $i=j$ — wartości funkcji autokorelacji (dla przesunięcia n) ciągu binarnego o numerze i.

W niniejszej pracy będą analizowane trzy następujące realizacje generatorów sekwencji wielowartościowych, w których wykorzystuje się k równolegle pracujących generatorów binarnych ciągów pseudolosowych:

a) *rejestry liniowe generujące sekwencje o różnych długościach,*

b) *rejestry liniowe generujące sekwencje o jednakowej długości, lecz opisane różnymi wielomianami charakterystycznymi;*

c) *rejestry liniowe generujące jednakowe sekwencje binarne przesunięte wzajemnie o* δ_{ij} *taktów zegarowych.*

2. ANALIZA WIELOWARTOŚCIOWYCH CIĄGÓW PSEUDOLOSOWYCH GENEROWANYCH PRZEZ NIEZALEŻNE REJESTRY LINIOWE O RÓŻNYCH DŁUGOŚCIACH

W tej części pracy zostanie przeanalizowany generator sekwencji wielowartościowych wykorzystujący k równolegle pracujących, różnych rejestrów liniowych generujących binarne sekwencje o maksymalnej długości. W najogólniejszym przypadku wszystkie rejestry mogą posiadać różną długość m_j, $j \in \{1,\ldots,k\}$, a zatem generowane za pomocą tych rejestrów ciągi binarne $\{X_j\}$ posiadają różną długość l_j. Okres sekwencji wynikowej jest równy najmniejszej wspólnej wielokrotnej okresów ciągów składowych

$$L = \mathrm{NWW}(l_1, l_2, \ldots, l_k) = \frac{\prod_{j=1}^{k} l_j}{\prod_{l_s, l_r} \mathrm{NWD}(l_s, l_r)} \tag{6}$$

gdzie: NWW – najmniejsza wspólna wielokrotna, NWD – największy wspólny dzielnik.

Jeżeli każda z długości sekwencji l_i, l_j jest wzajemnie pierwsza, wówczas okres sekwencji całkowitej jest równy:

$$L = \prod_{j=1}^{k} l_j. \tag{7}$$

W pracy dokonana została analiza tego właśnie przypadku.

Na początek przebadana zostanie równomierność rozkładu pojawiania się liczb ze zbioru $\{0,1,\ldots,2^k-1\}$ w generowanej sekwencji wynikowej. Jak wiadomo w j-tej $(j=1,2,\ldots,k)$ binarnej sekwencji o maksymalnej długości liczba pojawiających się zer jest równa:

$$a_j(0) = (l_j-1)/2 = (2^{m_j}-2)/2 = 2^{m_j-1}-1,$$

zaś liczba jedynek:

$$a_j(1) = (l_j+1)/2 = 2^{m_j}/2 = 2^{m_j-1}$$

gdzie: m_j oznacza długość j-tego rejestru (liczbę przerzutników wchodzących w skład tego rejestru).

Prawdopodobieństwo pojawienia się liczby zero lub jeden na wyjściu dowolnego przerzutnika i-tego rejestru liniowego wynosi:

$$p_j(0) = \frac{2^{m_j-1}-1}{2^{m_j}-1} \simeq 1/2-\delta_j$$

$$p_j(1) = \frac{2^{m_j-1}}{2^{m_j}-1} \simeq 1/2+\delta_j \tag{8}$$

gdzie: $\delta_j = 1/(2(2^{m_j}-1))$.

Oczywisty jest fakt że, prawdopodobieństwo pojawienia się liczb zero i jeden jest tym bliższe wartości 0.5, im długość rejestru liniowego jest większa.

Rozwinięcie binarne liczby $X_i \in \{0,\ldots,2^k-1\}$, która jest elementem ciągu wynikowego posiada następującą postać:

$$X_i = x_{i_0}2^0 + x_{i_1}2^1 + \ldots + x_{i_{(k-1)}}2^{(k-1)} = \sum_{s=0}^{k-1} x_{i_m}2^s,$$

gdzie: współczynniki $x_{is} \in \{0,1,\}$, $s=0,1,\ldots,k-1$, są elementami ciągu binarnego uzyskanego z wyjścia dowolnego przerzutnika i-tego rejestru liniowego.

Przyjmijmy dalej, że dla pewnej liczby $X_i \in \{0,1,\ldots,(2^k-1)\}$, podzbiór $\{i_1, i_2,\ldots,i_m\}$ jest podzbiorem indeksów tych elementów składowych ciągów binarnych, dla których współczynniki x_i są równe 1, zaś podzbiór $\{j_1,j_2,\ldots,j_{k-m}\}$ jest podzbiorem indeksów tych elementów składowych ciągów binarnych, dla których współczynniki x_j są równe 0. Ponadto przyjmijmy także, że m_{jk}, m_{jl} oznaczają długość rejestrów generujących pseudolosowe ciągi binarne i_k oraz j_l. Wówczas X_i pojawi się podczas pełnego okresu sekwencji następującą ilość razy:

$$N_{X_i} = \prod_{k=1}^{s} 2^{(m_{ik}-1)} \prod_{l=1}^{k-s} (2^{(m_{ji}-1)}-1). \tag{9}$$

Zatem częstość pojawiania się liczby X_i w całej sekwencji MPS jest równa:

$$\gamma_{X_i} = \frac{N_{Xi}}{L} = \frac{\prod_{k=1}^{s} 2^{(m_{ik}-1)} \prod_{l=1}^{k-s} (2^{(m_{ji}-1)}-1)}{\prod_{i=0}^{k-1} (2^{(m_i-1)}-1)}. \tag{10}$$

Ponieważ interesuje nas zbadanie własności całej sekwencji, wobec tego w dalszej części pracy częstość γ_{Xi} pojawiania się liczby X_i będzie traktowana jako prawdopodobieństwo pojawienia się tej liczby.

Z zależności (9) łatwo wyznaczyć prawdobieństwo pojawiania się liczby 0, które jest równe:

$$P(0) = \prod_{j=1}^{k} (1/2-\delta_j) \simeq 1/2^k - \sum_{j=1}^{k} \delta_j \tag{11}$$

oraz prawdopodobieństwo pojawiania się liczby 2^k-1, które jest określone następującą zależnością:

$$P(2^k-1) = \prod_{j=1}^{k} (1/2+\delta_j) \simeq 1/2^k + \sum_{j=1}^{k} \delta_j. \tag{12}$$

Z wyrażeń (9)–(12) wynika, że prawdopodobieństwo pojawiania się liczby 0 w generowanym ciągu jest najmniejsze, zaś liczby (2^k-1) jest największe.

Rozkład pozostałych liczb pojawiających się w sekwencji wielowartościowej jest zależny od wag w_i, jakie zostaną przypisane składowym sekwencjom binarnym. Przykładowo, jeżeli do generacji –MPS użyjemy rejestrów liniowych opisanych wielomianami charakterystycznymi:

$$f_1(x) = x^2 \oplus x \oplus 1 \qquad f_2(x) = x^3 \oplus x \oplus 1$$

wówczas, w zależności od wag przypisanych binarnym ciągom składowym otrzymane zostaną następujące sekwencje:

$$\{X_{11}\} = \{3\ 2\ 1\ 3\ 3\ 0\ 3\ 3\ 1\ 2\ 2\ 1\ 2\ 3\ 1\ 3\ 2\ 0\}$$

bądź też

$$\{X_{12}\} = \{3\ 1\ 2\ 3\ 3\ 0\ 3\ 3\ 2\ 1\ 1\ 2\ 1\ 3\ 2\ 3\ 1\ 0\}.$$

W pracy przyjęto zasadę, że sekwencji najkrótszej przypisywana najmniejsza waga $\{2^0\}$, zaś najdłuższej – waga największa $\{2^{k-1}\}$.

Przykładowe rozkłady generowanych liczb pojawiających się w sekwencji wielowartościowej zamieszczono w tabeli 1. Analizując je nietrudno zauważyć, że w miarę równomierny rozkład prawdopodobieństwa pojawiania się liczb ze zbioru

Tabela 1

Długości rejestrów liniowych			Otrzymany rozkład generowanych liczb w sekwencji MPS							
m_1	m_2	m_3	0	1	2	3	4	5	6	7
3	4	–	21	28	24	32	–	–	–	–
3	5	–	45	60	48	64	–	–	–	–
3	7	–	189	252	192	256	–	–	–	–
2	3	5	45	90	60	120	48	96	64	128
2	3	7	189	378	252	504	192	384	256	512

$\{0,1,\ldots,(2^k-1)\}$ można otrzymać stosując rejestry o bardzo dużej długości. Wartość średnia generowanego ciągu wielowartościowego wynosi:

$$E(X) = \sum_{X_i=0}^{2^k-1} X_i \gamma_{x_i}. \tag{13}$$

Zgodnie z przyjętymi warunkami losowości, oprócz sprawdzenia równomierności rozkładu generowanych liczb pojawiających się w sekwencjach pseudolosowych, należy dokonać oceny przebiegu ich funkcji autokorelacji. Ogólna formuła określająca funkcję autokorelacji generowanych sekwencji wielowartościowych otrzymanych z wyjść równolegle pracujących rejestrów liniowych przedstawiona została wyrażeniem (2a). Dla dokładnej oceny przebiegu tej funkcji w każdym z rozpatrywanych przypadków należy wyliczyć wartość $R'_{ij}(n)$ zgodnie z zależnością (5).

Wartości funkcji $R'_{ij}(n)$, i,j$=1,2,\ldots,k$, analizowanego w tej części pracy generatora zostaną wyznaczone na podstawie poniższego twierdzenia:

Twierdzenie 1

Jeżeli długość generowanej sekwencji jest dana wyrażeniem (7), czyli długości ciągów składowych są wzajemnie pierwsze, wówczas funkcja $R'_{ij}(n)$, gdzie i oraz $j=1,2,\ldots,k$ sekwencji liczb należących do zbioru $\{0,1,\ldots,(2^k-1)\}$ posiada następującą postać

(a1)

$$R'_{ii}(n) = \begin{cases} 2^{(mi-1)}/(2^{(mi-1)} & \text{dla } n=r\ (2^{(mi-1)}) \\ 2^{(mi-2)}/(2^{(mi-1)}) & \text{dla } n\neq r\ (2^{(mi-1)}) \end{cases}$$

(b1)

$$R'_{ij}(n) = 2^{(mi-1)}2^{(mj-1)}/(2^{(mi)}-1)(2^{(mj)}-1) \quad \text{dla dowolnego } n.$$

zaś dla sekwencji liczb należących do zbioru $\{-(2^k-1), -(2^k-3), \ldots, -1,1, \ldots, (2^k-3), (2^k-1)\}$ jest określona wyrażeniem:

(a2)

$$R'_{ii}(n) = \begin{cases} 1 & \text{dla } n=r\ (2^{mi}-1) \\ -1/(2^{mi}-1) & \text{dla } n\neq r\ (2^{mi}-1) \end{cases}$$

(b2)

$$R'_{ij}(n) = 1/(2^{mi}-1)(2^{mj}-1) \quad \text{dla dowolnego } n.$$

gdzie: $r=1,2,\ldots,L/(2^{mi}-1)$.

Dowód

a) Udowodnienie słuszności wyrażeń (a1) i (a2) twierdzenia jest proste, gdyż określają one wartości funkcji autokorelacji pojedynczego pseudolosowego ciągu binarnego.

b) Dla dowodu słuszności wyrażeń (b1) i (b2), przedstawiamy wartość korelacji skrośnej zgodnie z definicją (2a):

$$R'_x(n) = 1/L \sum_{l=1}^{L} x_l x_{(l+n)} \tag{14}$$

gdzie x_l i $x_{(l+n)} \in \{0,1\}$ lub $\{-1,1\}$.

Weźmy pod uwagę wielowartościową sekwencję o długości $(2^{mi}-1)((2^{mj}-1)$ otrzymaną z wyjść przerzutników i-tego oraz j-tego rejestru. Analizowana sekwencja MPS zawiera zatem $L/(2^{mi}-1)(2^{mi}-1)$ podsekwencji obserwowanych na wyjściu dwóch przerzutników. Ponieważ w podsekwencji wielowartościowej obserwowanej na wyjściu obu rejestrów każdy element x_{is} i-tego ciągu pojawia się dokładnie raz z każdym elementem x_{jr} j-tego ciągu, stąd w każdej takiej podsekwencji istnieje:

$2^{mi-1}2^{mj-1}$	par elementów $(1_i,1_j)$ lub $(1_i,1_j)$,
$2^{mi-2}(2^{mj-1}-2)$	par elementów $(1_i,0_j)$ lub $(1_i,-1_j)$,
$(2^{mi-1}-2)2^{mj-2}$	par elementów $(0_i,1_j)$ lub $(-1_i,1_j)$,
$(2^{mi-1}-1)(2^{mj-1}-1)$	par elementów $(0_i,0_j)$ lub $(-1_i,-1_j)$.

Jeśli, jako ciąg MPS rozpatrywać będziemy ciągi dwubitowe obserwowane na wyjściach i-tego oraz j-tego rejestru, wówczas funkcja korelacji skrośnej dla takiej sekwencji wyznaczonej w okresie L może być przedstawiona w następującej postaci:

$$R'_{ij}(n) = \frac{1}{L}\frac{LR'^{*}_{ij}(n)}{(2^{mi}-1)(2^{mj}-1)} = \frac{R'^{*}_{ij}(n)}{(2^{mi}-1)(2^{mj}-1)} \tag{15}$$

gdzie: $R'^{*}_{ij}(n)$ oznacza wartość funkcji autokorelacji dla podsekwencji o długości $(2^{mi}-1)(2^{mj}-1)$.

Rozpatrując wszystkie pary elementów x_i i x_j funkcję $R'_{ij}(n)$ można wyrazić następująco:

dla przypadku (b1):

$$R'_{ij}(n) = 2^{mi-1}2^{mj-1}(1*1) + 2^{mi-2}(2^{mj-1}-2)(1*0) +$$
$$+ (2^{mi-1}-2)2^{mj-2}(0*1) + (2^{mi-1}-1)(2^{mj-1}-1)(0*0) = 2^{mi-1}2^{mj-1}$$

dla przypadku (b2):

$$R'_{ij}(n) = 2^{mi-1}2^{mj-1}(1*1) + 2^{mi-2}(2^{mj-1}-2)(1*(-1)) +$$
$$+ (2^{mi-1}-2)2^{mj-2}((-1)*1) + (2^{mi-1}-1)(2^{mj-1}-1)((-1)*(-1)) = 1$$

Podstawienie do wyrażenia (15) wartości $R'_{ij}(n)$ dla obu rozpatrywanych przypadków dowodzi twierdzenia 1.

Wnioski wynikające z twierdzenia 1 mogą być wykorzystane do obliczenia wartości funkcji autokorelacji analizowanych sekwencji wielowartościowych.

Dla ilustracji sposobu obliczania funkcji autokorelacji sekwencji wielowartościowej rozważony zostanie następujący przykład.

Niech będzie dany generator, który wytwarza ciąg wielowartościowy składający się z liczb należących do zbioru $\{-(2^k-1),-(2^k-3),\ldots,-1,1,\ldots,(2^k-3),(2^k-1)\}$.

Wartość funkcji autokorelacji dla $n=0$ można wyznaczyć z zależności:

$$R_{x(0)} = \sum_{i=0}^{k-1} 2^{2i} + \sum_{i=0}^{k-1}\sum_{\substack{j=0 \\ i\neq j}}^{k-1} 2^{i+j}(1/l_i l_j) - E(X)^2 \tag{16}$$

gdzie: $l_i=2^{m_i-1}-1$, $l_j=2^{m_j-1}-1$.

Jeżeli długość najkrótszej sekwencji składowej jest znacznie większa niż wartość 2^k, wówczas wartość funkcji autokorelacji jest zbieżna do wielkości danej wyrażeniem:

$$R_{x(0)} \simeq \sum_{i=0}^{k-1} 2^{2i} = (4^k-1)/3\,. \tag{17}$$

Wartość funkcji autokorelacji dla dowolnego przesunięcia n, gdzie $n=0$, $n=rl_{ij}$ $(i=1,2,\ldots,k)$, można wyliczyć zgodnie z zależnością:

$$R_{x(n)} = \sum_{i=0}^{k-1}(-1/l_i)2^{2i} + \sum_{i=0}^{k-1}\sum_{\substack{j=0 \\ i\neq j}}^{k-1}(1/l_i l_j)2^{i+j} - E(X)^2\,. \tag{18}$$

Wartość funkcji autokorelacji dla przesunięcia n o postaci $n=r_{i_1}l_{i_1}$, $n=r_{i_2}l_{i_2}$, ..., $n=r_{is}l_{is}$, $s\in\{1,2,\ldots,k-1\}$ może być wyznaczona wykorzystując poniższą zależność:

$$R_{x(n)} = \sum_{s\in S} 2^{2i_m} + \sum_{k\in K/S}(-1/l_{ik})2^{2i_k} +$$
$$+ \sum_{i=0}^{k-1}\sum_{\substack{j=0 \\ i\neq j}}^{k-1}(1/l_i l_j)2^{i+j} - E(X)^2\,, \tag{19}$$

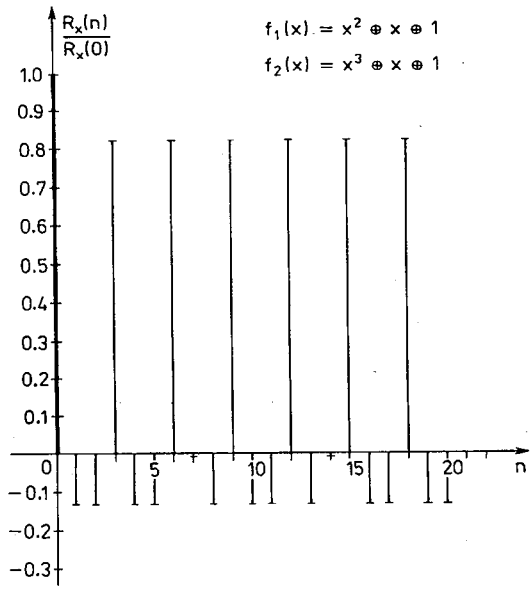
gdzie: $S=\{i_1,i_2,\ldots,i_s\}$ oznacza zbiór indeksów tych sekwencji binarnych, dla których spełniony jest warunek $n=r_{im}l_{im}$; $K=\{0,1,\ldots,k-1\}$ to zbiór indeksów wszystkich ciągów.

Z wyrażeń (16,18,19) wynika, że przebieg funkcji autokorelacji sekwencji wielowartościowych generowanych za pomocą równolegle pracujących różnych generatorów binarnych pseudolosowych ma postać funkcji grzebieniastej, bez względu na to,

Tabela 2

Przesunięcie n	Wartości znormalizowanej funkcji autokorelacji sekwencji MPS	
	Długości rejestrów liniowych	
	$m_1=2\ m_2=3\ m_3=5$	$m_1=3\ m_2=4\ m_3=5$
0	1.000	1.000
1	−0.078	−0.047
3	−0.014	−0.047
7	0.141	0.008
15	−0.078	0.157
21	0.206	−0.047
31	0.716	0.742
93	0.780	−0.047
105	−0.078	0.211
217	0.936	0.796
465	−0.078	0.945

czy liczby pojawiające się w sekwencji wielowartościowej będą należały do zbioru $\{0,1,\ldots,(2^k-1)\}$, czy też do zbioru $\{-(2^k-1),-(2^k-3),\ldots,-1,1\ldots,(2^k-3),(2^k-1)\}$. Wyniki obliczeń wartości znormalizowanej funkcji autokorelacji funkcji dla rejestrów różnej długości zamieszczono w tabeli 2. Przykładowy wykres przebiegu takiej funkcji przedstawiono na rysunku 2.



Rys. 2. Przykładowy wykres znormalizowanej funkcji autokorelacji pseudolosowej sekwencji wielowartościowej generowanej przez dwa rejestry liniowe różnej długości

Obecnie dokonana zostanie analiza spełnienia trzeciego postulatu losowości — zachowania warunku seryjności, przy założeniu, że długości składowych sekwencji binarnych są wzajemnie pierwsze.

Na wstępie należy poczynić pewne uwagi dotyczące pojawiania się serii liczb, a wynikające z zasady działania rejestrów liniowych. Nietrudno zauważyć, że maksymalna długość serii liczb ze zbioru $\{0,1,\ldots,(2^k-1)\}$ nie może przekraczać długości najdłuższej serii zer, lub jedynek występujących w ciągu binarnym o najkrótszym okresie. Dla przykładu, w ciągu utworzonym z trzech sekwencji binarnych otrzymywanych z trzech rejestrów liniowych o długościach $m_1=3$, $m_2=4$, $m_3=5$, najdłuższa seria liczb nieparzystych może być równa co najwyżej 3 (maksymalna liczba jedynek w sekwencji otrzymanej z rejestru o długości 3), zaś najdłuższa seria liczb parzystych — co najwyżej 2, czyli m_1-1.

Rozkład serii liczb pojawiających się w sekwencji wielowartościowej jest zależny, podobnie jak i rozkład generowanych liczb, od wag w_i, jakie zostaną przypisane składowym ciągom binarnym. Eksperymentalnie otrzymane wyniki badania seryjności umieszczono w tabelach 3a i 3b. Obecnie zostanie przeprowadzona dyskusja tych wyników.

Tabela 3a

Rozkład serii liczb w sekswencji MPS						
Liczby	Długość serii					Długość rejestrów
	1	2	3	4	5	
0	3	0	0	0	0	$m_1 = 2$
1	4	1	0	0	0	$m_2 = 3$
2	4	0	0	0	0	
3	4	2	0	0	0	
0	15	3	0	0	0	$m_1 = 3$
1	17	4	1	0	0	$m_2 = 4$
2	16	4	0	0	0	
3	18	4	2	0	0	
0	66	15	3	0	0	$m_1 = 4$
1	70	17	4	1	0	$m_2 = 5$
2	68	16	4	0	0	
3	72	18	4	2	0	
0	276	66	15	3	0	$m_1 = 5$
1	284	70	17	4	1	$m_2 = 6$
2	280	68	16	4	0	
3	288	72	18	4	2	
0	31	7	0	0	0	$m_1 = 3$
1	35	8	3	0	0	$m_2 = 5$
2	32	8	0	0	0	
3	36	8	4	0	0	
0	270	63	15	0	0	$m_1 = 2$
1	286	71	16	7	0	$m_2 = 3$
2	272	64	16	0	0	
3	288	72	16	8	0	

Tabela 3b

Rozkład serii liczb w sekswencji MPS				
Liczby	Długość serii			Długości rejestrów
	1	2	3	
0	273	21	0	$m_1 = 3$
1	339	36	3	$m_2 = 4$
2	304	28	0	$m_3 = 5$
3	374	44	6	
4	288	24	0	
5	356	40	4	
6	320	32	0	
7	392	48	8	
0	1137	93	0	$m_1 = 3$
1	1407	156	15	$m_2 = 4$
2	1264	124	0	$m_3 = 7$
3	1550	188	30	
4	1152	96	0	
5	1424	160	16	
6	1280	128	0	
7	1568	192	32	
0	18558	652	0	$m_1 = 3$
1	23927	1217	45	$m_2 = 4$
2	20965	869	0	$m_3 = 5$
3	26874	1559	91	$m_4 = 7$
4	19680	774	0	
5	25308	1368	60	
6	22208	992	0	
7	28408	1744	120	
8	18835	672	0	
9	24262	1248	48	
10	21266	897	0	
11	27251	1607	96	
12	19968	768	0	
13	25664	1408	64	
14	22528	1024	0	
15	28800	1792	128	

Przyjęto następujące oznaczenia:

m_0 — oznacza długość rejestru liniowego generującego najkrótszą sekwencję;

m_i — oznacza długość i-tego rejestru liniowego;

$l_{m0}(i)$ — określa liczbę najdłuższych serii liczby i;

$l_{(m0-j)}(i)$ — liczba serii liczby i o długości $m_0 - j$, gdzie $j = 1,\ldots,(m_0 - 1)$.

Na podstawie otrzymanych wyników zamieszczonych w tablicach 3a i 3b, wykorzystując zdefiniowane powyżej oznaczenia, sformułowane zostały następujące wyrażenia pozwalające określić liczność serii liczby $(2^k - 1)$.

$$
\begin{aligned}
l_{m0}(2^k-1) &= \prod_{i=1}^{k-1} 2^{(mi-m0)},\\
l_{(m0-1)}(2^k-1) &= (2^k-2)\, l_{m0}(2^k-1)\\
l_{(m0-2)}(2^k-1) &= (2^k-1)^2\, l_{m0}(2^k-1)\\
l_{(m0-3)}(2^k-1) &= (2^k-1)^2\, 2^{m0} l_{m0}(2^k-1)\\
&\cdot\\
&\cdot\\
&\cdot\\
l_{(m0-j)}(2^k-1) &= (2^k-1)^2\, 2^{(j-2)m0} l_{m0}(2^k-1)
\end{aligned}
$$

Wyrażenie określające liczność serii maksymalnej długości liczb nieparzystych posiada następującą postać:

$$l_{m0}(2^k-1-2p) = \prod_{r\in I_1} 2^{(mir-m0)} \prod_{m\in I_0} (2^{(nis-n0)}-1)$$

gdzie: I_1 — podzbiór indeksów tych składników wchodzących w skład binarnego rozwinięcia generowanej liczby, które są równe jeden; I_0 — podzbiór indeksów tych składników wchodzących w skład binarnego rozwinięcia generowanej liczby, które są równe zero; $p=0,1,\ldots,2^{(k-1)}-1$.

Liczność serii liczb parzystych o długości (n_0-1) jest określona następującym wyrażeniem:

$$l_{(m0-1)}(2^k-2-2p) = \prod_{r\in I_1} 2^{(mir-m0+1)} \prod_{m\in I_0} (2^{(mis-m0+1)}-1)$$

gdzie I_1, I_0 oraz p są określone jak powyżej.

Z przedstawionych wyników i podanych zależności wynika następujący wniosek:

Wniosek 1

Aby spełniony był warunek równomierności rozkładu generowanych liczb w sekwencji wielowartościowej oraz warunek seryjności, koniecznym jest, by różnice długości użytych rejestrów liniowych, generujących składowe sekwencje binarne, były jak najmniejsze.

Najkorzystniejszym zatem rozwiązaniem byłoby wykorzystanie rejestrów liniowych generujących sekwencje pseudolosowe o jednakowej długości, stąd w dalszej części niniejszej pracy dokonano analizy wielowartościowych ciągów pseudolosowych otrzymywanych z wyjść równolegle pracujących rejestrów liniowych jednakowej długości opisanych różnymi wielomianami charakterystycznymi.

3. ANALIZA PSEUDOLOSOWYCH SEKWENCJI WIELOWARTOŚCIOWYCH GENEROWANYCH PRZEZ RÓWNOLEGLE PRACUJĄCE, RÓŻNE REJESTRY LINIOWE JEDNAKOWEJ DŁUGOŚCI

W tej części pracy zakładamy, że generator wytwarzający sekwencję MPS jest zbudowany z k równolegle pracujących rejestrów liniowych o jednakowej długości. Poszczególne rejestry są opisane różnymi wielomianami charakterystycznymi tego samego stopnia nad ciałem Galois GF(2). Z przyjętego założenia wynika oczywisty fakt, że długość każdej generowanej sekwencji wynosi $L=2^m-1$, gdzie m — długość rejestru. Oczywistym jest, że w sekwencji wielowartościowej, wygenerowanej zgodnie z przyjętą zasadą, każda podsekwencja generowanego składowego ciągu binarnego o długości m pojawia się tylko jeden raz w całym okresie sekwencji. Ma to istotny wpływ zarówno na rozkład liczb pojawiających się w sekwencji MPS, jak i na przebieg funkcji autokorelacji.

Zanim zostanie przedstawiona dokładniejsza analiza pracy tak zbudowanego generatora sekwencji MPS, rozważony zostanie następujący przykład. Niech będą dane dwie sekwencje binarne o długości $L=15$, generowane przez rejestry liniowe opisane wielomianami:

$$f_1(x) = x^4 \oplus x^3 \oplus 1 \qquad f_2(x) = x^4 \oplus x \oplus 1 .$$

Jeżeli przyjąć założenie, że wektory początkowe obu rejestrów są jednakowe i równe $<1000>$, wówczas na wyjściach obu rejestrów wygenerowane zostaną następujące ciągi binarne:

$$\{X_{1\,1}\} = \{1\,1\,1\,1\,0\,1\,0\,1\,1\,0\,0\,1\,0\,0\,0\}$$
$$\{X_{1\,2}\} = \{1\,0\,0\,1\,1\,0\,1\,0\,1\,1\,1\,1\,0\,0\,0\}.$$

Tabela 4

Przesunięcie fazy ciągów binarnych	Rozkład liczb w sekswencji MPS			
δ_{12}	0	1	2	3
0	3	4	4	4
1	4	3	3	5
2	3	4	4	4
3	5	2	2	6
4	4	3	3	5
5	2	5	5	3
6	3	4	4	4
7	4	3	3	5
8	5	2	2	6
9	2	5	5	3
10	4	3	3	5
11	2	5	5	3
12	2	5	5	3
13	3	4	4	4
14	3	4	4	4

Łatwo zaobserwować, że w wynikowej sekwencji wielowartościowej, w zależności od wzajemnego przesunięcia pomiędzy ciągami składowymi, (przypisując ciągom odpowiednio wagi 2^0 i 2^1) uzyskane zostaną różne rozkłady generowanych liczb należących do zbioru $\{0,1,2,3\}$. Wszystkie możliwe rozkłady uzyskane dla różnych przesunięć umieszczone zostały w tabeli 4. Z tabeli tej wynika, że równomierny rozkład liczb w sekwencji MPS uzyskany zostanie dla przesunięć 0,2,6,13,14. Przykład powyższy wykazuje, że koniecznym staje się określenie reguły doboru przesunięć dla sekwencji binarnych tak, aby prawdopodobieństwo otrzymania każdej liczby ze zbioru $\{0,\ldots,2^k-1\}$ było jednakowe.

Jedynym, znanym dotychczas autorom, sposobem prowadzącym do doboru prawidłowych przesunięć pomiędzy ciągami jest działanie empiryczne, polegające na kolejnym przeglądaniu wynikowej sekwencji MPS, tak aby spełniała ona pierwszy postulat losowości.

Tabela 5

Rozkład liczb pojawiających się w sekwencji MPS								
Liczba:	0	1	2	3	4	5	6	7
Przypadek I	4	5	3	3	3	3	5	5
Przypadek II	2	5	5	3	5	3	3	5
Przypadek III	3	4	4	4	4	4	4	4

W tabeli 5 umieszczono przykładowe rozkłady liczb pojawiających się w ciągach wielowartościowych dla różnych przesunięć składowych ciągów binarnych. W wyniku badań sekwencji wielowartościowych otrzymywanych z zastosowaniem omawianej reguły sformułowano następujący wniosek, na podstawie którego określić będzie można zasadę wyboru pracujących równolegle składowych rejestrów liniowych.

Wniosek 2

Niezbędnym i wystarczającym warunkiem, aby w generowanej sekwencji wielowartościowej, liczby ze zbioru $\{0,\ldots,2^k-1\}$ miały rozkład równomierny jest spełnienie następującej zależności: dla każdego podzbioru ciągów binarnych $S_i, S_{i_2}, \ldots, S_{i_1}$, $l < k$

$$\mathrm{WH}(S_1 \oplus S_{i_2} \oplus \ldots \oplus S_{i_1}) = \mathrm{WH}(S_{i_1} \oplus S_{i_2} \oplus \ldots \oplus S_{i_1}) + 1 \qquad (20)$$

gdzie: WH — **waga Hamminga sekwencji binarnej,** S_{i_1} — **dopełnienie sekwencji** S_{i_1}, $\oplus$ — **suma mod (2).**

Wyrażenie (2a), które określa funkcję autokorelacji, jest również słuszne dla rozpatrywanego w tej części pracy rodzaju generatora sekwencji MPS. Jednakże jest tutaj znacznie trudniej wyznaczyć zwartą formułę określającą wartość funkcji korelacji skrośnej $R'_{ij}(n)$.

Dla pseudolosowej sekwencji wielowartościowej, której elementy należą do zbioru $\{0,1,...2^k - 1\}$ wartość funkcji korelacji skrośnej wynosi:

$$R'_{ij}(n) = 1/LN_{bp} <a_i, a_j> \tag{21a}$$

gdzie: $N_{bp}<a_i, a_j>$ oznacza liczbę par elementów i-tej oraz j-tej sekwencji przesuniętych względem siebie o n taktów zegarowych.

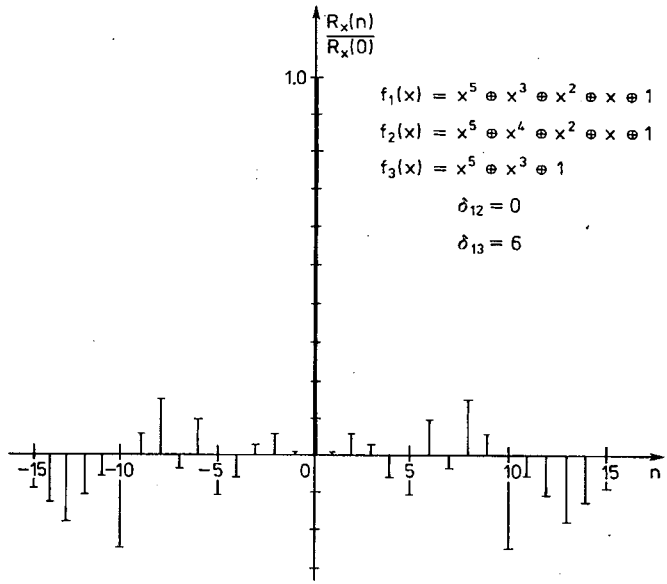
Dla pseudolosowej sekwencji wielowartościowej, której elementy należą do zbioru $\{-(2^k-1),-(2^k-3),...,-1,1,...,(2^k-3),(2^k-1)\}$ wartość funkcji korelacji wzajemnej wynosi z kolei:

$$R'_{ij}(n) = 1/L\,[(N_{bp} <0_i, 0_j> + N_{bp} <1_i, 1_j>) + - (N_{bp} <1_i, 0_j> + N_{bp} < 0_i, 1_j>)]. \tag{21b}$$

Tabela 6

$k = 7$	L = 127		Wielomiany charakterystyczne		
	Liczba		$x^7 \oplus x \oplus 1$ $x^7 \oplus x \oplus 1$	$x^7 \oplus x^3 \oplus x^2 \oplus x \oplus 1$ $x^7 \oplus x^6 \oplus 1$	$x^7 \oplus x \oplus 1$ $x^7 \oplus x^6 \oplus 1$
$R'_x(n)$	0	1	Liczba punktów o wartości funkcji korelacji skrośnej $R_x(n)$		
−41/127	43	84	–	1	–
−21/127	53	74	–	–	7
−17/127	55	72	28	14	14
−57/127	57	70	–	–	14
−9/127	59	68	–	28	7
−5/127	61	66	–	–	21
−1/127	63	64	63	35	14
3/127	65	62	–	–	7
7/127	67	60	–	28	21
11/127	69	58	–	–	8
15/127	71	56	36	14	7
19/127	73	54	–	–	7
23/127	75	52	–	7	–

Ponieważ dla większej liczby składowych ciągów binarnych obowiązuje takie samo prawo, jak dla dwóch ciągów, zatem wartość funkcji $R'_{ij}(n)$ zmienia się w zależności od przesunięcia n. Przeprowadzone badania funkcji korelacji wzajemnej sekwencji uzyskanych z rejestrów o jednakowej długości pokazują, że funkcja ta posiada postać funkcji grzebieniastej, tak jak to pokazano w tabeli 6, a zatem również funkcja autokorelacji ciągu wielowartościowego posiada postać funkcji grzebieniastej. Na rys. 3 zaprezentowano przebieg funkcji autokorelacji sekwencji wielowartościowej otrzymanej z trzech różnych ciągów binarnych o długości $L=31$ dla wzajemnych przesunięć fazowych wybranych losowo.



Rys. 3. Przykładowy wykres znormalizowanej funkcji autokorelacji pseudolosowej sekwencji wielowartościowej generowanej przez dwa rejestry liniowe jednakowej długości opisane różnymi wielomianami prymitywnymi

Z przedstawionych rozważań wynika, że dla pseudolosowych sekwencji wielowartościowych, otrzymywanych z równolegle pracujących rejestrów liniowych o jednakowej długości trudno jest sformalizować wyrażenie na wyznaczanie wartości funkcji autokorelacji, koniecznym zatem jest obliczanie jej metodą *„krok po kroku"*.

Nie znaleziono również zasady określającej warunek konieczny na to, by był spełniony postulat seryjności w sekwencjach wielowartościowych uzyskanych z generatorów MPS utworzonych z równolegle pracujących różnych rejestrów liniowych o tej samej długości. Przykładowe rozkłady serii liczb pojawiających się w sekwencji MPS dla różnych przesunięć pomiędzy binarnymi ciągami składowymi podano w tabeli 7.

Przedstawione w tej części pracy rozważania i wyniki prowadzą jednak do stwierdzenia, że parametry sekwencji wielowartościowych otrzymywanych z generatorów zbudowanych z k różnych rejestrów liniowych, odbiegają znacznie z punktu widzenia ich własności losowych, od parametrów sekwencji w pełni losowych, tzn. sekwencji spełniających postulaty losowości sprecyzowane przez Golomba i Daviesa, a dotyczących równomierności rozkładu generowanych liczb, przebiegu funkcji autokorelacji oraz zasady zachowania seryjności.

Z dotychczasowej analizy wynikają jednakże pewne wnioski dotyczące sposobu generacji pseudolosowych ciągów liczbowych nad GF(2^k). Z rozważań zawartych w pierwszej części artykułu wynika, że dla spełnienia postulatów losowości, koniecznym jest stosowanie rejestrów o jednakowej długości. Z kolei na podstawie drugiej